Karakus *et al.*
Large-scale Assessments in Education (2024) 12:42
https://doi.org/10.1186/s40536-024-00228-5

Large-scale Assessments in Education

RESEARCH

Open Access

Illuminating the shadows: the role of private supplementary tutoring on student math performance in PISA 2022

Mehmet Karakus[1*], Alisher Tlessov[2], Anas Hajar[2] and Matthew Courtney[3]

*Correspondence: mehmet.karakus@coventry.ac.uk

[1] Research Centre for Global Learning, Coventry University, 5 Whitefriars Street, Coventry CV12DS, UK
[2] Centre for Multilevel Modeling, University of Bristol, Bristol, Kazakhstan
[3] Research & Analytics, Nazarbayev University, Astana, Kazakhstan

Abstract

Recent decades have witnessed a global expansion of private supplementary tutoring, known as shadow education—additional support in academic subjects that takes place outside of regular school hours. Using the data of 55 countries from the Programme for International Student Assessment (PISA) 2022, this study examines the association between students' participation in shadow education and their academic achievement in mathematics and how these relationships change for developed, developing, and East Asian regions. This study also sheds light on the moderating effect that socio-economic status has on the relationship between engagement in shadow education and math performance. Multilevel regression analyses revealed that, at the between-school level, only one of the five forms of shadow education, namely the degree to which students received *asynchronous* video-recorded instruction, was associated with improved math performance. Similarly, at the within-school level, only the degree to which students received video-recorded instruction was statistically significantly associated with improved mathematical outcomes. Incidentally, we find that institutional-related school-level factors such as the overall quality of student–teacher relationships, provision of math-related extra-curricular activities, and support in terms of peer-to-peer tutoring appear to be associated with improved school math performance, while the quality of student–teacher relationships and level of math teacher support also feature positively for students within-schools. Regarding differential effects by region, at the between-school level, only the degree to which students received large-group study or practice was associated with math performance in East Asian countries. With regards to the moderating effect of SES on the positive relationship between *asynchronous* shadow education and math performance, we find the effects to be negative, suggesting that *asynchronous* engagement might be more beneficial for students from lower SES backgrounds. Finally, while the overall negative relationships between synchronous forms of shadow education and math performance appear counter-intuitive, we speculate that this may reflect the use of such tutoring for remedial purposes. We suggest that further research into these "remedial reverse effects" in post-COVID AI-bot-capable educational contexts may provide a more definitive understanding of the role that student engagement in shadow education has on their academic performance.

Keywords: Shadow education, Academic performance, Mathematics achievement, Socioeconomic status, PISA 2022, Multilevel analysis

Introduction

Mainstream schools typically adhere to a recognized curriculum and are generally regarded as the primary means through which societies educate their youth. However, there are alternative educational options available globally. One of the main options is shadow education, which is additional support in academic subjects and takes place outside of regular school hours (Bray, 2023). This definition does not include school-managed extra-curricular classes or programs related to culture or sports. Nevertheless, shadow education corresponds to and has a strong relationship with mainstream education and encompasses a wide range of educational activities undertaken outside of formal schooling, including private tutoring, enrichment programs, and extra-curricular activities (Kobakhidze & Suter, 2020). In terms of instructional mode, shadow education can be delivered one-to-one, in small groups, in large classes, and over the Internet (Zhang, 2023). This phenomenon has received significant attention due to its impact on the upbringing of new generations, economic growth, the functioning of formal education systems, and cultural and social development (Hajar & Karakus, 2024). This is evident from the books published on the topic, covering various regions, including Africa (Bray, 2021), Central Asia (Silova, 2009), East Asia (Kim & Jung, 2022), Europe (Bray, 2011), and the Middle East (Bray & Hajar, 2023).

The implications of shadow education have also been recognized in a report by the United Nations Special Rapporteur on Human Rights (Singh, 2015) and UNESCO's (2021) Global Education Monitoring Report on non-state actors in education. As shadow education has expanded, it has further attracted corresponding attention from large-scale international surveys such as the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA) (e.g., Byun et al., 2018; Entrich, 2021; Liang et al., 2022; Wiseman, 2021). For example, Wiseman (2021) utilized data from the PISA 2012 to disentangle the effect of shadow education on students' academic achievement. The findings revealed that students' involvement in three types of shadow education (commercial company, personal, and family tutoring) was negatively associated with student achievement in the cross-national sample. However, the evidence suggests that the impact of shadow education on student achievement may vary depending on the student and national-level educational equity factors as well as different contexts (Wiseman, 2021). Entrich (2021), in turn, used data from the 2012 PISA and official sources to investigate the differences in socioeconomic accessibility to shadow education across 63 societies. The results indicated that differences in access to shadow education are closely related to the incentives for high-performing students to engage in it, especially in societies with more differentiated education systems, such as early or mixed-tracking schooling models. In such societies, where institutional stratification is higher, high-performing students may be more motivated to seek additional academic support through shadow education to maintain or enhance their academic standing. However, this does not imply a direct causal relationship between shadow education and improved performance; rather, it suggests that those already inclined towards higher achievement are more likely to participate in shadow education, potentially introducing selection bias into the findings.

In societies with less stratified education systems, access to shadow education tends to be more equitable, reducing its potential to exacerbate social inequality. In these settings, where students face less pressure to differentiate themselves through shadow education, the impact of supplementary education may be less pronounced. This highlights the importance of considering the broader educational context when evaluating the effects of shadow education. Furthermore, Byun et al. (2018) explored how school-level factors, such as the quality of student–teacher relationships and the availability of extra-curricular activities, contribute to creating supportive school environments. These factors may also play a role in student achievement, offering an additional layer of understanding to the complex relationship between shadow education and student outcomes.

The present study aims to delve deeper into the complex dynamics of shadow education and its association with student academic achievement. By examining the moderating effects of socioeconomic status, the role of the school environment, and the effectiveness of various supplementary education programs, this study seeks to contribute to a comprehensive understanding of how shadow education influences educational outcomes. Through the analysis of PISA 2022 data, this study endeavours to inform educational policymakers, practitioners, and researchers about how student engagement in shadow education is associated with school and student academic performance and how these associations might be influenced by region and student socio-economic status.

Theoretical considerations

The emergence of shadow education and its impact on academic achievement

The practice of shadow education probably dates back as far as formal schooling itself (Zhang, 2023). For decades and even centuries, parents have used it as a means to improve their children's academic performance (Zhang, 2023; Zhang & Bray, 2020). In Japan, for instance, the first juku (an academic tutoring enterprise) was opened in 1911 by a teacher in Tokyo. The parents of his former students asked him to provide paid tutoring to their children to help them advance to lower secondary schools (Sato, 2012). In Russia, Mikhaylova (2019) identified advertisements by private tutors in the mid-nineteenth century. In South Korea, the earliest *hagwons* were established in 1903 by the Hwang-Sang Young Men's Christian Association to develop Korean awareness of Western culture and provide additional support in English and mathematics (Kim, 2016, p. 15). In the Middle East, Bray and Hajar (2023) note that shadow education has been prevalent since the 1960s and 1970s (e.g., Junaidi, 1977; Kuwait, 1962). However, shadow education only became a specific topic in academic literature during the 1990s (e.g., Bray, 1999; Marimuthu et al., 1991; Stevenson & Baker, 1992).

According to Bray and Hajar (2023: 36), in most countries, the subjects in high demand for tutoring are mathematics and languages (especially English). In all education systems worldwide, the fundamental driver of demand for shadow education is social competition (Hajar & Karakus, 2022). Improved academic performance in schooling is the main driver for this competition, and most types of shadow education aim to achieve high—or at least adequate—scores in examinations. Zwier et al. (2020) suggest that education systems with high-stakes assessments at key points are more likely to have a high incidence of shadow education. However, Bray and Hajar (2023) posit that shadow education can be widespread during low-stakes assessments throughout the academic year, as these assessments tend to impact students' peer dynamics, self-confidence, and teachers' attitudes towards students.

Shadow education is also facilitated by neoliberal and consumerist features that have challenged the earlier socialist notions of education in society (Hajar & Tabaeva, 2024; Verger et al., 2016). Moreover, shadow education experienced significant growth during the COVID-19 pandemic due to the closure of schools worldwide (Hajar & Karakus, 2023a; Lee et al., 2023). In many countries, face-to-face tutoring was also put on hold, prompting the owners of tutorial centres to transition online more quickly than traditional school systems. Additionally, shadow education played a crucial role in helping students catch up on missed learning during the closures. As a result, the pandemic considerably boosted the volume of shadow education, solidifying its position in many contexts.

Concerning the impact of shadow education on students' academic achievement, Guill et al. (2019) found that the empirical evidence is contradictory. Some researchers (e.g., Berberoğlu & Tansel, 2014; Domingue & Briggs, 2009; Han & Suh, 2020; Kuan, 2011) have reported a positive influence in certain subjects, while others have found minimal or no effects (e.g., Cheo & Quah, 2005; Park et al., 2016; Ryu & Kang, 2013; Sung & Kim, 2010; Wiseman, 2021). Even studies within the same country or region exhibited contradictory outcomes, though this may have been due to variations in research designs, statistical methods, student populations, and tests used (e.g., Choi et al., 2012; Ryu & Kang, 2013 in Korea; Kuan, 2011; Liu, 2012 in Taiwan). As noted by Bray (2014) and Byun (2014), the empirical evidence regarding shadow education's impact on students' academic achievement has been inconsistent, contradictory, and even confusing. Therefore, some researchers (e.g., Bray & Hajar, 2023; Kim & Lee, 2010) highlight the importance of considering various factors in the analysis of such effects, including the quality of tutoring, students' motivation and socioeconomic status, and educational contexts and policies.

