LEADERSHIP AND POLICY IN SCHOOLS
https://doi.org/10.1080/15700763.2025.2539806

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Instructional Leadership Moderating the Impact of (In)Congruency Between Peer and Individual Student SES on Achievement

Mehmet Şükrü Bellibaş[a,b], Burak Aydın[b,c], Alex Bowers[d], Laura Perry[e], and Marcus Pietsch[b]

[a]Institute of Leadership in Higher Education and College of Public Policy, University of Sharjah, Sharjah, UAE; [b]Institute of Educational Sciences, School of Education, Leuphana University, Lüneburg, Germany; [c]Department of Educational Sciences, Ege University, Izmir, Turkey; [d]Teachers College, Columbia University, New York, NY, USA; [e]School of Education, College of Health and Education, Murdoch University, Murdoch, Australia

ABSTRACT

The present study aims to investigate how school segregation, as well as the (in) congruence between the school and individual SES, can explain the variation in student achievement. Additionally, it examines the role of instructional leadership in mitigating this association. Using international large-scale assessments (PISA-TALIS link data) from seven countries – Australia, Colombia, the Czech Republic, Denmark, Georgia, Malta, and Turkey – we applied several multilevel polynomial regressions with response surface analyses. The results showed that both individual SES and school segregation have a profound impact on student achievement, with varying results across countries. Second, we found differential school composition effects, with the school composition effect strongest for low SES students in high SES schools. Third, our results do not support congruence theory, but they do somewhat favor (in)congruence theory. Finally, strong leadership magnifies benefits for low-SES students in high-SES schools and for all students at low-SES schools. Implications for policy, practice, and further research are discussed.

Introduction

School socioeconomic segregation – the clustering of students by socioeconomic background into separate schools – has been increasing in many countries (Bonal & Bellei, 2019). The causes, effects, and implications of this trend, and of school composition effects more generally, have been receiving increased attention from researchers and policymakers in efforts to mitigate negation consequences for educational equity and effectiveness (Crosnoe, 2009; Darling-Hammond, 2014; George & Darling-Hammond, 2019; Paloyo, 2020; Perry et al., 2022; Schindler Rangvid, 2003). However, policy solutions for mitigating negative effects are not straightforward because several questions remain unanswered or lack definitive answers. These include the extent to which school segregation and composition negatively impact students, whether students from different socioeconomic groups are impacted differentially, and whether and how school-level processes and system-level structures mediate effects (Caldas & Bankston, 1997; Gümüş et al., 2022; Perry et al., 2022; Sirin, 2005).

Research indicated that school composition effects are generally positive, with increases in average school SES associated with increases in student achievement for all students, regardless of their individual SES. In addition, research provided evidence that school leadership, particularly instructional leadership, is important for student learning (Bellibaş et al., 2025; R. Goddard et al., 2015; Mitchell et al., 2015; Sebastian et al., 2019), and it could also play a role in mitigating the negative consequences of SES on student learning (Gümüş et al., 2022). However, some questions remain (Gümüş et al., 2022; Gustafsson et al., 2018; Langenkamp & Carbonaro, 2018; van Ewijk & Sleegers, 2010). First, are these positive effects the same for all students, or do students from particular socioeconomic backgrounds benefit more? Second, can school leadership moderates the effects of school composition? Third, is it possible that student achievement

CONTACT Marcus Pietsch pietsch@leuphana.de Institute of Educational Sciences, School of Education, Leuphana University, Lüneburg 21335, Germany

is greatest when student and school SES are similar, as suggested by convergence theory, or conversely, greatest when they diverge, as suggested by (in)convergence theory? While (in)congruence theory has not been extensively examined in the school composition literature, research from social and educational psychology suggests that it may play a role (Boucher et al., 2022; Crosnoe, 2009; Gutiérrez, 2022; Paloyo, 2020). Understanding whether it does is important because it would suggest very different policy solutions, and perhaps even different policy agendas.

This study aims to examine these questions. We integrate secondary data from two high-quality, large-scale international assessments for seven countries: Australia, Colombia, the Czech Republic, Denmark, Georgia, Malta, and Turkey. We applied a polynomial regression with response surface analysis (Shanock et al., 2010) to examine the joint combination of two individual and school-level SES in relation to students' reading, math, and science achievements. Additionally, building on Zyphur et al. (2016) and Nestler et al. (2019), we utilized a multilevel polynomial regression model to compute response surface parameters separately for each country to visualize the effect of (in)congruence between individual and school SES on student achievement. Finally, we estimated the same model with(out) moderating the role of IL to investigate the contribution of school leadership in mitigating disparities in and across schools derived from differences in SES.

Literature Review

Individual SES

Educational researchers have long been interested in SES and its measures because it has dramatic consequences for key educational outcomes (Gümüş et al., 2022). "SES corresponds to a complex bundle of social and economic factors that are generally but imperfectly correlated" (Farah, 2017, p. 56). This definition links SES to material wealth, which is accompanied by noneconomic factors such as occupational prestige, social network, and education. Besides, drawing upon the cultural capital framework, researchers suggested home educational resources, such as books, computers, and the internet are relevant elements for gaining an idea about students' SES status (Yang & Gustafsson, 2004). In sum, key variables used to estimate SES included parental education and occupation status, family income, and the number of certain types of possessions available in the student's home.

OECD has long been using the "Index of Economic, Social, and Cultural Status (ESCS)" as an estimate of students' SES. Considering this measurement, a socioeconomically advantaged student would have parents with higher levels of education and occupation – most likely with a degree of post-secondary education and semi-skilled occupation. Moreover, on average, high-SES students report having more books and other educational possessions at home, such as educational software, computers, and a quiet place to study (OECD, 2016). Marks and O'Connell (2021) summarized several factors compromising the quality of ESCS, such as constructs not being equal across countries, poor model-to-data consistency on various subscales, poor cultural comparability, and poor fit indices of home possessions items.

The impact of SES on student achievement, learning, and development has been well-established in the educational literature. This literature is often traced back to the prominent Coleman Report (Coleman et al., 1966), which found a strong influence of family background and home environment on student outcomes. A cross-cultural comparison conducted by the OECD (2004) revealed that educational inequalities are a pervasive global concern, yet with large variations in the extent of inequality observed among different countries. To illustrate, the gap between high and low SES students is 56 score points in Iceland, while it is 110 points in Belgium and Hungary. Additionally, while the link between SES and student achievement is strongest among students with lower SES levels for some countries (e.g., Italy and Hungary), it is larger with students at the higher SES level for other countries (e.g., Australia, USA, and Turkey) (OECD, 2004).

Peer Effect

While the effect of individual SES on student learning and development has been extensively examined, research inquiring into the impact of school-level SES (SSES) – often referred to as peer effect – on pupils has remained scarce (Gümüş et al., 2022). More research on this is indispensable because it encompasses meanings, interpretations as well as consequences distinct from those associated with individual SES, which are not fully understood (Caldas & Bankston, 1997; Gümüş et al., 2022; Sirin, 2005). Yang and Gustafsson (2004) indicated that SSES, which refers to "the community's sociocultural and economic environment" highlights the school differences regarding community resources and is used as an indication of school quality (p.48). This definition highlights institutional or contextual factors, associating low-SES schools with, for example, lower-quality teachers, less involved parents, lower expectations, and less rigorous curricula (Gutiérrez, 2022; Yang & Gustafsson, 2004).

Another perspective emphasizes the "contagion" mechanisms, which are often translated as peer effect/ or compositional effect (Thrupp et al., 2002) and exhibit their influence as a result of peer interactions. Disadvantaged schools are often associated with larger numbers of students who have lower aspirations, achievement, and attainment, as well as poorer study habits and more disruptive classroom behavior, which collectively has a negative impact on students at these schools (Armor et al., 2018). The implication for peer effect extends beyond the mere interaction among peers but also involves various instructional and school organizational and managerial processes as well as teachers' decisions regarding instructional processes (Gutiérrez, 2022). This suggests that students would be affected by their peers because the dominant student group will be the reference and others will be treated and served in reference to those students (Wilkinson, 2002). However, as Harris (2010) argues, school-level SES impacts achievement via several mechanisms, making it powerful and complex.

Often traced back to the famous report of Coleman et al. (1966), researchers have reported that besides students' individual SES, the SES composition of a school has some consequences for students. A meta-analysis reported that a one standard deviation increase in average peer SES was associated with a .32 standard deviation increase in student achievement after students' individual backgrounds were taken into account, with not much variation across countries (van Ewijk & Sleegers, 2010). In a recent study, utilizing the PISA 2018 dataset, Gümüş et al. (2022) found a significant effect of SSES on student achievement in all domains (reading, math, and science), favoring students attending a school with affluent peers. Similarly, a study based on the 2018 Australian PISA sample concluded that the effect of SSES is stronger than that of individual SES for all student performance quintiles (Perry et al., 2022). In another meta-analysis, using International Large Scale Achievement Data, Holzberger et al. (2020) showed that SSES is the strongest predictor of student outcomes ($r = 0.30$), even when controlled for other school-level variables, and that these effects do not differ between subjects, i.e., student achievement in both math and science.

More recently, Tan et al. (2023), in another meta-analysis based on 480 effect sizes from 97 studies, found that SSES, on the one hand, is highly related to student achievement ($r = 0.50$) but that both the strength and magnitude of this association are dependent on specific school process; "The associations of school SES with school leadership and teacher quality suggest that higher-SES students may have better learning outcomes because competent school leaders and teachers cultivate a positive school climate" (Tan et al., 2025, p. 291). An OECD report, however, indicated that SSES is not significantly related to student achievement in reading literacy in some countries (e.g., Denmark, Korea, Finland, and Iceland), calling for a meticulous investigation of the issues at the cross-cultural level (Perry, 2007).

Finally, evidence is mixed when comparing the impact of individual and school-level SES. While some studies suggest that the effect of SSES on student learning is only slightly less than that of individual SES (Caldas & Bankston, 1997). In contrast, other studies have found that the SES of peer groups has a greater impact on academic achievement than individual SES (Gustafsson et al., 2018).

Assumptions About Mechanisms of Action

Besides the discussion of which SES has the greater effect on students, educational researchers have investigated how the mix of students with different SES in more homogeneous or heterogeneous schools influences their achievement. In the present study, we introduce the concept of congruence (or incongruence) to explain the mechanisms underlying the relative role of SSES to SES in predicting student achievement, based on recent work (e.g., Barranti et al., 2017; Lai et al., 2022; Nestler et al., 2019). Congruence suggests that students tend to perform better when their background aligns with that of their peers or school environment. For example, when low-SES students are in low-SES schools or high-SES students are in high-SES schools. In contrast, incongruence (where a student's SES differs significantly from the school's average SES) can create social and psychological mismatches that hinder achievement (see Boucher et al., 2022; Crosnoe, 2009; Gutiérrez, 2022; Paloyo, 2020).

Despite a well-established argument that student peers have a considerable impact on student outcomes, opposing perspectives and research findings regarding the peer effect render it challenging to ascertain policy recommendations (Paloyo, 2020). While the "boutique" perspective, for instance, suggests that students perform better when attending a school filled with peers of the same ability level, the "rainbow" perspective argues that heterogeneity in the school population may benefit all students (Gutiérrez, 2022). The two perspectives are closely related to the (in)congruency effect (Barranti et al., 2017). Accordingly, the congruence between individual and peer SES (boutique perspective) would yield better student outcomes; whereas, the rainbow perspective suggests that students perform better when attending a school where there is high incongruence (mismatch) between their individual SES and that of their peers (Gutiérrez, 2022). A "little frog big pond" perspective, however, emphasizes the adverse effects that may arise from evaluating students based on their relative position within the school, pointing to the disadvantage that a low SES student who attends a school with predominantly high SES students might face because of being compared with other students in a more competitive context (Crosnoe, 2009), particularly in the case of large segregation among high and low achieving students with a lower population of moderate achieving ones (Bolletta, 2021). This idea is closely related to the concept of group-based "contagion" (Harris, 2010), which posits that a student follows the lead of other students, especially those perceived as members of their own group. According to this approach, there may be multiple groups in schools that directly influence each other, while also influencing school resources and processes (e.g., teachers and instruction), and thus have indirect effects on other peers within the school. Similarly, a more nuanced finding regarding the group-based peer effect contradicts the mainstream idea of the linear effect assumption that students are influenced by the average outcome of their peers; rather, for the positive behaviors (e.g., GPA and study effort) the more active agents among their peers mattered the most while for the negative behaviors (troublemaking), the less active individuals had a larger influence (Boucher et al., 2022).

Policy recommendations are often based on research that focuses on the peer effect on the educational experience and outcome of low-SES students. One problem in the present literature is that insufficient attention is given to high-SES students in low-SES schools. Using a longitudinal data set that tracks students' transition from elementary to middle school, Langenkamp and Carbonaro (2018) found that non-disadvantaged students might suffer more than advantaged students from high-poverty elementary schools. Using Free and Reduced Lunch (FRL) as an indication of SES, they reported "Non-FRL students are more sensitive to percentage FRL in [elementary schools]: the 20% versus 80% FRL gap for non-FRL students is a difference in gains of 13.9 points, whereas it is only 4.2 points for FRL students. Indeed, non-FRL students in high-poverty elementary schools have the lowest growth rate among the six groups". Moreover, a recent review (Paloyo, 2020) found that both measures and results are highly context-specific, raising doubts about the generalizability and transferability of empirical findings and policy solutions, pointing out the need for more data and empirical investigation of the issue.

Cross-Cultural Differences

The research evidence regarding segregation and its impact is mixed and likely to vary across countries, research publications, time, status, and research methodologies (Boman, 2023; Thrupp et al., 2002). It is found that the majority of developing countries are suffering from school segregation, leading to severe SES differences among schools, and as a result, compared to developed countries, students in these countries are more likely to carry the burden associated with their SES (Gustafsson et al., 2018). In the USA, the busing policy supporting the incongruence of peer and individual SES for desegregation purposes has led to a reduction in educational disparities. Students who are transported to "better" schools through busing show enhanced academic performance, without suffering a decline, compared to those who are already assigned to high-quality schools, providing supporting evidence for the half-century effort of U.S. educational policy to desegregate schools (George & Darling-Hammond, 2019). This is consistent with the results from Danish schools, where the positive effect of being among high-achieving students is larger for low-performing disadvantaged students than high-performing ones

(Schindler Rangvid, 2003). On the other hand, there is also opposing evidence that elicits the need for further investigation of the issue. For example, the work of Crosnoe (2009) supported the "frog pond" effect, showing evidence that disadvantaged students in higher SSES schools in the USA experience reduced coursework levels and increased psychosocial complications relative to similar students in lower SES schools, which is likely to hinder the potential achievement benefits associated with positive peer effect.

Instructional Leadership (IL) and Student Achievement

The notion of IL has often been traced back to the 70s, indicating that principals in these schools were instructional leaders who spent time to improve teaching and learning in their schools (Edmonds, 1979). They hire high-quality teachers, protect their instructional time, create a positive school environment, observe teaching, provide feedback (Guthrie & Willower, 1973), and use teaching data to increase teacher performance (Zechman, 1977). Their focus is on improving teaching and promoting student learning (Edmonds, 1979). Synthesizing the early research, Hallinger and Murphy (1985) developed a more structural and extensive definition of IL, which has received considerable recognition from educational researchers worldwide. This suggests that instructional leaders develop academic goals and communicate these goals to the staff; they coordinate curriculum, monitor student academic achievement, and observe teaching. Such leaders protect instructional time, promote professional learning, reward high-quality teaching and student learning, and maintain visibility in schools (Hallinger, 2005).

Since the introduction of IL to educational literature, researchers have been extremely concerned with the association between this leadership type and student achievement (R. Goddard et al., 2015; Larsen, 1987; Mitchell et al., 2015; O'Donnell & White, 2005; Sebastian et al., 2019). With a wide variety of analytic strategies and approaches to the investigation of the association between leadership and student achievement, they have consistently reported that principals' enactment of IL is essential for school improvement and enhanced student learning (O'Donnell & White, 2005; Robinson et al., 2008; Sebastian et al., 2019). For example, consistent with school effectiveness research, early researchers often conducted a comparative analysis between the frequency of IL in schools that exhibited high academic achievement and those that demonstrated low academic performance and reported that in schools where student achievement was higher, the principals were also ranked significantly higher in terms of their involvement in IL (Larsen, 1987). Another group of researchers has examined the direct associations. For instance, after controlling for students' socio-economic status in a regression model, O'Donnell and White (2005) found a direct relationship between the frequency of principals' involvement in practices to promote a learning climate in the school and student reading and math achievement. Similarly, Sebastian et al. (2019) found a direct link between principals' IL profiles and student learning. Another line of research examined the indirect association between IL and student achievement using mediated-effects models and found a myriad of factors mediating this relationship including but not limited to academic press (Mitchell et al., 2015), teacher collaboration and collective efficacy (R. Goddard et al., 2015) and conditions for high-quality teaching and learning (Day & Sammons, 2020). Finally, Robinson et al. (2008) compared the effect of IL on pupil learning with that of other leadership types in an extensive meta-analysis and concluded that "the average effect of IL on student outcomes was three to four times that of transformational leadership" (p. 635). However, recent studies, drawing on International Large-Scale Assessment (ILSA) data, show that both the understanding and measurement (Eryilmaz & Sandoval Hernandez, 2021) and the effects of IL on student achievement significantly vary between contexts (Pietsch et al., 2023).

The Role of IL in Educational Disparities

The issue of educational disparities underlies the social engineering of schooling (Wilkinson, 2002). State-level educational policies across the globe can be effective to the extent that they reduce the negative consequences of SES differences between students and schools (OECD, 2004). The policies aimed at reducing the student performance gap due to individual or institutional socio-economical and cultural differences have been diverse and varied, and so are the underlying causes contributing to this gap (Jeynes, 2015). The research found variations in countries' ability to address this issue. Some countries, mostly developed ones, were more effective than others in mitigating the impact of individual and peer SES on student learning through compensatory measures, such as fostering a safe and orderly school climate and prioritizing high-quality teaching and student academic success (Gustafsson et al., 2018).

These systems focused on school leadership, with a particular emphasis on teaching and learning, which was positively related to improved schools and student learning, especially in disadvantaged contexts (Demie, 2023; R. D. Goddard et al., 2017). Teddlie et al. (2002) found that in effective low-SES schools, principals had stronger control of instruction and a task orientation, whereas in effective high-SES schools, the control of instruction was low to moderate, and moderate concerning task orientation (Teddlie & Stringfield, 1985). Consistently, Darling-Hammond (2014) has emphasized key aspects of interventions that can help reduce the achievement gap among students in low SES schools: supporting students with high-quality teaching and teachers, developing a high-quality curriculum, and establishing a positive learning environment conducive to teaching and learning. R. D. Goddard et al. (2017) found empirical evidence that IL reduces inequalities in student outcomes. In a more nuanced empirical study, Gümüş et al. (2022) derived evidence from the 2018 PISA dataset that supports the IL role of school principals, reporting that IL intervenes in the mechanism of SSES and student learning and weakens the strength of the connection.

Present Study

Specifically, it examines whether these positive effects are uniform across all students or disproportionately benefit those from particular socioeconomic backgrounds. It also investigates whether instructional leadership (IL) can moderate the influence of school SES composition on student outcomes. Finally, it explores whether student achievement is maximized when individual and school SES levels align, in line with congruence theory, or when they diverge, as suggested by incongruence perspectives.

Our review of the literature showed that despite a general tendency of empirical findings to favor attending an affluent school (Gümüş et al., 2022; Holzberger et al., 2020; Tan et al., 2023), the match or mismatch between student individual SES and school composition could lead to various results in terms of their impact on student achievement (e.g., Burke & Sass, 2013; Gibbons & Telhaj, 2016). While certain studies have indicated favorable outcomes resulting from the placement of low-ability students alongside high-performing peers (Burke & Sass, 2013), contrasting research has demonstrated either no effects or even proposed a detrimental influence (Gibbons & Telhaj, 2016). The research evidence within this corpus is likely to vary across nations and studies (Paloyo, 2020). Particularly based on previous assumptions about SES and student achievement that were highlighted in the literature above (Crosnoe, 2009; Gutiérrez, 2022; Harris, 2010), it is essential to comprehend how the match and mismatch (Barranti et al., 2017) between peer and student SES could influence their learning outcomes in different nations. Aiming to address this issue, we pose the following research question:

(1) How does the effect of (in) congruence between students' individual SES and peer SES on student achievement in reading, science, and math across different countries?

Additionally, school factors that might intervene in this mechanism are not fully understood across countries (Perry, 2007). Particularly IL, which is strongly related to student learning (Robinson et al., 2008), might serve as a key mechanism in addressing issues of disparities in schools (Demie, 2023). Improving teaching and learning through strong leadership might lessen inequalities in student achievement (R. D. Goddard et al., 2017). The work of Gümüş et al. (2022) reported that principals IL could reduce the deleterious effect of negative peer effect on student learning, but did not address whether it might make a different impact on different students (e.g., high versus low SES students in low SES schools). Additionally, the fact that the impact could vary across countries, as indicated in the literature, is often ignored. Therefore, this research will examine the role of principals' IL in shaping the effect of match or mismatch between students' individual and peer SES on achievement. The second key research question that will be addressed in this research is:

(2) How do principals' IL moderate the effect of (in)congruence between students' individual SES and peer SES on student achievement in reading, science, and math across different countries?

Method

Data Source

In this study, we used the linked datasets of the Teaching and Learning International Survey (TALIS) and Programme on International Student Assessment (PISA) conducted in 2018. Ideally, the OECD aims first to select 150 schools from each country and ensure the inclusion of at least 20 students from each school. TALIS, the second dataset used in this study, was also designed, and administered by OECD and involves a wide variety of information about schools, teachers, and principals. The main sampling strategy of TALIS is to select 200 schools from each participating country and then select 20 teachers from each school. OECD created school weights (SCHWGT) to adjust for a more representative sample in the data analysis (OECD, 2019a, 2019b).

In this study, we utilized PISA-TALIS-linked data. The link data, however, is available for only nine jurisdictions: Australia, Buenos Aires, Colombia, the Czech Republic, Denmark, Georgia, Malta, Turkey, and Vietnam. We used school-level variables of TALIS as the host data and combined that with the PISA dataset using the PISA-TALIS link identifier and pairing the school data with each student in that school. After combining both datasets, the initial student sample was 32,032. We excluded Buenos Aires from the analysis as it did not provide a representative sample for the entire country. Similarly, we did not include Vietnam due to the lack of plausible values for student achievement, which were crucial for our analysis.

Variables and Measures

A combined TALIS and PISA data include variables at two levels: the student and the school. At the student level, we used student achievement scores in three subject areas: math, science, and reading as the key dependent variables. Instead of using a single achievement score, OECD developed 10 plausible values for each subject area. In this study, we employed each of the plausible values for the analysis as recommended by many researchers (Lorah, 2018). At the student level, we used the Index of Social, Economic, and Cultural Status (ESCS) for the socioeconomic status (SES) of individual students as one of our main independent variables (OECD, 2019a).