Understanding the differential impacts of these various shadow education modalities is crucial for educators, policymakers, and researchers aiming to understand the relationship between shadow education and student academic achievement. Therefore, this study aims to explore the effects of different types of shadow education on students' mathematics achievement, examining both global trends and specific regional variations. By analyzing recent data and considering the unique educational challenges posed by the pandemic, this research seeks to provide nuanced insights into the role that student engagement in shadow education might have on their performance in mathematics.

Socioeconomic status and the use of shadow education

The relationship between socioeconomic status (SES) and educational outcomes has long been a focus of educational research (Entrich, 2021; Jansen et al., 2021). Socioeconomic disparities often translate into unequal access to educational resources, significantly impacting student performance. According to UNESCO's (2017) Sustainable Development Goals, one of the targets is ensuring that everyone has access to high-quality education and lifelong learning opportunities by 2030. However, Zhang (2023) postulates that achieving this goal may be challenging without addressing and regulating the shadow education market, which is often exclusive and inequitable. Bray (2024, pp. 2–3) points out that in the past, students from disadvantaged backgrounds were less likely to complete secondary education, let alone post-secondary education. However, the push for universal primary education led to demands for universal lower-secondary education and expanded upper-secondary and post-secondary education. This expansion has made upper-secondary and post-secondary education more accessible to students from disadvantaged backgrounds who previously faced barriers (Bray, 2024). Bray (2024) further argued that education systems remain stratified despite this progress, shifting the focus from whether education is attainable at each level to which institutions can be accessed. As a result, Bray (2024) argues that shadow education is still necessary to achieve high marks on high-stakes exams to gain entry into highly selective universities.

The impact of shadow education on educational inequality has sparked heated debates (Hajar & Karakus, 2023b; Jansen et al., 2021; Wiseman, 2021). Evidence from various countries suggests that shadow education can exacerbate educational disparities, and students from higher-income families benefit disproportionately from shadow education because these families can afford more and higher-quality tutoring than low-income families (Bray et al., 2020; Byun & Park, 2012; Gupta, 2021; Hajar & Karakus, 2023b; Holloway & Kirby, 2019; Ku et al., 2022; Liao & Huang, 2018; Marshall & Fukao, 2019). Contrarily, in contexts where shadow education is deeply integrated into the educational system, such as South Korea, the disparities related to SES are less pronounced. Lee and Shouse (2011) observed that in South Korea, shadow education's pervasive and normalized nature diminishes SES-related disparities associated with its effectiveness. This suggests that country-specific cultural attitudes significantly influence the impact of shadow education. Therefore, understanding the role that these contextual factors play may be useful for understanding shadow education in different regions.

The present study explores the moderating effect of SES on the relationship between student engagement in shadow education and mathematics achievement, shedding light on how shadow education might mitigate or exacerbate educational inequalities. Shadow education can be a crucial tool for bridging educational gaps and enhancing academic performance for socioeconomically disadvantaged students. This notion is supported by Kim's (2016) study in South reporting that shadow education could mitigate educational inequality by providing customized support to struggling students, often at affordable rates. Thus, while shadow education can potentially increase educational disparities in some contexts, it can also provide essential support to disadvantaged students, depending on the surrounding cultural and economic conditions.

The aim and significance of the study

This study addresses some gaps in the existing literature on shadow education and its association with student academic achievement, particularly in mathematics. Utilizing the most recent PISA 2022 dataset, this study comprehensively examines the relationship between student engagement in shadow education and mathematics performance across diverse contexts, including global, developed, developing, and East Asian countries. By incorporating a robust set of control variables that span personal and family backgrounds, SES, provisions and support inside and outside school, and study habits, this study offers a nuanced understanding of the factors influencing math achievement beyond the scope of previous research.

We have primarily focused on mathematics achievement in this study due to the emphasis placed on mathematics in the PISA 2022 cycle. PISA 2022 specifically measured only mathematics-related private tutoring, allowing for a more precise analysis of the association of shadow education with math performance. This singular focus on mathematics enables us to delve deeply into the factors influencing math achievement, providing targeted insights that can inform educational strategies and policies aimed at improving math outcomes for students globally.

One significant contribution of this study is its focus on the moderating role of SES in the relationship between shadow education and math performance. Previous studies have often acknowledged the influence of SES on educational outcomes (e.g., Broer et al., 2019; Harwell et al., 2016), but less attention has been paid to how SES interacts with the relationship between student engagement in shadow education and student achievement (e.g., Entrich, 2021; Wiseman, 2021). Using data from the 2022 PISA cycle, this study sheds light on the differential benefits of shadow education for students from varying regional backgrounds, providing valuable insights for policymakers aiming to understand the role that shadow education plays in student academic development.

Furthermore, this research is noteworthy for its examination of the impact of various modes of shadow education, including one-on-one tutoring, Internet/computer tutoring, video-recorded instruction, and small and large group study sessions. The study aims to examine the role that these modes of shadow education have across different educational settings. This comparative analysis is especially relevant in the post-COVID-19 era, as digital and asynchronous learning tools have become more prevalent (Lee et al., 2023). As highlighted by Hajar and Karakus (2023a), the effectiveness and intensity of online private tutoring require further investigation, mainly because both tutors and students have become more accustomed to this type of tutoring after many countries' government bodies banned all forms of face-to-face tutoring to contain the spread of COVID-19.

Additionally, the inclusion of a wide range of control variables enhances the robustness and validity of our findings. Factors such as the quality of student–teacher relationships, school provision of mathematics-related extra-curricular activities, and support systems at home and school have been shown to influence academic outcomes (e.g., Bray & Hajar, 2023; Gao et al., 2023; Han & Suh, 2020). By considering these variables, this study offers a valuable perspective on the relationship between shadow education and student mathematics performance. However, the findings should be interpreted with caution, as the absence of measures for prior academic achievement or cognitive ability limits our ability to fully account for all potential confounding factors.

The originality of this study also lies in its multi-contextual analysis, which considers developed, developing, and East Asian countries separately. This approach recognizes the significant cultural and systemic differences that can affect the implementation and effectiveness of shadow education. By doing so, the present study contributes to a more regional-based understanding of shadow education practices and their varying impacts, offering tailored recommendations for different educational contexts.

In summary, this study significantly contributes to educational research by filling crucial gaps in the literature on shadow education and math achievement. Through its comprehensive and multi-contextual analysis, consideration of socio-economic status as a moderating factor, and inclusion of a wide array of control variables, this research provides valuable insights that can inform both policy and practice. The findings have the potential to guide educators, policymakers, and researchers in developing more effective and equitable educational interventions, ultimately enhancing student outcomes across diverse settings. The following research questions were specified to achieve the study's objectives and guide the analyses:

RQ1: What degree of variance in students' math ability can be attributable to within-school, between-school, and between-country effects?
RQ2: What is the relationship between student engagement in shadow education and math performance for all PISA countries?
RQ3: What is the relationship between student engagement in shadow education and math performance for (a) developed countries, (b) developing countries, and (c) East Asian countries?
RQ4: What is the within-school moderation effect of socio-economic status on the relationship between student engagement in shadow education and math performance for (a) all PISA countries, (b) developed countries, (c) developing countries, and (d) East Asian countries?

Methods

Participants

Participants for the current study were drawn from the most recent PISA 2022 dataset and included 274,888 students nested in 9887 schools nested in 58 countries (see Appendix A for a list of countries). For RQ3, countries were assigned to the developed (24 countries), developing (34 countries), and East Asian (six states) subsamples based on classifications by the United Nations (n.d.).

Sampling

Sampling for PISA involved a two-stage stratified sampling design. In the first stage, whole schools are sampled using probability proportional to size sampling (based on the number of PISA-eligible 15-year-old students). In the second stage, the students within the sampled schools were selected at random (for full details on target cluster sizes, see the PISA 2022 technical report Chapter 6, OECDa, 2022).

Data preparation

The school and student data[1] were downloaded from the OECD website (OECDb, 2022) and were read into the R software with the assistance of the haven package (Wickham et al., 2023). As one of the focal areas of this study is the amount of variance attributable to school and country effects, we decided to remove schools with fewer than ten students (Hox & McNeish, 2020). Cases were also removed listwise due to missing

[1] CY08MSP_SCH_QQQ.sav school and CY08MSP_STU_QQQ.sav student files.

data.[2] It should also be noted that while gender was modelled as a between-school variable, it was not modelled as a within-school variable, so single-sex schools could also be included in the study. For this study, we examined the effect of some student-level variables at both within- and between-school levels. To explore the pure within-group effects of these variables, each variable was group-mean centered and modelled at the within-school level (Huang, 2022). Simultaneously, to explore the between-group effects of these variables, we aggregated up the original variables to generate group means for each school and modelled these effects at the between-school level (see Huang, 2022).