At the school level, we have two main independent variables, school-level socio-economic status (SSES) and IL. We used PISA's ESCS index and aggregated it to the school level to create a school composition of SES. As for IL, OECD created an index of school leadership (T3PLEADS) based on Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). In the construction of this scale, principals were requested to indicate the frequency with which they engaged in the following activities within their school over the past 12 months. An example is "I took actions to support co-operation among teachers to develop new teaching practices" (TC3G22D).

Analytic Strategy

For our analyses, we applied a polynomial regression with response surface analysis (Shanock et al., 2010), as this non-linear approach allows us to investigate whether the joint combination of two predictors relates to a specific outcome. Hence, building on Zyphur et al. (2016) and Nestler et al. (2019), we utilized a multilevel polynomial regression model to compute response surface parameters separately for each country. Equation 1 provides the combined model in which all variables are continuous, and a student-level outcome (Y_{ij}) is predicted by observed variables. All predictors were created based on a single student-level (X_{ij}) and a single school-level (W_j) predictor;

$$\begin{aligned} Y_{ij} = {} & \gamma_{00} + \gamma_{10}\left(X_{ij} - \underline{X_j}\right) + \gamma_{20}\left(X_{ij} - \underline{X_j}\right)^2 + \gamma_{30}\left(X_{ij} - \underline{X_j}\right) \times \left(\underline{X_j}\right) + \gamma_{40}\left(X_{ij} - \underline{X_j}\right) \times \left(W_j\right) + \\ & \gamma_{50}\left(X_{ij} - \underline{X_j}\right)^2 \times \left(W_j\right) + \gamma_{60}\left(X_{ij} - \underline{X_j}\right) \times \left(\underline{X_j}\right) \times \left(W_j\right) + \gamma_{01}\left(\underline{X_j}\right) + \gamma_{02}\left(\underline{X_j}\right)^2 + \\ & \gamma_{03}\left(W_j\right) + \gamma_{04}\left(\underline{X_j}\right) \times \left(W_j\right) + \gamma_{05}\left(\underline{X_j}\right)^2 \times \left(W_j\right) + u_{0j} + e_{ij} \end{aligned} \tag{1}$$

where Y_{ij} is a standardized plausible value for one of the three achievement domains, X_{ij} is the standardized student ESCS, $\underline{X_j}$ is the school average of X_{ij} and W_j is the standardized T3PLEADS scores. As Snijders and Bosker (2012, p. 88) noted including random slopes for group-mean centered level-1 variables (i.e., $X_{ij} - \underline{X_j}$) should have strong theoretical justifications, otherwise researchers are advised to be reluctant to include such slopes; hence, Equation 1 is a random intercept model. Nestler et al. (2019) defined Equation 1 as the conditional multilevel response surface analysis (MRSA) model and demonstrated a pick-a-point approach to compute surface parameters with a level-2 moderator (i.e., W_j).

To account for interactions and investigate curvilinear associations between SES and SSES, we created squared terms of these variables as well as interaction terms between both variables and IL. Further, we estimated MRSA parameters a_1-a_5 for the impact of IL on the SSES and SES association with student achievement. Both a_1 and a_2 indicate the line of congruence i.e., the extent to which the two individual and peer SES match matters. a_3 – a_4 indicates the lines of incongruence and deals with to what extent mismatches matter (Barranti et al., 2017). Here, a_1 examines the extent to which SSES and SES match; a_2 investigates if this matching is of similar importance at all levels of SSES and SES and, thus, curvilinearity; a_3 assesses the extent to which the direction of discrepancy or incongruence between SSES and SES predicts student achievement; a_4 tests if the curvilinear association between SSES and SES matter with regard to student achievement; and, a_5 evaluates whether a congruence effect is observable, that is, it tests if corresponding values on SSES and SES are better than discrepancies (Lai et al., 2022). However, interpretation of MRSA parameters must be done holistically with regard to congruence effects: Strict congruence is supported if (i) a_1, a_2, and a_3 do not significantly differ from zero, (ii) a_4 is significantly negative, and (ii) a_5 is not significantly different from zero (Nestler et al., 2019). Humberg et al. (2022) provide a comprehensive demonstration of holistic interpretation examples, offering valuable insights for readers. For our study, if $a_4 \neq 0$, a1 0, and $a5 = 0$, a broad congruence effect is observable, which indicates that SSES-SES congruence influences student achievement, while, at the same time, predictor variables are allowed to have common main effects in addition to a congruence effect. As we intended to investigate how the effect of IL on student achievement is mediated by these relations, we modeled this multi-level association and, hence, reported various values for the MRSA parameters, following Holzberger et al. (2020): a_1 -a_5 values for average IL (suffix: lead = 0), and one SD above (suffix: lead = 1) and below (suffix: lead = −1) the average. Consequently, these values thus show whether differences in the MRSA parameters can be found in schools with very high or very low IL and to what extent student performance differs as a result.

In our analyses, we first utilized an R code to automatically split the PISA-TALIS data set into subsets across countries, achievement domains, and further across each plausible value, resulting in a subset that has an achievement score, ESCS, T3PLEADS, and relevant sampling weights. We then rescaled variables (i.e., separately for each country) to have a mean of zero and a standard deviation of 1. Following the z-standardization, we computed predictors in Equation 1 to run the multilevel model using M*plus* (Muthén & Muthén, 2021) with the model constraint command to compute the surface parameters along with the difference test of these parameters when $W_j = -1$ and $W_j = 1$. The multilevel model with M*plus* utilized an imputation approach pioneered by Rubin (1987) and Schafer (1997) to combine results obtained from the analysis of each plausible value (Arıkan et al., 2020; OECD, 2009). Further, in the multilevel model, both student and school weights are considered, assuming missing values occurred completely at random the full information maximum likelihood to address missing data was used. In the final analytical step, we simulated datasets based on the multilevel model outputs to employ the R package *RSA* (Schönbrodt & Humberg, 2023) and *multilevelRSA* function (Nestler et al., 2019) to create response surface graphs.

Results

Descriptive Results

The plausible values and predictors were standardized at the country level and hence had a mean of 0 and a standard deviation of 1, an approach to compute partially standardized regression coefficients as an effect

Table 1. Sample, ICC and descriptive statistics.

	Sample Size		ICC			Min. Value			Max. Value		
Country	J	n	Math	Read	Scie	SES	S-SES	Lead.	SES	S-SES	Lead.
AUS	131	18.35	.22	.18	.18	−3.31	−1.49	−2.89	2.67	1.27	1.77
COL	152	30.93	.40	.39	.33	−2.41	−2.40	−2.54	3.15	1.72	1.49
CZE	182	23.07	.37	.41	.38	−7.01	−1.69	−3.05	4.23	1.16	2.38
DNK	102	22.46	.17	.15	.15	−3.64	−1.32	−2.33	2.57	0.98	2.46
GEO	134	23.58	.30	.31	.31	−6.24	−1.74	−3.70	5.09	1.26	2.37
MLT	49	68.37	.21	.24	.20	−3.99	−1.03	−3.09	3.98	1.00	2.40
TUR	143	38.95	.54	.56	.56	−2.57	−1.26	−3.04	3.49	2.03	2.08

Note: J (number of schools) and n (average cluster size) were taken from the M*plus* output for the conditional model, SES is the group mean centered standardized-ESCS, S-SES is the group mean of standardized-ESCS, Lead is the standardized-T3PLEADS; ICC; Intra-class correlation, AUS: Australia, COL: Colombia, CZE: Czeck Republic, DNK: Denmark, GEO: Georgia, MLT: Malta, TUR: Turkey.

size measure (Lorah, 2018). Table 1, separately for each country, reports the number of schools, average cluster size, intraclass correlation coefficient (ICC) for the outcome, and minimum and maximum values for the predictors. The number of schools varied between 49 and 182 and the average cluster size varied between 18.35 and 68.37. The smallest ICC values were observed for Denmark (i.e., .15) and the largest ICCs were observed for Turkey (i.e., .56) and justified the use of multilevel modeling, indicating a comparatively high between-school segregation in Colombia, the Czech Republic, Georgia and Turkey, and a rather low social segregation in the other countries under study. The largest minimum predictor values were −7.01, −2.40, and −3.70, respectively, for SES, SSES, and IL. The largest maximum predictor values were 5.09, 2.03, and 2.46, respectively, for SES, SSES, and IL.

Model Testing

The results of multilevel polynomial regression model are reported separately for each achievement domain (i.e., Mathematics, Reading and Science) in Tables 2–4 for the seven countries. Consequently, 21 different model outputs were reported following the approach suggested by several researchers (Barranti et al., 2017; Humberg et al., 2022; Nestler et al., 2019) and the results are mainly threefold: (a) covariate effects, (b) surface parameters (i.e., a_1, a_2, a_3, a_4, a_5) for $Lead = -1$ and $Lead = 1$, and (c) difference test for the surface parameters within a model.

It is worth noting that student-level SES had a significant and positive effect across all models. The effect of individual SES varied across jurisdictions and achievement domains (see Tables 2–4). The largest SES coefficient was detected for Denmark's reading outcome ($B_{\text{SES}} = 0.30$, $p < .01$); whereas the smallest SES coefficient was detected for Turkey's mathematics outcome ($B_{\text{SES}} = 0.06$, $p = .02$). This means that individual SES has a more prominent influence on reading achievement in Denmark compared to mathematics achievement in Turkey. Similarly, SSES also had a significant and positive effect across all models, showing evidence of peer effect in all jurisdictions and achievement domains. The largest SSES coefficient was detected for the Czech Republic's mathematics outcome ($B_{\text{S-SES}} = 0.88$, $p < .01$); whereas the smallest S-SES coefficient was detected for Georgia's science outcome ($B_{\text{SES}} = 0.36$, $p < .01$). The effect size of SSES is overall larger than that of individual SES, meaning that school composition effect exerts more impact than students' individual backgrounds. Overall, SSES is consistently greater than that of individual SES. This implies that the backgrounds and outcomes of peers have a more significant influence compared to the individual socioeconomic backgrounds of students. However, the effect of SES and SSES was not linear except for Austria and Turkey. For example, the quadratic term for SES (i.e., SES2) was negative and significant when predicting Malta's achievement scores (e.g., ${B_{\text{SES}}}^2 = -0.04$, $p = .01$ for mathematics). This indicates that at lower levels of SES, an increase in SES is associated with a greater positive impact on student achievement; however, as SES continues to rise, the additional impact on achievement becomes smaller. The quadratic term for S-SES (i.e., S-SES2) was positive and significant for Columbia and Georgia (e.g., ${B_{\text{S-SES}}}^2 = 0.19$, $p = .01$ for Columbia reading scores), meaning that in both countries the impact of peer effect on student achievement will increase in advantageous schools (high SSES); the cross-level interaction between SES and S-SES was positive and significant when predicting Denmark's achievement scores (e.g., $B_{SES\ x\ S\text{-}SES} = 0.18$, $p = .02$ for mathematics). At last, the leadership scores formed interactions with S-SES when predicting Denmark's reading scores (i.e., BS-SES × Lead = 0.22, $p = .02$). This shows that the impact of individual SES

Table 2. The results for math achievement across countries.