Descriptive statistics

Descriptive statistics for the independent and dependent variables are provided in Table 1 (readers interested in student performance by specific countries are encouraged to use IEA's IDB Analyzer[3]). Adjustments to variables exhibiting raw skewness beyond |2.00| were made with the assistance of the [blinded for review] R package (Courtney & Chang, 2018). All presented mean, *SD* and skewness statistics were generated with the assistance of the TAM package's weighted functions (Robitzsch et al., 2022) using adjusted senate weights (see subSect. "Analysis" below for details).

Analysis

All data preparation and analysis were performed in the R programming language (R Core Team, 2023). Multilevel models, also known as mixed-effects models or hierarchical regression models, were used due to their prevalence in examining nested and cross-sectional data (Finch et al., 2019; Merlo et al., 2018). For all null and fixed effects models, the lme4 package (Bates et al., 2015) was used with the optimx package optimizer (Nash, 2014). All three-level multilevel models were run 10 times for each of the 10 plausible values to accurately estimate the mean coefficients in accordance with Huang (2024). Each set of results was pooled to ensure that the final reported coefficients account for the variability introduced by using multiple plausible values (Huang, 2024). Due to missing data, the original senate weights were adjusted (i.e., multiplied by a constant specific to each country) prior to modelling so that the total weights for each country summed to 5000. Senate weights were chosen so that the contribution from all PISA participating countries was equivalent despite the differing sample sizes of students (Courtney et al., 2023; Jerrim et al., 2017). For RQs 2 and 3, we also undertook sensitivity analyses to check for the robustness of the main results. For this, we ran additional models with a reduced set of control variables at the between-school level (i.e., ESCS, gender, and school type) and at the within-school level (i.e., ESCS). Our focus was to check for the stability of the shadow-education-related coefficients at *both* levels in the final models. To this end, instances when (1) the direction of the coefficients changed or (2) the associated statistical significance ($p < 0.05$) was attained or lost were highlighted (see Appendix A, Table A2, notes). In addition, tests for collinearity between the focal shadow education variables are presented in Table A3. We ran this model with the assistance of the lavaan package (Rosseel, 2012) and ran a two-level model specifying all correlational

[2] Due to missing data, the total PISA sample size of 613,744 students, nested in 21,629 schools, nested in 80 countries was reduced to 274,888 students, nested in 9,887 schools, nested in 58 countries.

[3] See the following YouTube for instructions: https://www.youtube.com/watch?v=2nczoGDLGXc&t=241s.

Table 1 Descriptive statistics for dependent and independent variables

Abbreviation	Variables (coding)	*M*	*SD*	Skew/Ratio
MATH_PV	Plausible values for mathematics	451.24	99.85	0.33
SCHLTYPE	School type (private = 0, public = 1)	0.79	0.40	21:79
SC212Q01JA	Study help for students: Room(s) where the students can do their homework (yes = 1, no = 0)	0.64	0.48	36:64
SC212Q02JA	Study help for students: Staff help with homework (yes = 1, no = 0)	0.58	0.49	39:61
SC212Q03JA	Study help for students: Peer-to-peer tutoring (yes = 1, no = 0)	0.6	0.49	36:64
ST004D01T	Student (Standardized) Gender (male = 1, female = 0)	0.48	0.50	52:48
REPEAT	Grade repetition (repeated at least once = 1, else = 0)	0.07	0.25	94:6
MISSSC	Missing school for more than 3 months (Missed school for > 3 months at least once = 1, else = 0)	0.08	0.27	92:8
MATHEXC_NO_DIF[a]	Mathematics extension courses offered without differentiation[1] (yes = 1, no = 0)	0.02	0.14	98:2
MATHEXC_EN_O_REM[a]	Mathematics extension courses offered for enrichment or remediation (yes = 1, no = 0)	0.2	0.4	81:19
MATHEXC_EN_A_REM[a]	Mathematics extension courses offered for enrichment and remediation (yes = 1, no = 0)	0.42	0.49	54:46
MACTIV	Mathematics-related extra-curricular activities at school (0 = none, 5 five or more)	2.65	1.57	0.00
TDTEST	Use of teacher-developed tests (WLE)	0.23	1.08	− 0.4
STDTEST	Use of standardized tests (WLE)	0.27	1.07	− 0.15
ST297Q01JA	[Additional math instruction] received: One-on-one tutoring with a person (checked = 1, not checked = 0)	0.27	0.45	1.02
ST297Q03JA	[Additional math instruction] received: Internet or comp. tutoring with a program/application (checked = 1, not checked = 0)	0.25	0.43	1.15
ST297Q05JA	[Additional math instruction] received: Video-recorded instruction by a person (checked = 1, not checked = 0)	0.24	0.43	1.21
ST297Q06JA	[Additional math instruction] received: Small group study or practice (2–7 students) (checked = 1, not checked = 0)	0.23	0.42	1.26
ST297Q07JA	[Additional math instruction] received: Large group study or practice (8 or more students) (checked = 1, not checked = 0)	0.16	0.36	1.88
WORKHOME	Working in household/take care of family members before or after school (zero per week = 0; 10 or more times per week = 10)	5.53	3.53	− 0.08
WORKPAY[b]	Working for pay before or after school (zero work for pay per week = 0; 10 or more times working for pay per week = 10)	1.37	2.71	0.97
EXERPRAC	Exercise or practice a sport before or after school (no exercise or sports = 0; 10 or more times… = 10)	4.64	3.57	0.23
ST296Q01JA	How much time spent on homework in: Mathematics homework (up to 30 min per day = 0; more than 4 h per day = 6)	1.99	1.19	1.35
RELATST	Quality of student–teacher relationships (WLE)	0.07	1.04	0.20
TEACHSUP	Mathematics Teacher Support (WLE)	0.12	1.16	− 0.40
ESCS	Index of economic, social and cultural status	− 0.25	1.08	− 0.54
RELATST	Quality of student–teacher relationships (WLE)	0.07	1.04	0.2
STUDYHMW	Studying for school or homework before or after school	5.98	3.22	− 0.2

All statistics at the student level; MATH_PV = average of 10 plausible values

[1] Student admission to extension program only dependent on prior level of achievement

[a] Dummy coded based on the original categorical variable (MATHEXC) with NAs coded as zero; *M* = weighted mean; *SD* = weighted standard deviation; Skew = weighted skewness; the ratio of 1 to 0 for dichotomous variables only

[b] Normalized variable by way of optimal lambda implemented (see Courtney & Chang, 2018); WLE = Warm's (1989) likelihood estimate based on developed IRT-scale; "checked: refers to instances when the student provided a positive response

relationships between the shadow education variables (senate weights applied). RQ4 involved testing the interaction effect of student SES on the relationships between engagement in shadow education and math performance (at the student level). This involved testing the moderation effect of five shadow education variables for each of the four country classifications (5×4). Consequently, 20 *separate* models (each involving the 10 PVs and adjusted senate weights, and the single additional interaction effect of interest) were run to address this research question. Equation 1 provides a representation of the formula for testing the interaction effect of student SES on the relationship between student engagement in the *first listed* shadow education variable (i.e., ST297Q05JA, "student received video-recorded instruction from a person") and math performance:

$$\begin{aligned} Math\,Performance_{ijk} = {} & \gamma_{000} + (\gamma_{010})STUDYHMW_{0jk} + \cdots + (\gamma_{020})ST297Q05JA_{0jk} + \ldots \\ & + (\gamma_{030})RELATST_{0jk} \\ & + \cdots + (\gamma_{040})ST004D01T_{0jk} + \cdots + (\gamma_{100})STUDYHMW_{ijk} \\ & + \cdots + [(\gamma_{200})ST297Q05JA_{ijk} \times ESCS_{ijk}] \\ & + (\gamma_{300})ST297Q05JA_{ijk} + (\gamma_{400})ST297Q01JA_{ijk} \\ & + (\gamma_{500})ST297Q07JA_{ijk} \\ & + (\gamma_{600})ST297Q03JA_{ijk} + (\gamma_{700})ST297Q06JA_{ijk} \\ & + (\gamma_{800})RELATST_{ijk} + \cdots + (\gamma_{900})ESCS_{ijk} \\ & + \cdots + \upsilon_{00k} + \nu_{0jk} + \varepsilon_{ijk} \end{aligned} \tag{1}$$

where, $MathPerformance_{ijk}$ is the math performance of student i in school j in country k; γ_{000} is the fixed intercept; γ_{010} is the school-level coefficient associated with the $STUDYHMW_{0jk}$ variable; the first "$+\ldots$" is the set of associated coefficients and variables listed under the "Personal Study" category in Table 2; γ_{020} is the school-level coefficient associated with the $ST297Q05JA_{0jk}$ variable; and the second "$+\ldots$" is the set of coefficients and variables listed under the "Shadow Education" category in Table 2…; γ_{200} is the focal coefficient of interest for the single-tested interaction effect for the model; υ_{00k} is the country-level residual; ν_{0jk} is the school-level residual, and ε_{ijk} is the student-level residual.

Finally, with the trend toward more strict interpretations of statistical significance (Benjamin et al., 2017) and the large sample size of the PISA studies, we use a threshold of $p < 0.01$ to interpret coefficients as substantive at the within- and between-school levels. Note that the R code is available upon reasonable request from Dr Matthew Courtney (matthew.courtney@nu.edu.kz).

Results

With regard to RQ1, i.e., the degree of variance in student performance attributable to within-school, between-school, and between-country effects, we found that 49.5% could be attributable to within-school effects, 23.6% could be attributable to between-school effects, and 26.8% could be attributable to between-country effects.

Research Question 2 focused on the relationship between student engagement in shadow education and math performance for all PISA countries. For this question, we found mixed results (Table 2, "Global"). Globally, at the between-school level, only one of the five shadow education related variables was positively associated with student

Table 2 Multilevel models for the role of shadow education on student math ability

Item code	Predictors (coding)	Global	Developed	Developing	East Asian
–	Intercept	440.74****	455.67****	423.38****	464.67****
Between-School Effects					
Personal Study					
STUDYHMW	Studying for school or homework before or after school	**27.31******	**18.19******	**26.93******	**52.99******
ST296Q01JA	How much time spent on homework in: Mathematics homework	7.23****	5.54**	11.99****	17.17****
Shadow Education ["Additional math instruction"]					
ST297Q05JA	Student received video-recorded instruction from a person [asynchronous]	**4.92****	**6.48****	4.12*	11.07
ST297Q03JA	Student received Internet or computer tutoring with a programme or application [synchronous]	**− 11.77******	− 4.47*	**− 13.76******	− 11.09
ST297Q01JA	Student received one-on-one tutoring with a person [synchronous]	**− 11.16******	**− 17.71******	**− 9.38******	**− 30.76******
ST297Q07JA	Student received large group study or practice (8 or more students) [synchronous]	**− 5.21*****	**− 14.50******	− 0.78	***17.26******
ST297Q06JA	Student received small group study or practice (2 to 7 students) [synchronous]	− 1.93	− 1.70	− 1.10	− 5.51
Institutional Related Factors					
RELATST	Quality of student–teacher relationships (WLE)	**9.40******	**10.50******	**6.31******	3.2
MACTIV	Mathematics-related extra-curricular activities at school	**5.83******	**6.96******	**4.58******	4.28*
SC212Q03JA	Study help for students: Peer-to-peer tutoring (yes = 1, no = 0)	**2.61*****	**2.56****	1.01	3.73
TEACHSUP	Mathematics Teacher Support (WLE)	2.57*	2.95	3.39*	12.07*
SC212Q01JA	Study help for students: Room(s) where the students can do their homework (yes = 1, no = 0)	1.71*	2.75*	0.70	− 0.55
MISSSC	Missing school for more than 3 months (Missed school for > 3 months at least once = 1, else = 0)	**− 57.39******	**− 53.14******	**− 61.85******	**− 54.26****
REPEAT	Grade repetition (repeated at least once = 1, else = 0)	**− 33.36******	**− 60.71******	**− 22.02******	**− 81.19******
MATHEXC_EN_O_REM	Mathematics extension courses offered for enrichment or remediation (yes = 1, no = 0)	**− 5.65******	**− 6.13******	**− 3.90****	− 6.07
MATHEXC_NO_DIF	Mathematics extension courses offered without differentiation[1] (yes = 1, no = 0)	− 5.32*	− 5.38*	− 7.62*	− 18.75
MATHEXC_EN_A_REM	Mathematics extension courses offered for enrichment and remediation (yes = 1, no = 0)	**− 4.90******	**− 8.67******	− 1.00	− 3.19
SC212Q02JA	Study help for students: Staff help with homework	**− 3.99******	**− 4.40******	− 1.53	− 4.41

Table 2 (continued)

Item code	Predictors (coding)	Global	Developed	Developing	East Asian
STDTEST	Use of standardized tests (WLE)	**<u>− 1.44***</u>**	**<u>− 1.71**</u>**	− 1.04*	− 1.55
TDTEST	Use of teacher-developed tests (WLE)	**<u>− 1.35***</u>**	− 0.96	**<u>− 1.41**</u>**	0.01
Controls					
ST004D01T	Student gender (male = 1, female = 0)[a]	**<u>33.10****</u>**	**<u>29.43****</u>**	**<u>35.89****</u>**	**<u>44.88****</u>**
ESCS	Index of economic, social and cultural status	**<u>52.07****</u>**	**<u>70.80****</u>**	**<u>42.74****</u>**	**<u>48.56****</u>**
WORKHOME	Working in household/take care of family members before or after school	**<u>− 27.86****</u>**	**<u>− 26.05****</u>**	**<u>− 24.85****</u>**	**<u>− 47.10****</u>**
WORKPAY	Working for pay before or after school	**<u>− 21.85****</u>**	**<u>− 19.28****</u>**	**<u>− 27.88****</u>**	**<u>− 28.20****</u>**
EXERPRAC	Exercise or practice a sport before or after school	**<u>− 21.19****</u>**	**<u>− 21.38****</u>**	**<u>− 21.78****</u>**	**<u>− 34.39****</u>**
SCHLTYPE	School type (private = 0, public = 1)	**<u>6.38****</u>**	**<u>5.85****</u>**	2.25	**<u>11.84****</u>**
Within-school effects					
Personal study					
STUDYHMW	Studying for school or homework before or after school$_{gmc}$	**<u>− 0.45**</u>**	**<u>− 3.54****</u>**	***<u>1.86****</u>***	***<u>6.08****</u>***
ST296Q01JA	How much time spent on homework in: Mathematics homework$_{gmc}$	**<u>− 0.64****</u>**	**<u>− 2.70****</u>**	***<u>0.43**</u>***	***<u>4.12****</u>***
Shadow education ["Additional math instruction"]					
ST297Q05JA	Student received video-recorded instruction from a person [asynchronous]	**<u>0.98****</u>**	0.20	**<u>1.48****</u>**	0.98
ST297Q01JA	Student received one-on-one tutoring with a person [synchronous]	**<u>− 9.15****</u>**	**<u>− 10.53****</u>**	**<u>− 8.17****</u>**	**<u>− 10.50****</u>**
ST297Q07JA	Student received large group study or practice (8 or more students) [synchronous]	**<u>− 2.38****</u>**	**<u>− 2.16****</u>**	**<u>− 2.73****</u>**	0.64
ST297Q03JA	Student received Internet or computer tutoring with a programme or application [synchronous]	**<u>− 2.10****</u>**	**<u>− 2.68****</u>**	**<u>− 1.63****</u>**	**<u>− 2.29****</u>**
ST297Q06JA	Student received small group study or practice (2–7 students) [synchronous]	**<u>− 1.34****</u>**	**<u>− 1.91****</u>**	**<u>− 1.07****</u>**	− 0.76
Institutional related factors					
RELATST	Quality of student–teacher relationships (WLE) $_{gmc}$	**<u>5.28****</u>**	**<u>7.78****</u>**	**<u>3.67****</u>**	**<u>2.82****</u>**
TEACHSUP	Mathematics teacher support (WLE) $_{gmc}$	**<u>1.96****</u>**	**<u>3.69****</u>**	**<u>0.63***</u>**	**<u>3.99****</u>**
REPEAT	Grade repetition (repeated at least once = 1, else = 0)	**<u>− 30.35****</u>**	**<u>− 38.49****</u>**	**<u>− 25.84****</u>**	**<u>− 29.42****</u>**
MISSSC	Missing school for more than 3 months (Missed school for > 3 months at least once = 1, else = 0)	**<u>− 16.68****</u>**	**<u>− 16.88****</u>**	**<u>− 16.39****</u>**	**<u>− 17.00****</u>**
Controls					
ESCS	Index of economic, social and cultural status$_{gmc}$	**<u>16.42****</u>**	**<u>22.58****</u>**	**<u>12.68****</u>**	**<u>15.94****</u>**

Table 2 (continued)

Item code	Predictors (coding)	Global	Developed	Developing	East Asian
WORKPAY	Working for pay before or after school$_{gmc}$	**– 6.82******	**– 7.09******	**– 6.76******	**– 9.65******
EXERPRAC	Exercise or practice a sport before or after school$_{gmc}$	**– 4.79******	**– 5.82******	**– 4.16******	**– 9.65******
WORKHOME	Working in household/take care of family members before or after school$_{gmc}$	**– 4.10******	**– 6.18******	**– 2.81******	**– 8.48******

[a] Variable is continuous with the proportion of males between 0 and 1 specified for each school; substantive effects with $p < 0.01$ underlined and in bold; notable reverse effects for groups of countries underlined and in bold; gmc = variable is group-mean centered on each school; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$

math performance, i.e., "Student received video-recorded instruction by a person." However, three shadow education factors (Internet or computer tutoring, one-on-one tutoring, and large group study) were negatively associated with math performance, and one (small group study) was not related. At the between-school level, three institutional related factors appeared to have a positive relationship with math performance, namely, (1) quality of student–teacher relationships, (2) math extra-curricular activities at school, and (3) provision of peer-to-peer study help for students.