	AUS			COL			CZE			DNK			GEO			MLT			TUR		
	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p
Coefficient																					
Constant	−0.02	0.04	0.62	**−0.14**	0.07	0.03*	**−0.16**	0.06	0.01*	0.04	0.06	0.47	−0.23	0.12	0.06	0.05	0.06	0.39	**−0.38**	0.13	0.00*
SES	**0.20**	0.03	0.00*	**0.09**	0.03	0.01*	**0.18**	0.03	0.00*	**0.27**	0.04	0.00*	**0.18**	0.03	0.00*	**0.23**	0.03	0.00*	**0.06**	0.03	0.02*
SES^2	−0.03	0.02	0.19	0.02	0.03	0.56	−0.04	0.02	0.08	0.02	0.03	0.41	−0.01	0.02	0.62	**−0.04**	0.02	0.01*	0.02	0.03	0.56
SES x S-SES	0.05	0.06	0.38	0.06	0.04	0.12	−0.06	0.05	0.18	**0.18**	0.08	0.02*	0.00	0.09	0.99	0.08	0.06	0.15	0.03	0.05	0.46
SES x Lead	0.00	0.03	0.87	−0.02	0.02	0.48	−0.03	0.03	0.30	0.06	0.04	0.16	−0.03	0.04	0.41	0.03	0.02	0.29	−0.02	0.03	0.61
SES^2 x Lead	−0.01	0.02	0.73	−0.03	0.02	0.21	0.01	0.01	0.25	0.02	0.02	0.45	−0.01	0.02	0.68	0.02	0.01	0.24	0.00	0.02	0.87
SES x S-SES x Lead	−0.02	0.04	0.72	−0.04	0.04	0.37	0.01	0.03	0.75	−0.08	0.06	0.15	0.00	0.09	0.97	0.01	0.06	0.92	0.05	0.04	0.12
S-SES	**0.64**	0.07	0.00*	0.69	0.08	0.00*	**0.88**	0.09	0.00*	**0.71**	0.11	0.00*	**0.38**	0.13	0.00*	0.81	0.12	0.00*	**0.66**	0.15	0.00*
$S\text{-}SES^2$	0.08	0.10	0.44	0.15	0.08	0.06	0.08	0.10	0.45	−0.05	0.16	0.74	0.21	0.11	0.06	−0.04	0.15	0.82	0.17	0.10	0.09
Lead	0.00	0.04	0.97	0.03	0.07	0.65	−0.06	0.04	0.15	−0.03	0.05	0.56	0.14	0.10	0.18	−0.05	0.06	0.38	−0.04	0.11	0.69
S-SES x Lead	−0.04	0.06	0.51	0.00	0.07	0.96	−0.05	0.06	0.42	0.18	0.12	0.14	0.11	0.14	0.43	−0.02	0.08	0.85	0.13	0.12	0.25
$S\text{-}SES^2$ x Lead	−0.02	0.08	0.78	0.01	0.07	0.93	−0.04	0.07	0.57	0.15	0.13	0.26	0.16	0.16	0.32	0.15	0.15	0.31	−0.14	0.10	0.16
Surface test																					
$a_{1\mid lead=-1}$	**0.88**	0.12	0.00*	**0.80**	0.12	0.00*	**1.15**	0.10	0.00*	**0.75**	0.17	0.00*	0.47	0.20	0.02*	1.02	0.13	0.00*	**0.61**	0.19	0.00*
$a_{2\mid lead=-1}$	0.14	0.17	0.41	**0.28**	0.13	0.04*	−0.01	0.17	0.97	0.06	0.23	0.79	0.05	0.26	0.84	−0.17	0.21	0.42	**0.31**	0.15	0.04*
$a_{3\mid lead=-1}$	**−0.47**	0.12	0.00*	**−0.60**	0.13	0.00*	**−0.72**	0.12	0.00*	−0.32	0.20	0.11	−0.05	0.25	0.84	**−0.62**	0.15	0.00*	**−0.45**	0.18	0.01*
$a_{4\mid lead=-1}$	0.01	0.17	0.96	0.09	0.11	0.44	0.14	0.16	0.39	−0.46	0.31	0.14	0.04	0.23	0.85	−0.32	0.25	0.21	**0.35**	0.15	0.02*
$a_{5\mid lead=-1}$	−0.12	0.15	0.43	−0.10	0.11	0.38	−0.17	0.14	0.24	0.21	0.26	0.43	−0.05	0.21	0.81	0.13	0.22	0.57	**−0.30**	0.14	0.03*
$a_{1\mid lead=1}$	**0.79**	0.09	0.00*	0.76	0.10	0.00*	**0.99**	0.12	0.00*	**1.22**	0.16	0.00*	**0.64**	0.15	0.00*	1.05	0.14	0.00*	**0.84**	0.21	0.00*
$a_{2\mid lead=1}$	0.05	0.12	0.69	0.17	0.13	0.20	−0.04	0.11	0.74	0.23	0.18	0.21	0.35	0.22	0.12	0.17	0.21	0.41	0.14	0.16	0.41
$a_{3\mid lead=1}$	**−0.40**	0.08	0.00*	**−0.62**	0.10	0.00*	**−0.68**	0.13	0.00*	**−0.56**	0.16	0.00*	**−0.34**	0.17	0.04*	**−0.54**	0.16	0.00*	**−0.75**	0.19	0.00*
$a_{4\mid lead=1}$	−0.02	0.12	0.90	0.12	0.12	0.32	0.07	0.10	0.47	0.05	0.14	0.75	0.35	0.21	0.09	0.00	0.23	0.99	−0.04	0.16	0.81
$a_{5\mid lead=1}$	−0.09	0.10	0.36	−0.17	0.12	0.18	−0.06	0.11	0.54	−0.06	0.15	0.71	**−0.38**	0.18	0.03*	−0.14	0.21	0.50	−0.01	0.16	0.95
$a_{1\text{-}difference}$	−0.09	0.14	0.52	−0.04	0.15	0.79	−0.16	0.13	0.23	0.46	0.25	0.06	0.17	0.26	0.52	0.02	0.15	0.88	0.24	0.24	0.33
$a_{2\text{-}difference}$	−0.09	0.19	0.64	−0.11	0.19	0.56	−0.03	0.16	0.85	0.17	0.28	0.55	0.30	0.41	0.47	0.34	0.30	0.25	−0.17	0.20	0.39
$a_{3\text{-}difference}$	0.07	0.13	0.57	−0.03	0.15	0.85	0.05	0.14	0.75	−0.24	0.25	0.33	−0.29	0.33	0.38	0.08	0.17	0.63	−0.30	0.24	0.22
$a_{4\text{-}difference}$	−0.03	0.19	0.90	0.03	0.15	0.85	−0.07	0.13	0.59	0.50	0.29	0.09	0.31	0.31	0.32	0.32	0.33	0.33	−0.39	0.23	0.10
$a_{5\text{-}difference}$	0.03	0.16	0.85	−0.07	0.16	0.66	0.10	0.14	0.46	−0.26	0.27	0.34	−0.33	0.32	0.31	−0.27	0.29	0.37	0.29	0.21	0.17

*indicates a p-value below 0.05. Significant coefficients are bold. Lead: instructional leadership, SES: individual socio-economic status, S-SES: school socio-economic status.

Table 3. Results for reading achievement across countries.

	AUS			COL			CZE			DNK			GEO			MLT			TUR		
	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p
Coefficient																					
Constant	−0.02	0.04	0.70	**−0.15**	0.07	0.03*	**−0.19**	0.07	0.01*	0.03	0.05	0.55	−0.18	0.14	0.18	0.02	0.07	0.77	**−0.39**	0.16	0.01*
SES	**0.22**	0.03	0.00*	**0.08**	0.02	0.00*	**0.15**	0.03	0.00*	**0.30**	0.04	0.00*	**0.14**	0.03	0.00*	**0.17**	0.02	0.00*	**0.09**	0.02	0.00*
SES^2	−0.02	0.02	0.29	0.00	0.02	0.95	**−0.04**	0.02	0.01*	0.01	0.03	0.63	−0.03	0.02	0.09	**−0.05**	0.01	0.00*	−0.01	0.02	0.52
SES x S-SES	0.06	0.06	0.32	0.05	0.04	0.17	−0.04	0.05	0.35	**0.15**	0.07	0.03*	0.03	0.08	0.74	0.07	0.05	0.18	−0.01	0.05	0.83
SES x Lead	0.01	0.03	0.66	0.00	0.02	0.98	−0.03	0.03	0.23	0.05	0.04	0.20	−0.03	0.03	0.42	0.03	0.02	0.12	−0.02	0.03	0.56
SES^2 x Lead	−0.02	0.02	0.42	−0.02	0.02	0.38	0.01	0.01	0.18	0.04	0.03	0.22	0.00	0.02	0.93	0.01	0.01	0.37	0.02	0.02	0.32
SES x S-SES x Lead	−0.01	0.05	0.85	−0.01	0.03	0.69	0.00	0.03	0.88	−0.07	0.05	0.19	−0.04	0.07	0.60	−0.02	0.05	0.67	0.05	0.03	0.07
S-SES	**0.63**	0.06	0.00*	**0.77**	0.08	0.00*	**0.87**	0.11	0.00*	**0.68**	0.10	0.00*	**0.43**	0.12	0.00*	**0.71**	0.15	0.00*	**0.63**	0.18	0.00*
$S\text{-}SES^2$	0.05	0.10	0.61	**0.19**	0.07	0.01*	0.08	0.11	0.49	−0.01	0.14	0.94	**0.22**	0.10	0.03*	0.11	0.22	0.61	0.15	0.12	0.22
Lead	−0.01	0.05	0.89	0.05	0.07	0.49	−0.05	0.04	0.25	**−0.10**	0.04	0.03*	0.13	0.12	0.26	0.00	0.07	0.99	−0.14	0.11	0.21
S-SES x Lead	−0.04	0.06	0.44	−0.04	0.07	0.52	−0.11	0.06	0.09	**0.22**	0.10	0.02*	0.15	0.14	0.31	−0.02	0.11	0.87	0.19	0.12	0.13
$S\text{-}SES^2$ x Lead	0.07	0.08	0.36	−0.03	0.06	0.67	−0.06	0.06	0.31	**0.28**	0.12	0.02*	0.15	0.13	0.27	−0.05	0.19	0.80	−0.08	0.10	0.42
Surface test																					
$a_{1\vert lead=-1}$	**0.88**	0.10	0.00*	**0.89**	0.13	0.00*	**1.17**	0.11	0.00*	**0.70**	0.15	0.00*	**0.45**	0.22	0.04*	**0.86**	0.19	0.00*	**0.55**	0.21	0.01*
$a_{2\vert lead=-1}$	0.03	0.14	0.81	**0.30**	0.12	0.01*	0.04	0.16	0.81	−0.10	0.21	0.66	0.10	0.22	0.65	0.19	0.26	0.47	0.14	0.15	0.34
$a_{3\vert lead=-1}$	**−0.47**	0.10	0.00*	**−0.73**	0.13	0.00*	**−0.80**	0.13	0.00*	−0.20	0.17	0.22	−0.13	0.25	0.60	**−0.59**	0.20	0.00*	−0.33	0.20	0.09
$a_{4\vert lead=-1}$	−0.10	0.17	0.58	0.17	0.11	0.12	0.13	0.16	0.43	**−0.54**	0.26	0.04*	−0.03	0.20	0.90	0.00	0.31	0.99	0.27	0.15	0.07
$a_{5\vert lead=-1}$	0.01	0.14	0.93	−0.20	0.11	0.07	−0.19	0.14	0.19	0.27	0.23	0.23	−0.10	0.17	0.56	−0.22	0.28	0.44	−0.27	0.15	0.08
$a_{1\vert lead=1}$	**0.81**	0.09	0.00*	**0.80**	0.07	0.00*	**0.89**	0.13	0.00*	**1.25**	0.14	0.00*	**0.69**	0.12	0.00*	**0.89**	0.16	0.00*	**0.89**	0.24	0.00*
$a_{2\vert lead=1}$	0.13	0.13	0.35	0.19	0.11	0.09	−0.06	0.11	0.59	**0.40**	0.17	0.02*	0.32	0.21	0.11	0.08	0.29	0.79	0.12	0.20	0.56
$a_{3\vert lead=1}$	**−0.36**	0.08	0.00*	**−0.65**	0.09	0.00*	−0.64	0.15	0.00*	**−0.56**	0.12	0.00*	**−0.47**	0.15	0.00*	**−0.48**	0.19	0.01*	**−0.74**	0.22	0.00*
$a_{4\vert lead=1}$	0.03	0.14	0.80	0.12	0.09	0.21	0.02	0.11	0.86	0.24	0.16	0.14	0.34	0.18	0.06	−0.02	0.31	0.94	0.03	0.18	0.86
$a_{5\vert lead=1}$	−0.16	0.12	0.17	**−0.18**	0.08	0.03*	−0.04	0.11	0.68	−0.23	0.16	0.15	**−0.39**	0.15	0.01*	−0.10	0.29	0.72	−0.06	0.18	0.72
$a_{1\text{-difference}}$	−0.07	0.13	0.60	−0.09	0.13	0.51	**−0.28**	0.13	0.04*	**0.54**	0.21	0.01*	0.24	0.27	0.37	0.03	0.21	0.88	0.34	0.26	0.19
$a_{2\text{-difference}}$	0.09	0.18	0.62	−0.11	0.15	0.47	−0.09	0.13	0.48	**0.49**	0.24	0.04*	0.23	0.32	0.48	−0.11	0.36	0.76	−0.03	0.19	0.89
$a_{3\text{-difference}}$	0.11	0.12	0.36	0.08	0.15	0.58	0.16	0.14	0.28	−0.35	0.20	0.07	−0.34	0.32	0.28	0.10	0.23	0.65	−0.41	0.25	0.10
$a_{4\text{-difference}}$	0.13	0.19	0.50	−0.06	0.13	0.66	−0.11	0.13	0.39	**0.78**	0.28	0.01*	0.37	0.30	0.21	−0.03	0.41	0.95	−0.24	0.23	0.30
$a_{5\text{-difference}}$	−0.17	0.15	0.27	0.02	0.12	0.89	0.15	0.13	0.25	−0.50	0.26	0.05	−0.29	0.26	0.27	0.12	0.37	0.75	0.20	0.22	0.36