At the within-school level, the only shadow education related variable to have a positive global effect was "Student received video instruction by a person" ($b = 0.98$, $p < 0.0001$). Student engagement in one-on-one tutoring, Internet/computer tutoring, and small- and large-group study appeared to be significantly negatively associated with math performance. In addition, with reference to institutional factors, the quality of student–teacher relationships and mathematics-teacher support were also positively associated with student math performance within schools.

For RQ3, that is, the relationship between student engagement in shadow education and math performance for developed, developing, and East Asian countries, the pattern of effects was generally the same but with some exceptions (Table 2, "Developed", "Developing", "East Asian" columns). At the between-school level, globally, aggregated incidents of students receiving large group study or practice were negatively associated with math performance ($b = -5.21$, $p < 0.001$). However, this effect was reversed for the East Asian region ($b = 17.26$, $p < 0.0001$), with more incidents of large group study in schools being positively associated with school math performance.

In terms of personal study, the average amount of time students studied outside of school (STUDYHMW) and time spent specifically on math homework was positively related to the average math performance of schools at the between-school level. However, the roles of these two factors are reversed at the within-school level. Note that while they are globally negatively related to students' math performance, these within-school effects reverse for students in developing and East Asian regions (readers are advised to review additional analyses in Tables A2 and A3 providing details on the sensitivity analysis and demonstration of minimal levels of collinearity between the shadow-education related variables [within-schools], respectively).

RQ4 examines the moderating effect that socio-economic status has on the relationship between student engagement in shadow education and student math performance. Table 3

Table 3 Moderation effects of socio-economic status on relationship between engagement in shadow education and math performance (within-schools)

Item code	Independent variable	Global (b)	Developed (b)	Developing (b)	East Asian (b)
ST297Q07JA	Student received large group study/practice$_{gmc}$	**− 1.18******	**− 1.36*****	**− 1.08******	**− 1.38 ****
ST297Q05JA	Student received video-recorded instruction by a person$_{gmc}$	**− 1.15******	**− 1.51******	**− 0.99******	**− 1.84****
ST297Q06JA	Student received small group study/practice$_{gmc}$	**− 0.83******	− 0.56	**− 0.93******	− 0.77
ST297Q03JA	Student received Internet or computer tutoring with a programme or application$_{gmc}$	**− 0.73******	**− 0.96****	**− 0.67****	− 0.78
ST297Q01JA	Student received one-on-one tutoring with a person$_{gmc}$	**− 0.81******	− 0.77*	**− 0.80******	− 0.16

All effects are modelled at the within-school level; exceptional effects underlined and in bold; gmc = group mean centered; factors ordered by size of global coefficient for each predictor group; substantive effects with $p < 0.01$ underlined and in bold; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$

provides the results, with the overall trend generally suggestive of negative moderation effects.

Results suggest that, globally, SES has a negative moderating effect on the relationship between shadow education and math performance. However, these effects appear to be less consistent and pronounced in the developed and East Asian educational contexts.

Discussion

The results of this study reveal the multifaceted nature of shadow education and its relationship to student academic performance, considering various regional factors and socio-economic conditions. In this discussion, we synthesize the findings from our study and the insights gained from previous research to explore the implications of shadow education for educational practice and policy. We delve into the nuanced dynamics of shadow education, considering its differential effects across socioeconomic strata, school environments, and supplementary education programs. By examining the interplay between shadow education and student outcomes, we aim to provide a comprehensive understanding of how educational stakeholders can leverage supplementary education to promote educational equity and improve student achievement. This discussion seeks to inform future research directions and practical strategies for addressing educational inequalities and fostering positive learning experiences for all students.

Shadow education and mathematics achievement

The findings of our study present a complex picture of the effects of shadow education on students' mathematics achievement, revealing both positive and negative associations depending on the type of student engagement in shadow education and the context in which it is used. Our results highlight the importance of understanding the nature and implementation of shadow education to discern its association with academic performance.

At the global level, the only shadow education variable positively associated with student mathematics performance was receiving video-recorded instruction by a person. This finding suggests that asynchronous learning, where students have the flexibility to engage with instructional material at their own pace and time, can be particularly beneficial. The flexibility of video-recorded instruction may better accommodate individual learning needs than synchronous formats. Some empirical studies have highlighted the effectiveness of asynchronous learning in providing personalized educational experiences and accommodating diverse learning paces (Yung, 2020; Yung & Wong, 2024; Zheng et al., 2022). This finding is especially relevant in the context of the COVID-19 pandemic, during which students faced significant disruptions to traditional schooling and may have benefited from the adaptable nature of asynchronous learning (Hajar & Karakus, 2023a; Yung & Wong, 2024; Zuo et al., 2021).

Conversely, other forms of shadow education, such as one-on-one tutoring, Internet/computer tutoring, and large-group study, were negatively associated with mathematics performance at the between-school level. The negative associations for these forms of shadow education, which were also generally observed at the within-school level, could be attributed to their predominant use for remedial purposes. Students struggling with mathematics may seek (or more likely be "assigned by parents/guardians to receive") additional help through these methods, and their lower performance could reflect their initial difficulties rather than the ineffectiveness of shadow education itself. Baker et al. (2001) and Wiseman (2021) also identified negative associations between shadow education and student mathematics achievement globally. They suggested that private tutoring is often utilized for remedial purposes, with lower-performing students opting for supplementary tutoring to receive individualized support. These findings align with previous research indicating that in contexts where shadow education is primarily used for remedial purposes, its positive effects on academic achievement are rarely observed and often questioned (Buchmann et al., 2010; Entrich, 2021; Guill et al., 2019). Conversely, in contexts where a significant number of high-performing students engage in shadow education for enrichment, more positive outcomes are evident (Baker et al., 2001; Byun, 2014). This suggests that the impact of shadow education varies greatly depending on its purpose and the profile of the students who participate in it. However, more research may be necessary to provide a more definitive understanding of this proposed "remedial reverse effect" in post-COVID educational contexts.

Moreover, as the current findings reveal, the negative association between large-group study sessions and academic performance in developed Western countries, compared to its positive effects in East Asian countries, highlights important cultural and systemic differences in how such educational interventions are received and utilized. In East Asian contexts, large-group study sessions tend to be more structured, with a stronger alignment between shadow education content and school curricula. This structured approach, coupled with culturally ingrained discipline and the high value placed on academic achievement, may explain the positive outcomes observed in these countries (Hau & Ho, 2010; Su & Lee, 2023; Zayet et al., 2023). In contrast, in Western contexts, large-group tutoring sessions may lack the necessary focus and personalization, leading to less effective outcomes. The absence of such structure, combined with differences in pedagogical approaches, may explain the relatively negative association observed in these settings.

The differences between these regions may also reflect varying motivations for participating in shadow education, which could introduce selection bias. In East Asian countries, shadow education is often viewed as a necessary supplement to formal schooling, particularly in preparing for high-stakes exams and future career prospects. As Byun et al. (2018) point out, the negative correlation between academic performance and engagement in shadow education is less pronounced in Southeast and East Asian countries than in other regions. In some cases, such as South Korea and Taipei, a positive correlation has also been observed between participation in shadow education and academic achievement. This could be attributed to the widespread prevalence of shadow education in these countries, where educational credentials are strongly tied to social mobility and economic opportunities, and Confucian values place a high emphasis on education.

However, it is crucial to distinguish between the observed associations and causal impacts. The positive outcomes in East Asian contexts may reflect not just the efficacy of shadow education itself, but also a selection effect where students who are already more motivated or academically inclined are more likely to engage in these supplementary sessions. Therefore, while shadow education may have a role in improving academic performance, especially in specific cultural contexts, the findings should be interpreted with caution, considering the complex interplay of motivation, cultural expectations, and structural differences in educational systems. At the within-school level, the pattern of results mirrored the global findings, with video-recorded instruction again being the sole positive influence on math performance. The consistency of this finding underscores the potential of asynchronous learning tools to enhance student outcomes, possibly by allowing students to revisit challenging content and learn at their own pace (Zheng et al., 2022).

The negative associations with other shadow education variables at the within-school level also suggest that the effectiveness of shadow education modes can vary significantly based on how they are integrated into students' overall learning experience. For instance, Internet/computer tutoring with a programme or application might not provide the same level of engagement or tailored instruction necessary for substantial improvements in math performance. Despite their potential benefits in collaborative learning environments, small-group and large-group types of tutoring might also lack the individual attention required to address specific learning gaps in math.

It is also crucial to consider the timing of the data collection of PISA 2022, which was conducted shortly after the COVID-19 pandemic. The pandemic profoundly affected educational systems worldwide, forcing a rapid shift to online and remote learning. This abrupt transition may have influenced the role and effectiveness of various shadow education forms during this period (Bray & Hajar, 2023; Hajar & Karakus, 2023a; Lee et al., 2023). The United Arab Emirates, for instance, warned that anyone providing face-to-face tutoring in public or private places would be fined 30,000 Emirati Dirham (approx. 8200 USD); host venues also fined 20,000 Emirati Dirham (approx. 5400 USD) (Bray & Hajar, 2023). The heightened reliance on asynchronous learning tools like video-recorded instruction could be partly due to their compatibility with remote learning environments, where flexibility and accessibility are paramount (Hajar & Karakus, 2023a; Zayet et al., 2023).