*indicates a p-value below 0.05. Significant coefficients are bold. Lead: instructional leadership, SES: individual socio-economic status, S-SES: school socio-economic status.

Table 4. Results for science achievement across countries.

	AUS			COL			CZE			DNK			GEO			MLT			TUR		
	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p
Coefficient																					
Constant	−0.01	0.04	0.74	**−0.16**	0.06	0.01*	**−0.16**	0.07	0.02*	0.02	0.06	0.74	−0.16	0.12	0.17	0.05	0.05	0.37	**0.40**	0.16	0.01*
SES	**0.21**	0.03	0.00*	**0.07**	0.03	0.01*	**0.17**	0.03	0.00*	**0.29**	0.04	0.00*	**0.16**	0.03	0.00*	**0.21**	0.03	0.00*	**0.09**	0.03	0.00*
SES^2	−0.03	0.03	0.19	0.02	0.03	0.49	−0.02	0.02	0.17	0.01	0.03	0.68	−0.03	0.02	0.10	**−0.06**	0.02	0.00*	−0.01	0.02	0.72
SES x S-SES	0.02	0.05	0.72	0.08	0.05	0.12	−0.02	0.05	0.66	**0.15**	0.07	0.04*	−0.03	0.08	0.75	0.04	0.06	0.50	0.04	0.05	0.46
SES x Lead	−0.01	0.03	0.85	0.01	0.03	0.77	−0.03	0.03	0.40	0.05	0.03	0.16	−0.03	0.03	0.36	0.04	0.02	0.12	−0.01	0.03	0.76
SES^2 x Lead	−0.01	0.02	0.64	−0.02	0.03	0.37	0.00	0.01	0.69	0.03	0.03	0.33	0.00	0.01	0.92	0.02	0.02	0.15	0.02	0.02	0.36
SES x S-SES x Lead	0.01	0.05	0.93	−0.04	0.04	0.40	0.04	0.03	0.19	−0.04	0.05	0.47	−0.02	0.07	0.77	−0.01	0.05	0.92	0.07	0.04	0.10
S-SES	**0.59**	0.07	0.00*	**0.71**	0.07	0.00*	**0.86**	0.11	0.00*	**0.65**	0.09	0.00*	**0.36**	0.13	0.00*	**0.78**	0.12	0.00*	**0.65**	0.18	0.00*
$S\text{-}SES^2$	0.07	0.11	0.50	**0.22**	0.08	0.01*	0.05	0.11	0.68	−0.02	0.13	0.90	0.16	0.10	0.12	0.00	0.17	0.99	0.14	0.13	0.26
Lead	0.00	0.05	0.93	0.05	0.06	0.38	**−0.09**	0.04	0.03*	−0.07	0.05	0.13	0.13	0.10	0.20	−0.05	0.05	0.34	−0.11	0.11	0.30
S-SES x Lead	−0.06	0.07	0.40	−0.04	0.06	0.48	−0.06	0.06	0.39	0.16	0.10	0.13	0.13	0.13	0.32	−0.05	0.08	0.48	0.17	0.12	0.17
$S\text{-}SES^2$ x Lead	−0.01	0.08	0.87	0.00	0.07	0.97	0.07	0.05	0.23	0.16	0.12	0.21	0.09	0.14	0.53	0.13	0.14	0.36	−0.09	0.10	0.38
Surface test																					
$a_{1\|lead=-1}$	**0.86**	0.12	0.00*	**0.81**	0.11	0.00*	**1.11**	0.12	0.00*	**0.74**	0.15	0.00*	**0.41**	0.21	0.05*	**1.00**	0.14	0.00*	**0.58**	0.22	0.01*
$a_{2\|lead=-1}$	0.07	0.16	0.64	**0.37**	0.15	0.01*	−0.11	0.17	0.53	−0.01	0.21	0.98	0.04	0.22	0.87	−0.16	0.21	0.46	0.17	0.16	0.29
$a_{3\|lead=-1}$	**−0.43**	0.11	0.00*	**−0.69**	0.12	0.00*	**−0.72**	0.13	0.00*	−0.25	0.16	0.11	−0.04	0.24	0.86	**−0.66**	0.16	0.00*	**−0.39**	0.20	0.05*
$a_{4\|lead=-1}$	0.05	0.19	0.81	0.15	0.14	0.28	0.01	0.16	0.95	−0.37	0.24	0.12	0.05	0.20	0.81	−0.25	0.27	0.35	0.24	0.15	0.11
$a_{5\|lead=-1}$	−0.11	0.16	0.49	−0.18	0.13	0.16	−0.01	0.14	0.96	0.16	0.21	0.45	−0.11	0.18	0.55	0.05	0.23	0.85	−0.26	0.15	0.09
$a_{1\|lead=1}$	**0.74**	0.10	0.00*	**0.75**	0.07	0.00*	**0.95**	0.14	0.00*	**1.14**	0.15	0.00*	**0.62**	0.14	0.00*	**0.96**	0.13	0.00*	**0.89**	0.24	0.00*
$a_{2\|lead=1}$	0.04	0.13	0.76	**0.25**	0.11	0.02*	0.11	0.10	0.29	0.29	0.19	0.13	0.17	0.20	0.40	0.13	0.21	0.53	0.17	0.20	0.40
$a_{3\|lead=1}$	**−0.33**	0.08	0.00*	**−0.59**	0.08	0.00*	**−0.66**	0.15	0.00*	**−0.47**	0.13	0.00*	**−0.36**	0.15	0.02*	**−0.48**	0.16	0.00*	**−0.74**	0.24	0.00*
$a_{4\|lead=1}$	−0.01	0.13	0.95	0.17	0.10	0.07	0.07	0.10	0.46	0.07	0.17	0.68	0.27	0.22	0.23	0.06	0.22	0.79	−0.04	0.18	0.83
$a_{5\|lead=1}$	−0.10	0.11	0.38	**-0.22**	0.09	0.02*	−0.13	0.10	0.17	−0.10	0.15	0.48	−0.28	0.17	0.10	−0.16	0.20	0.43	−0.04	0.19	0.83
$a_{1\text{-}difference}$	−0.12	0.15	0.42	−0.07	0.12	0.58	−0.16	0.14	0.25	0.40	0.22	0.06	0.21	0.25	0.41	−0.04	0.15	0.81	0.32	0.26	0.23
$a_{2\text{-}difference}$	−0.03	0.19	0.86	−0.12	0.18	0.49	0.22	0.13	0.10	0.29	0.27	0.28	0.13	0.31	0.67	0.29	0.28	0.30	0.00	0.21	1.00
$a_{3\text{-}difference}$	0.10	0.13	0.46	0.10	0.13	0.48	0.06	0.14	0.67	−0.22	0.21	0.30	−0.32	0.29	0.28	0.18	0.17	0.30	−0.36	0.24	0.14
$a_{4\text{-}difference}$	−0.05	0.20	0.80	0.02	0.15	0.88	0.06	0.12	0.61	0.44	0.28	0.11	0.22	0.32	0.49	0.31	0.31	0.32	−0.28	0.22	0.22
$a_{5\text{-}difference}$	0.01	0.16	0.95	−0.04	0.14	0.78	−0.12	0.11	0.26	−0.26	0.25	0.31	−0.17	0.28	0.53	−0.20	0.28	0.47	0.22	0.22	0.31

*indicates a p-value below 0.05. Significant coefficients are bold. Lead: instructional leadership, SES: individual socio-economic status, S-SES: school socio-economic status.

on math achievement is larger in high SES schools and leadership increases the strength of association between peer effect and student reading achievement only in Denmark.

The surface parameters were computed based on the regression coefficients following the *pick-a-point* approach introduced by Nestler et al. (2019), separately for low and high leadership scores (i.e., $Lead = -1$ and $Lead = 1$). A consistent pattern was observed for the first surface parameter (i.e., a1) across all countries, all achievement domains, and both low and high Lead scores. Interpretation of these values (a_1-a_5) should be carried out alongside the graphs (see Figure). A positive a_1 value indicates that the outcome is higher with a congruence at the higher level (Barranti et al., 2017). Our analysis revealed that all a1 values were positive and statistically significant, regardless of school leadership, suggesting that student achievement is expected to be higher for high SES students in high SES schools. The largest a_1 value was detected for Denmark's reading scores for high lead scores ($a_{1|lead=1} = 1.25$, $p < .01$), and the smallest a_1 value was detected for Georgia's reading scores for low lead scores ($a_{1lead=-1} = 0.45$, $p < .01$). This means that regardless of leadership, high SES student attending high SSES will score better. However, a significant difference between weak and strong leadership was observed only in the Czech Republic ($a_{1\text{-difference}} = -0.28$, $p = .04$) and Denmark ($a_{1\text{-difference}} = 0.54$, $p = .01$) for reading scores, meaning that while IL is making a positive difference in Denmark, the impact is negative in the Czech Republic. A positive a_2 value shows that the outcome is higher when two variables match at more extreme levels than at midrange levels (Barranti et al., 2017). Our results showed significant and positive a_2 values for a few models; for example, in Columbia the reading scores for both low and high leadership scores (e.g., $a_{2lead=-1} = 0.37$, $p = .01$); however, the a_2 values were insignificant in 15 out of 21 models, indicating that the relationship between achievement along the line of SES and SSES agreement is linear. Additionally, the difference between low and high leadership was only significant for the reading parameter in Denmark ($a_{4\text{-difference}} = 0.49$, $p = .04$). A positive a_3 value shows that the outcome is higher when one predictor is higher than the other predictor (Barranti et al., 2017). In our model, a_3 value was significantly lower than 0 in 13 out of 21 models (e.g., $a_{3lead=-1} = -0.47$, $p<.01$ for Austria's mathematics scores). All high leadership values were negative and significant for all domains. The non-significant values were detected for low leadership values for Denmark and Georgia across all achievement domains. a_4 deals with whether matches are better or worse than mismatches. A positive value means that the outcome is higher when the two predictors deviate from one another rather than match (Barranti et al., 2017). In our model, a_4 value was non-significant in 19 out of 21 models and significant only for Turkey's mathematics scores when leadership was low ($a_{4lead=-1} = 0.35$, $p = .02$) and Denmark's reading scores when leadership was low ($a_{4lead=-1} = -0.54$, $p = .04$). Additionally, the difference between low and high leadership was only significant for reading parameter in Denmark ($a_{4\text{-difference}} = 0.78$, $p = .01$). Finally, the fifth surface parameter a_5 value was also non-significant in 16 out of 21 models but negative and significant for Turkey's mathematics scores with low lead ($a_{5lead=-1} = -0.30$, $p = .03$), Georgia's mathematics ($a_{5|lead=1} = -0.38$, $p = .03$) and reading scores ($a_{5|lead=1} = -0.39$, $p = .01$) with high lead, and Columbia's reading ($a_{5|lead=1} = -0.18$, $p = .03$) and science scores ($a_{5|lead=1} = -0.22$, $p = .02$) with high lead.

The interpretation of MRSA parameters must, however, be made holistically about congruence effects (Humberg et al., 2022). Based on the criteria highlighted in the data analysis section (Nestler et al., 2019), we conclude that our results do not support either a broad or a strict congruence effect because all a_1 parameters are significantly different from zero. In addition, a_4 is not significant in almost all cases, which means that the surface does not predict that students will achieve better in the case of congruence between SES and SES. When a_3 is significantly different from 0, the surface also contradicts a congruence effect. This overall indicates that SSES-SES congruence does not matter for student achievement. In many cases, a_3 was negative and significant, while a_4 was always insignificant, indicating a negative linear LOIC. This means student achievement is higher when SSES and SES diverge. This seems to be the case in almost all domains and countries when there is stronger IL (lead = 1), but not everywhere when leadership is weak (lead = −1). We therefore suggest that the mixing of students from various backgrounds, in general, has positive effects, especially in schools with strong IL.

As seen in Figures 1, 2 and 3, low SES students' math, reading and science achievement in high SES schools declines when there is weak leadership. However, in the case of strong leadership, all students in low-SES schools and low-SES students in high-SES schools benefit to a large extent. The gain seems to be larger for reading, particularly for students attending a disadvantaged school. In the Czech Republic, IL

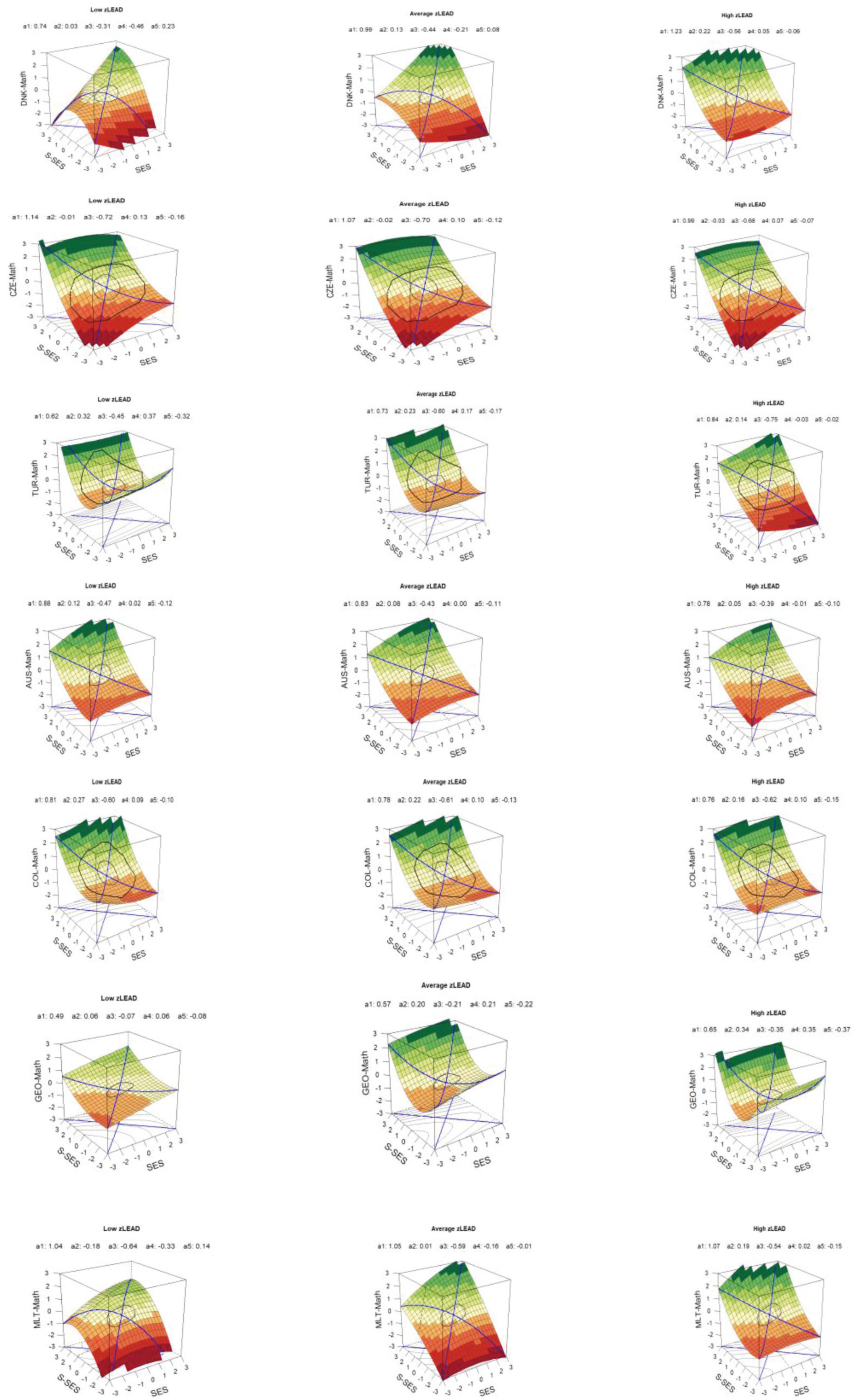

Figure 1. RSA of SES and SSES for math achievement at low, average and high leadership.

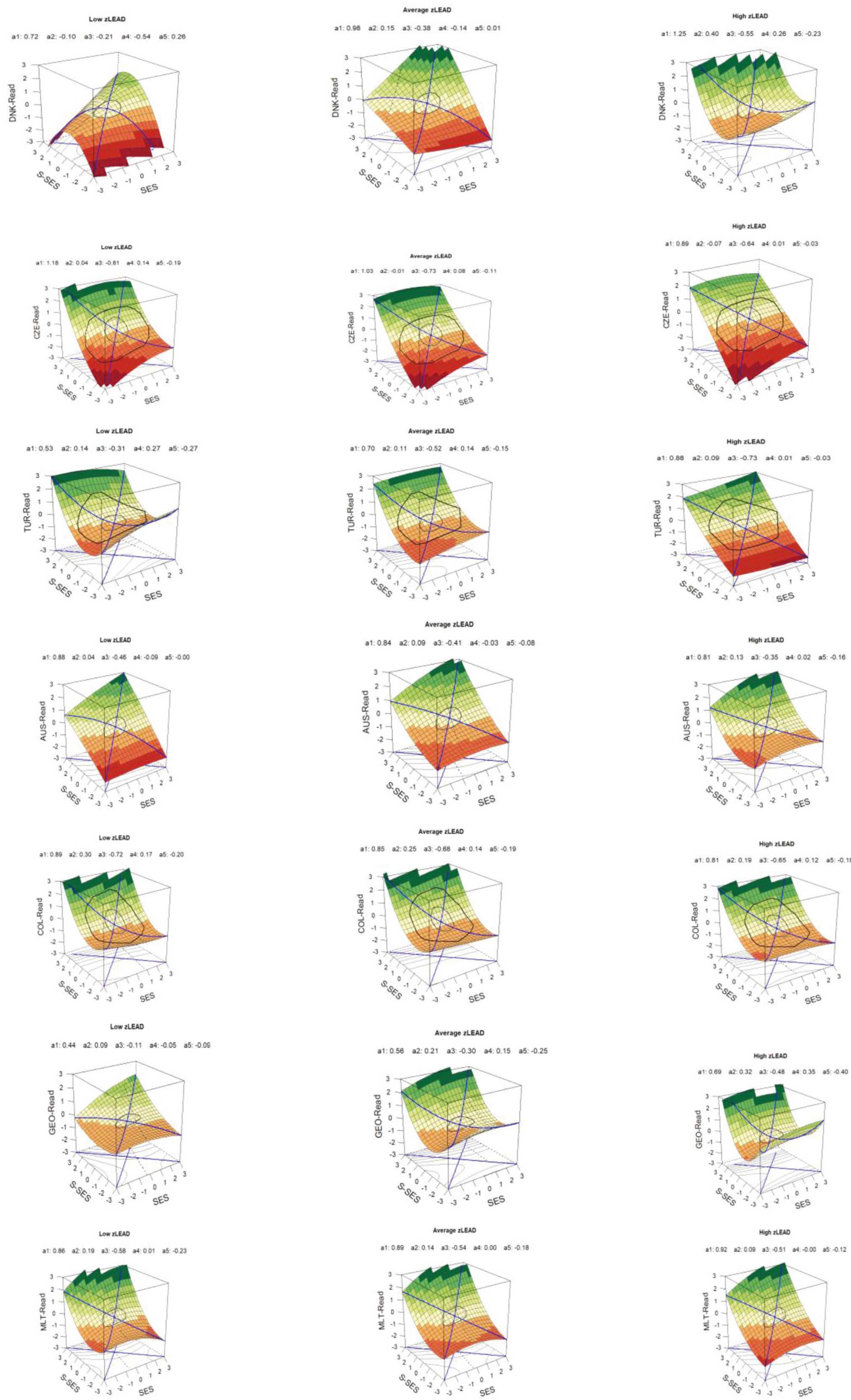

Figure 2. RSA of SES and SSES for reading achievement at low, average and high leadership.