While "remedial reverse effects" might offer some explanation for the negative relationship between synchronous supplementary tutoring and math achievement in the current study, global conditions should also be considered. It is also conceivable that synchronous supplementary online tutoring during the COVID-19 restrictive period may have been potentially psychologically burdensome to many students. Certainly, during this period, an unprecedented level of screentime was *imposed* upon students. Therefore, the additional imposition of asynchronous online supplementary tutoring may have been counter-productive in the given context.

In summary, our findings indicate that not all forms of shadow education are equally effective in improving math achievement. Asynchronous video-recorded instruction stands out as a beneficial method, likely due to its flexibility and alignment with students' learning needs during the pandemic. In contrast, other forms of shadow education, particularly those used for remedial purposes, may reflect students' struggles rather than the interventions' inefficacy. Future research should continue to explore these dynamics, considering the evolving educational landscape and students' varying needs across different cultural and disciplinary contexts. Related to this, the duration of receiving private tutoring (PT) can be associated with the academic achievement of students, and this factor warrants further research. That is, some researchers (e.g., Bray & Hajar, 2023; Hajar & Karakus, 2023b) found that the demand for shadow education often increases as student exams become more imminent; however, it is important to note that shadow education might not be quite effective for those who only receive it for a short period before sitting high-stakes exams. Therefore, future research might also explore temporal aspects of shadow education.

The moderating role of SES

Our investigation into the moderating effect of SES on the relationship between student engagement in shadow education and mathematics performance reveals intriguing and complex dynamics. With reference to the main models (Table 2), the significant positive impact of SES at both the between- and within-school levels suggests that students from higher SES backgrounds achieve better mathematics performance. However, further analysis suggested that SES also had a moderating effect on the relationship between student engagement in shadow education and math performance.

Findings suggested that SES functioned as a *negative moderator* between incidents of receiving "asynchronous video-recorded instruction from a person" and math performance. This meant that students with a lower SES status tended to benefit *more* from receiving such asynchronous video-recorded instructions. This finding may be informative to the early perspective that shadow education can act as a vital tool to bridge educational gaps and enhance academic outcomes for underprivileged students. Kim's (2016) study in South Korea indicates that shadow education can help minimize educational inequality by providing personalized assistance to students who are lagging behind, often at more affordable prices. However, it may be that asynchronous forms of shadow education may be a more optimal way to bridge educational gaps.

The moderating effect of SES on the relationship between student engagement in *synchronous* shadow education can be interpreted differently. For example, at the within-school level, the direct effect of student engagement in synchronous shadow education had a consistent negative effect on student math performance. However, moderation analysis suggests that the intensity of this negative effect is exacerbated for students with a higher SES status. Therefore, in the current study, students with lower SES backgrounds appear to be *less* negatively affected by engagement in synchronous shadow education. These findings herein appear counter-intuitive.

The literature presents mixed views on this general issue. Several studies from high-income countries such as the United States (Byun & Park, 2012), England (Holloway & Kirby, 2019), and South Korea (Ku et al., 2022), as well as middle- and low-income countries including Myanmar (Bray et al., 2020), Kazakhstan (Hajar & Karakus, 2023b), India (Gupta, 2021), and Cambodia (Marshall & Fukao, 2019) have demonstrated that shadow education tends to exacerbate educational inequalities. These studies argue that students from higher-income families can afford better and more extensive tutoring than those from lower-income families and low-achieving students, thus perpetuating the disparity in educational outcomes. For instance, Liao and Huang (2018) revealed that economically advantaged students benefited more from participation in shadow education, as evidenced by the PISA 2015 results from mainland China.

Nevertheless, based on PISA 2012 results globally, Wiseman (2021) discovered that private tutoring does not intensify educational inequality, as students from higher-income families do not necessarily receive more educational resources within schools, nor do they supplement their education outside of school more frequently than students from lower-income families. Similarly, Entrich (2021) showed that access to shadow education is more significantly influenced by cultural and educational resources than by a family's economic status.

Furthermore, institutional, economic, and cultural differences can play a crucial role in shaping the impact of shadow education.

Our results indicate that the negative moderating effects of SES on the relationship between synchronous engagement in shadow education and math performance are less consistent and pronounced in developed and East Asian educational contexts. Therefore, the mechanism by which engagement in shadow education is associated with lower math performance appears to be less pronounced in developed and East Asian regions.

Entrich (2021) notes that developed nations with high GDP per capita and less stratified, highly inclusive comprehensive schooling systems often provide fewer incentives for high performers to participate in shadow education. In such education systems, the drive to use supplementary education is reduced because the schooling system itself is designed to offer more equitable learning opportunities, leading to less pressure for students to seek external academic support. However, this does not necessarily mean that the impact of shadow education decreases in these systems. Instead, the lower participation rates among high-achieving students reflect the reduced need for additional academic support in a more equal system. The effectiveness of shadow education remains context-dependent, and the selection bias related to who chooses to participate in shadow education plays a crucial role. In some cases, even a reversed SES gap in access to shadow education is observed, where students from lower socioeconomic backgrounds may utilize shadow education to overcome educational challenges in systems with less selective schooling models.

In these contexts, shadow education can serve as a remedial tool that helps low-performing students bridge learning gaps and counteract their socioeconomic disadvantages, allowing them to achieve higher academic success (Entrich, 2021). This provides a possible explanation for the generally strong 'remedial reverse effects' demonstrated for student engagement in synchronous shadow education in developed countries observed in this study. However, these findings should not be interpreted as a direct causal impact of shadow education but rather as an association shaped by the differing motivations and needs of students in these education systems.

Implications and conclusion

This study provides several key implications for educators, policymakers, and researchers examining the role of shadow education in student academic achievement, particularly in mathematics. The positive influence of video-recorded instruction indicates that asynchronous learning tools can be more widely incorporated into supplementary education strategies. For practitioners, this entails developing high-quality video content aligned with curriculum standards to address common challenges in mathematics. A hybrid approach combining synchronous and asynchronous learning could better cater to diverse student needs, offering flexibility in post-pandemic educational environments (Yung & Wong, 2024; Zheng et al., 2022).

The mixed or negative associations found with other forms of shadow education, such as one-on-one tutoring, Internet/computer tutoring, and large- or small-group study, highlight the need for a more nuanced understanding of these interventions. It is important to clarify that these associations do not necessarily imply a causal impact of shadow education on student outcomes but could be influenced by selection biases. For example, students who are already struggling or have specific needs may be more likely to participate in certain forms of shadow education, which can affect the observed outcomes. Therefore, it is critical for practitioners to consider students' specific contexts, including their pre-existing academic levels and needs, when recommending these forms of shadow education, particularly for remedial purposes. Targeted, individualized interventions that address specific learning gaps may be more effective. Additionally, large-group tutoring sessions can be carefully structured to provide value, particularly in Western contexts where they may otherwise be less effective (Baker et al., 2001; Byun, 2014).

For researchers, this study underscores the importance of exploring the contextual factors that influence the effectiveness of different forms of shadow education. Future research should examine how cultural, socioeconomic, and educational system differences impact the success of shadow education interventions. Longitudinal studies are particularly crucial for tracking the long-term effects of shadow education and understanding how its impact evolves across various student populations. The effects of the COVID-19 pandemic on supplementary education practices and outcomes also warrant further investigation, given the potential long-term shifts in education toward remote and hybrid learning models (Lee et al., 2023).

Policymakers can use these findings to inform the design and regulation of shadow education markets, ensuring that high-quality, affordable, and accessible resources are available to all students, particularly those from disadvantaged backgrounds (Kim, 2016; Zhang, 2023). Bridging the gap between shadow education and formal schooling by integrating asynchronous learning tools into mainstream education could offer more equitable opportunities for students (Bray & Hajar, 2023; Lee et al., 2023). Subsidies or financial assistance programs may also be necessary to ensure that disadvantaged students can access these resources (Kim, 2016).

This study also highlights the importance of addressing socioeconomic disparities in academic achievement. Policymakers can consider expanding partnerships between schools and shadow education providers to offer services at reduced or no cost to students from lower SES backgrounds. For example, Zhang and Bray (2020) describe a US-based online tutoring company funded by the Department of Defence that provides free services to military families. In the UK, Kirby (2016) suggests expanding non-profit tuition programs like the Tutor Trust to benefit disadvantaged students.

Practitioners, particularly educators and tutors, can tailor shadow education programs to meet the specific needs of disadvantaged students. This could include developing targeted instructional strategies to close learning gaps and provide individualized support. Cultural and contextual differences can also be taken into account, with teaching methods adapted to maximize effectiveness across diverse settings (Lee & Shouse, 2011; Park, 2013).

In summary, our findings highlight the potential of asynchronous shadow education to promote educational equity, particularly for socioeconomically disadvantaged students. By providing targeted resources and tailored interventions, educators and policymakers can help ensure that all students have the opportunity to achieve their full academic potential. Further research is needed to refine these strategies and better understand the broader implications of shadow education in different educational contexts.

This study emphasizes the varied effects of shadow education on mathematics achievement, with asynchronous learning tools such as video-recorded instruction standing out as particularly effective. These findings point to the need for educators, policymakers, and researchers to consider the contextual and individual factors that shape the effectiveness of shadow education interventions. Integrating high-quality asynchronous tools, offering hybrid learning models, and implementing targeted interventions can help better support diverse student needs and improve overall educational outcomes.

Moving forward, research should continue exploring shadow education's long-term effects, particularly through longitudinal studies. Understanding the causal impact of shadow education is essential to developing evidence-based strategies that can maximize its potential to improve learning outcomes for all students, ensuring equitable and adequate educational opportunities on a global scale.

Limitations

Despite this study's comprehensive design, several limitations must be acknowledged. First, using the PISA 2022 secondary dataset restricts the analysis of available variables and does not capture all possible factors influencing the relationship between shadow education and math performance. For instance, qualitative aspects of shadow education, such as the quality of instruction or student motivation, were not assessed, which could provide deeper insights into the effectiveness of different forms of shadow education.

Second, the cross-sectional design of the PISA dataset limits the ability to identify any causal relationships. While the study identifies associations between shadow education and math performance, it cannot definitively determine whether it directly causes changes in math achievement. Additionally, a significant limitation is the absence of data on students' prior academic achievement or cognitive ability, which are important factors that could influence both participation in shadow education and subsequent academic performance. Without controlling for these variables, the observed effects may be partially driven by pre-existing differences in students' cognitive abilities or prior achievement levels. As a result, the findings should be interpreted with caution. Longitudinal studies that track students' performance over time and incorporate measures of prior achievement are needed to better understand the causal dynamics and more accurately isolate the true impact of shadow education.

Third, the study's findings may be influenced by cultural and contextual differences that are not fully accounted for in the analysis. The effectiveness of shadow education varies significantly across regions, particularly between developed, developing, and East Asian countries, due to differing educational systems, cultural attitudes towards education, and economic factors. Future research should consider more localized studies to understand these regional differences better.

Fourth, the study relies on self-reported data for variables such as student engagement in shadow education and socioeconomic status, which may be subject to bias or inaccuracies. Students and parents may overestimate or underestimate their engagement in shadow education activities, and socioeconomic status may not be uniformly reported across different countries.

Finally, while significant, SES's moderating effect was not uniformly consistent across all contexts. This suggests that other unmeasured variables may interact with SES to influence the relationship between shadow education and math performance. Further research is needed to identify these variables and understand their interactions.

In summary, while this study provides valuable insights into the impact of shadow education on math performance, the limitations highlight the need for more detailed, causal, and context-specific research. Addressing these limitations in future studies will help to develop more effective and equitable educational policies and practices. Moving forward, further research is needed to explore the long-term impacts of synchronous and asynchronous use of shadow education (and use of AI-bots; see Batyrbayeva, 2024) and develop evidence-based strategies to maximize its potential to enhance student learning outcomes.

Appendix A

See Tables 4, 5, 6

Table 4 Classification of countries into regional groups

Developed (Western)	S_1–S_2	N_1–N_2	Developing economies	S_1–S_2	N_1–N_2	East Asian	S_1–S_2	N_1–N_2
AUS: Australia	743–500	13437–8917	ALB: Albania	274–108	6129–2416	HKG: Hong Kong (China)	163–100	5907–3532
BEL: Belgium	285–61	8286–1511	ARE: United Arab Emirates	840–547	24600–15308	KOR: South Korea	186–177	6454–6096
BGR: Bulgaria	202–133	6107–3519	ARG: Argentina	457–319	12111–7000	SGP: Singapore	164–160	6606–6327
CHE: Switzerland	260–142	6829–3249	BRA: Brazil	598–289	10798–5105	TAP: Chinese Taipei	182–173	5857–5625
CZE: Czech Republic	430–281	8460–6383	CHL: Chile	230–171	6488–4491	THA: Thailand	279–219	8495–7717
DNK: Denmark	347–177	6200–2916	COL: Colombia	262–204	7804–5876	VNM: Viet Nam	178–148	6068–5479
EST: Estonia	196–154	6392–5559	DOM: Dominican Republic	253–97	6868–2198	–	–	–
FRA: France	282–144	6770–3478	GEO: Georgia	267–120	6583–2887	–	–	–
GBR: United Kingdom	451–281	12972–6527	GTM: Guatemala	290–70	5190–1498	–	–	–
GRC: Greece	230–169	6403–4808	HKG: Hong Kong (China)	163–100	5907–3532	–	–	–
HRV: Croatia	180–157	6135–4920	JOR: Jordan	260–222	7799–5393	–	–	–
HUN: Hungary	262–147	6198–4421	KAZ: Kazakhstan	571–480	19769–18041	–	–	–
ISL: Iceland	134–55	3360–1719	KOR: South Korea	186–177	6454–6096	–	–	–
LTU: Lithuania	292–199	7257–5831	KSV: Kosovo	229–77	6027–2532	–	–	–
MLT: Malta	46–36	3127–2307	MDA: Republic of Moldova	265–191	6235–5018	–	–	–
NLD: Netherlands	154–98	5046–2865	MEX: Mexico	280–167	6288–4649	–	–	–
NZL: New Zealand	169–99	4682–2529	MKD: North Macedonia	111–68	6610–3384	–	–	–
POL: Poland	240–197	6011–5010	MNE: Montenegro	63–45	5793–4552	–	–	–
PRT: Portugal	224–203	6793–5704	PAN: Panama	215–38	4544–754	–	–	–
ROU: Romania	262–155	7364–5808	PER: Peru	336–207	6968–4547	–	–	–
SVK: Slovak Republic	288–159	5824–4029	PRY: Paraguay	281–76	5084–1671	–	–	–
SVN: Slovenia	345–208	6721–5174	PSE: Palestinian Authority	273–232	7905–5894	–	–	–
SWE: Sweden	262–171	6072–3724	QAT: Qatar	229–194	7676–5306	–	–	–
USA: United States	154–121	4552–3264	QAZ: Baku (Azerbaijan)	199–65	7720–1498	–	–	–
–	–	–	QUR: Ukrainian regions[1]	164–99	3876–2468	–	–	–
–	–	–	SAU: Saudi Arabia	193–148	6928–4687	–	–	–

Table 4 (continued)

Developed (Western)	S_1–S_2	N_1–N_2	Developing economies	S_1–S_2	N_1–N_2	East Asian	S_1–S_2	N_1–N_2
–	–	–	SGP: Singapore	164–160	6606–6327	–	–	–
–	–	–	SLV: El Salvador	290–108	6705–2688	–	–	–
–	–	–	SRB: Serbia	183–164	6413–5276	–	–	–
–	–	–	TAP: Chinese Taipei	182–173	5857–5625	–	–	–
–	–	–	THA: Thailand	279–219	8495–7717	–	–	–
–	–	–	TUR: Türkiye	196–179	7250–6529	–	–	–
–	–	–	URY: Uruguay	222–178	6618–4274	–	–	–
–	–	–	VNM: Viet Nam	178–148	6068–5479	–	–	–

[1] 18 of 27 regions; S_1 = total number of schools sampled in original PISA dataset; S_2 = total number of schools included in the current study; N_1 = total number of students sampled in original PISA dataset; N_2 = total number of students included in the current study

Table 5 Multilevel models for the role of shadow education on student math ability (with demographic control variables only)

Item code	Predictors (coding)	Global	Developed	Developing	East Asian
–	Intercept	445.32****	443.94****	440.25****	506.14****
Between-School Effects					
Shadow Education ["Additional math instruction"]					
ST297Q05JA	Student received video-recorded instruction from a person [asynchronous]	**6.81*** **	5.7*	**8.06*** **	13.87
ST297Q03JA	Student received Internet or computer tutoring with a programme/application [synchronous]	**− 14.02**** **	− 5.36*	**− 16.48**** **	**^− 27.39*** **
ST297Q01JA	Student received one-on-one tutoring with a person [synchronous]	**− 14.03**** **	**− 16.85**** **	**− 18.8**** **	**− 48.2**** **
ST297Q07JA	Student received large group study or practice (8 or more students) [synchronous]	**− 15.01**** **	**− 22.91**** **	**^− 10.77**** **	**31.07**** **
ST297Q06JA	Student received small group study or practice (2–7 students) [synchronous]	− 2.16	− 2.87	− 0.05	6.84
Covariates					
ESCS	Index of economic, social and cultural status	**69.96**** **	**91.91**** **	**59.29**** **	**90.69**** **
ST004D01T	Student gender (male = 1, female = 0)[a]	2.93	5.91*	2.8	0.77
SCHLTYPE	School type (private = 0, public = 1)	**4.89**** **	3.92*	1.58	**14.39**** **
Within-School Effects					
Shadow Education ["Additional math instruction"]					
ST297Q05JA	Student received video-recorded instruction from a person [asynchronous] $_{gmc}$	**0.82**** **	− 0.39	**1.52**** **	0.49
ST297Q03JA	Student received Internet or computer tutoring with a programme/application [synchronous] $_{gmc}$	**− 2.66**** **	**− 3.85**** **	**− 1.94**** **	**− 2.8**** **
ST297Q01JA	Student received one-on-one tutoring with a person [synchronous] $_{gmc}$	**− 10.37**** **	**− 12.16**** **	**− 9.29**** **	**− 11.43**** **
ST297Q07JA	Student received large group study or practice (8 or more students) [synchronous] $_{gmc}$	**− 3.35**** **	**− 3.21**** **	**− 3.68**** **	0.32
ST297Q06JA	Student received small group study or practice (2–7 students) [synchronous] $_{gmc}$	**− 1.86**** **	**− 2.67**** **	**− 1.48**** **	^− 0.88*
Covariate					
ESCS	Index of economic, social and cultural status$_{gmc}$	**16.29**** **	**22.60**** **	**12.6**** **	**16.55**** **