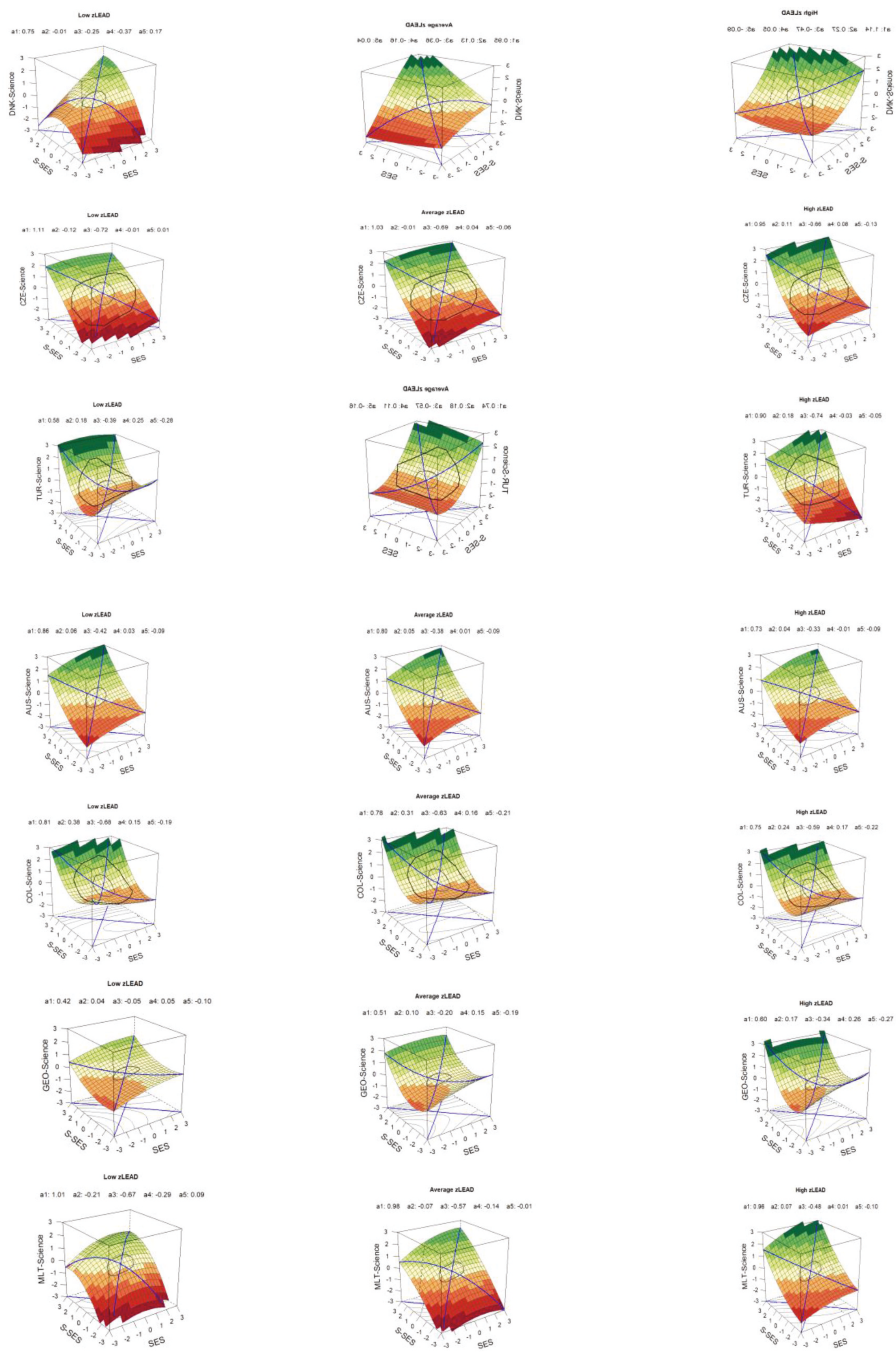

Figure 3. RSA of SES and SSES for science achievement at low, average and high leadership.

supports student science achievement, particularly for those attending disadvantaged schools. In Turkey, interestingly, stronger IL means less math and science achievement for students who attend a school that does not match their individual SES. In Georgia, IL contributes to the overall math and reading achievement of students attending a school surrounded by disadvantaged peers. In Malta, leadership seems to make some difference in the math achievement of students enrolled in low-SES schools and low-SES students enrolled in high-SES schools.

Discussion

Our findings have several theoretical contributions. First, we find that student achievement is associated with both school and student SES, and much more strongly with the former than the latter, confirming the work of Sirin (2005). Second, our results show differential school composition effects, with the school composition effect strongest for low SES students in high SES schools. Conversely, we found that high-SES students can suffer from attending schools dominated by low-SES communities. These results appear to support both boutique and rainbow perspectives, but for different student groups (Gutiérrez, 2022). It supports the boutique perspective by indicating that the match between individual and peer SES would yield better student outcomes. However, our results indicate that this is true only if the match occurs at a high SES level. Additionally, the rainbow perspective suggests that students perform better when attending a school where there is a mismatch between their individual SES and that of their peers (Gutiérrez, 2022). We found that this is also true, but for students from low-SES backgrounds. However, our results did not support the "little frog big pond" perspective, pointing to the disadvantage that a low SES student attending a school with predominantly high SES students might face due to being compared with other students in a more competitive context (Crosnoe, 2009).

Third, the findings in our second point explain why our results do not support congruence theory but do support somewhat (in)congruence theory. Incongruence theory is supported by the fact that low SES students benefit more from attending high SES schools than from attending low SES schools; high SES students also benefit from attending a high SES school, but not to the same extent as their low SES peers. Fourth, IL moderates the relationships between student SES and school SES, as well as student achievement. Strong IL magnifies benefits for low-SES students in high-SES schools, as well as for all students at low-SES schools. It also ensures optimal achievement in socially mixed schools. Finally, our results analyzing the interactive effect of SSES and SES as well as the role of leadership in predicting student achievement across three subject matters produce varying results based on country context. For example, we found that while IL is making a positive difference in Denmark, the impact is negative in the Czech Republic.

With regard to education policy, however, the results highlight that policy borrowing (Burdett & O'Donnell, 2016) may be problematic. This is because while we found positive effects of leadership in one country, we found negative effects in another. This means that what may be a good solution in one country to achieve social justice in education and more effective schools may have no effect, or even a negative one, elsewhere. This begs the question of why this is the case. As Burdett and O'Donnell (2016) note, education systems are complex and sociopolitical systems in their own right, as is the practice of policy borrowing itself. Accordingly, a variety of actors with different agendas and goals are involved in the development of education. This, again, occurs in various contexts, which in turn shape individual and institutional actions. In this regard, Silver (1994) has long shown that concepts of inequality vary across societies and encompass different mechanisms of exclusion. For example, in liberal societies, individuals and their families are responsible for their full inclusion and integration into society, and education serves as a means to achieve this (Alexiadou, 2002). In republican systems based on solidarity, on the other hand, education serves to forge stronger bonds between the individual and society (Zay, 2005). In this respect, the role of the social composition of a school and a school's leadership might also be linked to these orientations.

In terms of research, it should be noted that the RSA approach used also allows for the testing of additional hypotheses (Humberg et al., 2022; Schneider et al., 2022), i.e., 1. An optimal margin hypothesis, which examines whether student achievement can be maximized (or minimized) when school-level SES and individual SES differ by a certain amount, 2. An interaction hypothesis, which examines whether the effect of individual SES on achievement is linear, with school-level SES moderating this association, or 3. A reverse congruency effect hypothesis, which evaluates whether individual student achievement is associated with a high discrepancy between individual and school-level SES in either direction. Although computationally challenging, it would also be relevant and important to use longitudinal and dynamic RSA models (Deventer et al., 2019) to examine how the interplay of individual and school-level SES and their joint influence on student achievement trajectories over time.

Limitations and Suggestions for Future Research

Our study has several limitations. First, our findings cannot be used to establish causality as our data are cross-sectional and our model is correlational. Furthermore, our regression model is moderately complex, which requires careful interpretation. We treated each country as a separate case; therefore, we did not adjust for multiple comparisons. We standardized student achievement, SES, and IL variables separately for each country. Specifically, for the IL variable, we were interested in one standard deviation below and above the country's average; hence, we did not focus on measurement invariance. Future research could utilize longitudinal data and a quasi-experimental design to validate our findings.

Second, our seven countries have very different social, demographic, and economic contexts. Our results show broad patterns regarding the associations between school SES, student SES, student achievement, and IL, which suggests that these associations largely hold true even in very different national contexts. However, we also found contradicting results across countries; it is beyond the capacity of a single study to explain all the contradictions in seven countries. We believe that more in-depth analyses of individual contexts or comparisons of "like" countries would yield more nuanced results. For example, future research could test these associations in countries such as Australia, the US, Canada, the UK, and New Zealand, as well as in countries like those in Scandinavia (Denmark, Finland, Norway, Sweden, and Iceland). Such comparisons are powerful for controlling for societal-level contextual factors when experimentation is not possible (Manzon, 2007).

Conclusion and Implications

Some commentators have suggested that efforts to reduce school social segregation may be a zero-sum game, with gains in achievement for low SES students offset by losses for high SES students. Our results do not support this concern; instead, they point to the possibility of reduced achievement gaps and overall greater achievement, especially when supported by strong instructional leadership. We therefore conclude that education policymakers who are interested in increasing educational effectiveness, efficiency, and equity should reduce school social segregation and promote strong IL.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported by the Alexander von Humboldt Foundation through a Senior Researcher Fellowship to Mehmet Şükrü Bellibaş [TUR – 1227776 – HFST-E] and by the German Research Foundation (DFG) through a DFG Heisenberg Professorship to Marcus Pietsch [451458391].

ORCID

Alex Bowers http://orcid.org/0000-0002-5140-6428

References

Alexiadou, N. (2002). Social inclusion and social exclusion in England: Tensions in education policy. *Journal of Education Policy*, *17*(1), 71–86. https://doi.org/10.1080/02680930110100063

Arıkan, S., Özer, F., Şeker, V., & Ertaş, G. (2020). The importance of sample weights and plausible values in large-scale assessments. *Journal of Measurement & Evaluation in Education & Psychology*, *11*(1), 43–60. https://doi.org/10.21031/epod.602765

Armor, D. J., Marks, G. N., & Malatinszky, A. (2018). The impact of school SES on student achievement: Evidence from US statewide achievement data. *Educational Evaluation and Policy Analysis*, *40*(4), 613–630. https://doi.org/10.3102/0162373718787917

Barranti, M., Carlson, E. N., & Côté, S. (2017). How to test questions about similarity in personality and social psychology research: Description and empirical demonstration of response surface analysis. *Social Psychological & Personality Science*, *8*(4), 465–475. https://doi.org/10.1177/1948550617698204

Bellibaş, M. Ş., Polatcan, M., & Alzouebi, K. (2025). Instructional leadership and student achievement across UAE schools: Mediating role of professional development and cognitive activation in teaching. *Educational management administration & leadership*, 17411432241305702.

Bolletta, U. (2021). A model of peer effects in school. *Mathematical Social Sciences*, *114*, 1–10. https://doi.org/10.1016/j.mathsocsci.2021.09.003

Boman, B. (2023). Is the SES and academic achievement relationship mediated by cognitive ability? Evidence from PISA 2018 using data from 77 countries. *Frontiers in Psychology*, *14*, 1045568. https://doi.org/10.3389/fpsyg.2023.1045568

Bonal, X., & Bellei, C. (2019). The renaissance of school segregation in a context of globalization. In X. Bonal & C. Bellei (Eds.), *Understanding school segregation: Patterns, causes and consequences of spatial inequalities in education* (pp. 1–35). Bloomsbury.

Boucher, V., Rendall, M., Ushchev, P., & Zenou, Y. (2022). *Toward a general theory of peer effects*. SSRN. https://ssrn.com/abstract=4113310

Burdett, N., & O'Donnell, S. (2016). Lost in translation? The challenges of educational policy borrowing. *Educational Research*, *58*(2), 113–120. https://doi.org/10.1080/00131881.2016.1168678

Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics*, *31*(1), 51–82. https://doi.org/10.1086/666653

Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research*, *90*(5), 269–277. https://doi.org/10.1080/00220671.1997.10544583

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. U.S. Department of Health, Education, and Welfare, Office of Education.

Crosnoe, R. (2009). Low-income students and the socioeconomic composition of public high schools. *American Sociological Review*, *74*(5), 709–730. https://doi.org/10.1177/000312240907400502

Darling-Hammond, L. (2014). Closing the achievement gap: A systemic view. In J. V. Clark (Ed.), *Closing the achievement gap from an international perspective* (pp. 7–20). Springer.

Day, C., & Sammons, P. G. K. (2020). *Successful school leadership*. Education Development Trust.

Demie, F. (2023). Tackling educational inequality: Lessons from London schools. *Equity in Education & Society*, *2*(3), 243–266. https://doi.org/10.1177/27526461231161775

Deventer, J., Humberg, S., Lüdtke, O., Nagy, G., Retelsdorf, J., Wagner, J., Donnellan, M. B., & Donnellan, M. B. (2019). Testing competing hypotheses on the interplay of importance and support of the basic psychological needs at work and personality development with response surface analysis. *Collabra: Psychology*, *5*(1), 39. https://doi.org/10.1525/collabra.214

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, *37*(1), 15–24.

Eryilmaz, N., & Sandoval Hernandez, A. (2021). Improving cross-cultural comparability: Does school leadership mean the same in different countries? *Educational Studies*, *50*(5), 917–938. https://doi.org/10.1080/03055698.2021.2013777

Farah, M. J. (2017). The neuroscience of socioeconomic status: Correlates, causes, and consequences. *Neuron*, *96*(1), 56–71. https://doi.org/10.1016/j.neuron.2017.08.034

George, J., & Darling-Hammond, L. (2019). *The federal role and school integration*. Learning Policy Institute.

Gibbons, S., & Telhaj, S. (2016). Peer effects: Evidence from secondary school transition in England. *Oxford Bulletin of Economics and Statistics*, *78*(4), 548–575. https://doi.org/10.1111/obes.12095

Goddard, R. D., Skrla, L., & Salloum, S. J. (2017). The role of collective efficacy in closing student achievement gaps: A mixed methods study of school leadership for excellence and equity. *Journal of Education for Students Placed at Risk (JESPAR)*, *22*(4), 220–236. https://doi.org/10.1080/10824669.2017.1348900

Goddard, R., Goddard, Y., Sook Kim, E., & Miller, R. (2015). A theoretical and empirical analysis of the roles of IL, teacher collaboration, and collective efficacy beliefs in support of student learning. *American Journal of Education*, *121*(4), 501–530. https://doi.org/10.1086/681925

Gümüş, S., Bellibaş, M. Ş., & Pietsch, M. (2022). School leadership and achievement gaps based on socioeconomic status: A search for socially just instructional leadership. *Journal of Educational Administration*, *60*(4), 419–438.

Gustafsson, J. E., Nilsen, T., & Hansen, K. Y. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. *Evidence from 50 Countries in TIMSS 2011 Studies in Educational Evaluation*, *57*, 16–30. https://doi.org/10.1016/j.stueduc.2016.09.004

Guthrie, H. D., & Willower, D. J. (1973). The ceremonial congratulation: An analysis of principals' observation reports of classroom teaching. *The High School Journal*, *56*(6), 284–290.

Gutiérrez, G. (2022). Is it socioeconomic or academic? Disentangling sources of peer effects on student achievement. *British Journal of Sociology of Education*, *44*(1), 144–163. https://doi.org/10.1080/01425692.2022.2137465

Hallinger, P. (2005). IL and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools*, *4*(3), 221–239. https://doi.org/10.1080/15700760500244793

Hallinger, P., & Murphy, J. (1985). Assessing the instructional management behavior of principals. *Elementary School Journal*, *86*(2), 217–247. https://doi.org/10.1086/461445

Harris, D. N. (2010). How do school peers influence student educational outcomes? Theory and evidence from economics and other social sciences. *Teachers College Record*, *112*(4), 1163–1197. https://doi.org/10.1177/016146811011200404

Holzberger, D., Reinhold, S., Luedtke, O., & Seidel, T. (2020). A meta-analysis on the relationship between school characteristics and student outcomes in science and maths-evidence from large-scale studies. *Studies in Science Education*, *56*(1), 1–34. https://doi.org/10.1080/03057267.2020.1735758

Humberg, S., Schönbrodt, F. D., Back, M. D., & Nestler, S. (2022). Cubic response surface analysis: Investigating asymmetric and level-dependent congruence effects with third-order polynomial models. *Psychological Methods*, *27*(4), 622. https://doi.org/10.1037/met0000352

Jeynes, W. H. (2015). A meta-analysis on the factors that best reduce the achievement gap. *Education & Urban Society*, *47*(5), 523–554. https://doi.org/10.1177/0013124514529155

Lai, L., Sun, Q., Zhu, W., & Ren, Z. (2022). The relationship between responsibility attribution and session outcomes: Two-dimension attribution and two-person perspective. *Clinical Psychology & Psychotherapy*, *29*(6), 1928–1941. https://doi.org/10.1002/cpp.2762

Langenkamp, A. G., & Carbonaro, W. (2018). How school socioeconomic status affects achievement growth across school transitions in early educational careers. *Sociology of Education*, *91*(4), 358–378. https://doi.org/10.1177/0038040718802257

Larsen, T. J. (1987, April 20–24). *Identification of IL behaviors and the impact of their implementation on academic achievement* (ERIC Document Reproduction Service No. ED281286) [Paper presentation]. Annual Meeting of the American Educational Research Association, Washington, DC.

Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education*, *6*(1), 1–11. https://doi.org/10.1186/s40536-018-0061-2

Manzon, M. (2007). Comparing places. In M. Bray, B. Adamson, & M. Mason (Eds.), *Comparative education research: Approaches and methods* (pp. 85–122). Springer.

Marks, G. N., & O'Connell, M. (2021). Inadequacies in the SES-achievement model: Evidence from PISA and other studies. *Review of Education*, *9*(3), e3293. https://doi.org/10.1002/rev3.3293

Mitchell, R. M., Kensler, L. A., & Tschannen-Moran, M. (2015). Examining the effects of IL on school academic press and student achievement. *Journal of School Leadership*, *25*(2), 223–251. https://doi.org/10.1177/105268461502500202

Muthén, L. K., & Muthén, B. O. (2021). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nestler, S., Humberg, S., & Schönbrodt, F. D. (2019). Response surface analysis with multilevel data: Illustration for the case of congruence hypotheses. *Psychological Methods*, *24*(3), 291. https://doi.org/10.1037/met0000199

O'Donnell, R. J., & White, G. P. (2005). Within the accountability era: Principals' IL behaviors and student achievement. *NASSP Bulletin*, *89*(645), 56–71. https://doi.org/10.1177/019263650508964505

OECD. (2004). *Education at a glance*. OECD publications.

OECD. (2009). *PISA 2006 technical report*. OECD publications.

OECD. (2016). *Education at a glance*. OECD Publications.

OECD. (2019a). *PISA 2018 technical report*. OECD publications.

OECD. (2019b). *TALIS technical report*. OECD Publications.

Paloyo, A. R. (2020). Peer effects in education: Recent empirical evidence. In S. Bradley & C. Green (Eds.), *The economics of education* (pp. 291–305). Academic Press.

Perry, L. B. (2007). *School composition and student outcomes: A review of emerging areas of research* [Paper presentation]. Australian Association for Research Education Conference. Freemantle, Western Australia.

Perry, L. B., Saatcioglu, A., & Mickelson, R. A. (2022). Does school SES matter less for high-performing students than for their lower-performing peers? A quantile regression analysis of PISA 2018 Australia. *Large-Scale Assessments in Education*, *10*(1), 1–29. https://doi.org/10.1186/s40536-022-00137-5

Pietsch, M., Aydin, B., & Gümüş, S. (2023). Putting the instructional leadership-student achievement relation in context: A meta-analytical big data study across cultures and time. *Educational Evaluation and Policy Analysis*, *47*(1), 29–64. https://doi.org/10.3102/01623737231197434

Robinson, V. M., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, *44*(5), 635–674. https://doi.org/10.1177/0013161X08321509

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

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Chapman & Hall.

Schindler Rangvid, B. (2003). *Do schools matter? The influence of school inputs on student performance and outcomes*. Aarhus School of Business.

Schneider, S., Lösch, T., Schneider, D., Schütz, A., & Donnellan, B. (2022). How self-concept, competence, and their fit or misfit predict educational achievement, well-being, and social relationships in the school context. *Collabra: Psychology*, *8*(1), 37154. https://doi.org/10.1525/collabra.37154

Schönbrodt, F. D., & Humberg, S. (2023). *RSA: An R package for response surface analysis* (Version 0.10.6). https://cran.r-project.org/package=RSA

Sebastian, J., Allensworth, E., Wiedermann, W., Hochbein, C., & Cunningham, M. (2019). Principal leadership and school performance: An examination of IL and organizational management. *Leadership and Policy in Schools*, *18*(4), 591–613. https://doi.org/10.1080/15700763.2018.1513151

Shanock, L. R., Baran, B. E., Gentry, W. A., Pattison, S. C., & Heggestad, E. D. (2010). Polynomial regression with response surface analysis: A powerful approach for examining moderation and overcoming limitations of difference scores. *Journal of Business & Psychology*, *25*(4), 543–554. https://doi.org/10.1007/s10869-010-9183-4

Silver, H. (1994). Social exclusion and social solidarity: Three paradigms. *International Labour Review*, *133*(5/6), 531–578.

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

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis : An introduction to basic and advanced multilevel modeling* (2nd ed.). SAGE.

Tan, C. Y., Hong, X., Gao, L., & Song, Q. (2023). Meta-analytical insights on school SES effects. *Educational review*, Advance online publication.

Tan, C. Y., Hong, X., Gao, L., & Song, Q. (2025). Meta-analytical insights on school SES effects. *Educational Review*, *77*(1), 274–302.

Teddlie, C., & Stringfield, S. (1985). A differential analysis of effectiveness in middle and low socioeconomic status schools. *The Journal of Classroom Interaction*, *20*(2), 38–44. https://www.jstor.org/stable/43997876

Teddlie, C., Stringfield, S., & Reynolds, D. (2002). Context issues within school effectiveness research. In C. Teddlie & D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 174–200). Routledge.

Thrupp, M., Lauder, H., & Robinson, T. (2002). School composition and peer effects. *International Journal of Educational Research*, *37*(5), 483–504. https://doi.org/10.1016/S0883-0355(03)00016-8

van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, *5*(2), 134–150. https://doi.org/10.1016/j.edurev.2010.02.001

Wilkinson, I. A. (2002). Introduction: Peer influences on learning: Where are they? *International Journal of Educational Research*, *37*(5), 395–401. https://doi.org/10.1016/S0883-0355(03)00012-0

Yang, Y., & Gustafsson, J. E. (2004). Measuring socioeconomic status at individual and collective levels. *Educational Research & Evaluation*, *10*(3), 259–288. https://doi.org/10.1076/edre.10.3.259.30268

Zay, D. (2005). Preventing school and social exclusion: A French-British comparative study. *European Educational Research Journal*, *4*(2), 109–120.

Zechman, H. T. (1977). Are principals competent in the IL domain? *NASSP Bulletin*, *61*(413), 21–25. https://doi.org/10.1177/019263657706141304

Zyphur, M. J., Zammuto, R. F., & Zhang, Z. (2016). Multilevel latent polynomial regression for modeling (in) congruence across organizational groups: The case of organizational culture research. *Organizational Research Methods*, *19*(1), 53–79. https://doi.org/10.1177/1094428115588570