[a] Variable Student gender (male = 1, female = 0) is continuous with the proportion of males between 0 and 1 specified for each school; substantive effects with $p < 0.01$ underlined and in bold; notable reverse effects for groups of countries underlined and in bold; gmc = variable is group-mean centered on each school; ^coefficient became non-statistically significant in the main model that included all control variables, Table 2; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$

Table 6 Tests for collinearity between shadow education variables

Item code	Variable	ST297Q01JA	ST297Q03JA	ST297Q05JA	ST297Q06JA	ST297Q07JA
Within-School Effects						
Shadow Education ["Additional math instruction"]						
ST297Q01JA	Student received one-on-one tutoring with a person [synchronous]	1.00	–	–	–	–
ST297Q03JA	Student received Internet or computer tutoring with a programme/application [synchronous]	0.29****	1.00	–	–	–
ST297Q05JA	Student received video-recorded instruction from a person [asynchronous]	0.28****	0.45****	1.00	–	–
ST297Q06JA	Student received small group study or practice (2 to 7 students) [synchronous]	0.27****	0.30****	0.29****	1.00	–
ST297Q07JA	Student received large group study or practice (8 or more students) [synchronous]	0.25****	0.29****	0.30****	0.36****	1.00
Between-School Effects						
Shadow Education ["Additional math instruction"]						
ST297Q01JA	Student received one-on-one tutoring with a person [synchronous] $_{gmc}$	1.00	–	–	–	–
ST297Q03JA	Student received Internet or computer tutoring with a programme/application [synchronous] $_{gmc}$	0.55****	1.00	–	–	–
ST297Q05JA	Student received video-recorded instruction from a person [asynchronous] $_{gmc}$	0.54****	0.77****	1.00	–	–
ST297Q06JA	Student received small group study or practice (2 to 7 students) [synchronous] $_{gmc}$	0.52****	0.44****	0.45****	1.00	–

Table 6 (continued)

Item code	Variable	ST297Q01JA	ST297Q03JA	ST297Q05JA	ST297Q06JA	ST297Q07JA
ST297Q07JA	Student received large group study or practice (8 or more students) [synchronous] gmc	0.45****	0.52****	0.50****	0.61****	1.00

All values represent correlation coefficients, *r*; gmc = variable is group-mean centered on each school; substantive effects with $p < .01$ underlined and in bold; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$

Acknowledgements
Not applicable.

Author contributions
MK and AH conceptualized this paper. MC and AT carried out all statistical analyses. MK, AH, MC, and AT contributed to the literature review and discussion sections. All authors read and approved the final manuscript.

Funding
No funding was used in support of this paper.

Availability of data and materials
The datasets analysed during the current study are available from the corresponding author on reasonable request (also available online).

Declarations

Ethics approval and consent to participate
All data was taken from the PISA 2022 cycle. Therefore, it was assumed that ethical approval was gained by the OECD for all participants.

Consent for publication
All contributing authors give consent for this paper to be published pending acceptance.

Competing interests
The authors declare that they have no competing interests.

Received: 15 July 2024 Accepted: 21 October 2024
Published online: 04 December 2024

References

Baker, D. P., Akiba, M., LeTendre, G. K., & Wiseman, A. W. (2001). Worldwide shadow education: outside-school learning, institutional quality of schooling, and cross-national mathematics achievement. *Educational Evaluation and Policy Analysis, 23*(1), 1–17. https://doi.org/10.3102/01623737023001001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.1863/jss.v067.i01

Batyrbayeva, A. (2024). The impact of AI-bot usage on students' academic achievements and motivation in Kazakhstan. Graduate School of Education. [Master's thesis, Nazarbayev University]. Nazarbayev University Repository. https://nur.nu.edu.kz/handle/123456789/7803

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C. D., Clyde, M., Cook, T. D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., & Johnson, V. E. (2017). Redefine statistical significance. *Nature Human Behaviour, 2*(1), 6–10. https://doi.org/10.1038/s41562-017-0189-z

Berberoğlu, G., & Tansel, A. (2014). Does private tutoring increase students' academic performance? Evidence from Turkey. *International Review of Education, 60*(5), 683–701. https://doi.org/10.1007/s11159-014-9436-y

Bray, T. M., Kobakhidze, M. N., & Kwo, O. W. Y. (2020). *Shadow education in Myanmar: Private supplementary tutoring and its policy implications*. UNESCO & Comparative Education Research Centre (CERC), the University of Hong Kong.

Bray, M. (1999). *The shadow education system: private tutoring and its implications for planners. Fundamentals of educational planning No. 61*. UNESCO International Institute for Educational Planning (IIEP).

Bray, M. (2011). *The challenge of shadow education: private tutoring and its implications for policymakers in the European Union*. European Commission.

Bray, M. (2014). The impact of shadow education on student academic achievement: why the research is inconclusive and what can be done about it. *Asia Pacific Education Review, 15*, 381–389. https://doi.org/10.1007/s12564-014-9326-9

Bray, M. (2021). *Shadow education in Africa: private supplementary tutoring and its policy implications*. The University of Hong Kong, UNESCO.

Bray, M. (2023). Understanding private supplementary tutoring: metaphors, diversities and agendas for shadow education research. *Journal for the Study of Education and Development, 46*(4), 728–773. https://doi.org/10.1080/02103702.2023.2194792

Bray, M. (2024). Still in need of confronting: shadow education and its implications in the sustainable development goals. *International Journal of Educational Development, 104*, 102967. https://doi.org/10.1016/j.ijedudev.2023.102967

Bray, M., & Hajar, A. (2023). *Shadow education in the Middle East: private supplementary tutoring and its policy implications*. Routledge.

Broer, M., Bai, Y., & Fonseca, F. (2019). A review of the literature on socioeconomic status and educational achievement. In M. Broer, Y. Bai, & F. Fonseca (Eds.), *Socioeconomic inequality and educational outcomes: evidence from twenty years of TIMSS* (pp. 7–17). Springer.

Buchmann, C., Condron, D. J., & Roscigno, V. J. (2010). Shadow education, American style: test preparation, the SAT and college enrollment. *Social Forces, 89*(2), 435–461. https://doi.org/10.1353/sof.2010.0105

Byun, S. Y. (2014). Shadow education and academic success in South Korea. In H. Park & K.-K. Kim (Eds.), *Korean education in changing economic and demographic contexts* (pp. 39–58). Springer.

Byun, S. Y., Chung, H. J., & Baker, D. P. (2018). Global patterns of the use of shadow education: student, family, and national influences. In H. Park & G. Kao (Eds.), *Research in the sociology of education* (pp. 71–105). Emerald Publishing Limited.

Byun, S. Y., & Park, H. (2012). The academic success of East Asian American youth: the role of shadow education. *Sociology of Education, 85*(1), 40–60.

Cheo, R., & Quah, E. (2005). Mothers, maids and tutors: an empirical evaluation of their effect on children's academic grades in Singapore. *Education Economics, 13*(3), 269–285. https://doi.org/10.1080/09645290500073746

Choi, Á., Calero, J., & Escardíbul, J. O. (2012). Private tutoring and academic achievement in Korea: An approach through PISA-2006. *Kedi Journal of Educational Policy*, *9*(2).

Courtney, M. G. R., & Chang, K. C. (2018). Dealing with non-normality: an introduction and step-by-step guide using R. *Teaching Statistics, 40*(2), 51–59. https://doi.org/10.1111/test.12154

Courtney, M., Torrano, D. H., Karakus, M., & Singh, N. (2023). Measuring student well-being in adolescence: proposal of a five-factor integrative model based on PISA 2018 survey data. *Large-Scale Assessments in Education*. https://doi.org/10.1186/s40536-023-00170-y

Domingue, B., & Briggs, D. C. (2009). Using linear regression and propensity score matching to estimate the effect of coaching on the SAT. *Multiple Linear Regression Viewpoints, 35*(1), 12–29.

Entrich, S. R. (2021). Worldwide shadow education and social inequality: explaining differences in the socioeconomic gap in access to shadow education across 63 societies. *International Journal of Comparative Sociology, 61*(6), 441–475. https://doi.org/10.1177/0020715220987861

Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel modeling using R* (2nd ed.). Chapman and Hall/CRC.

Gao, X., Luo, J., Chen, H., Zhen, Y., Zhang, J., & Fu, X. (2023). Alleviating educational inequality in math with the aid of online shadow education–the impact of equal access and equal quality mechanisms. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-12214-5

Guill, K., Lüdtke, O., & Köller, O. (2019). Assessing the instructional quality of private tutoring and its effects on student outcomes: analyses from the German National Educational Panel Study. *British Journal of Educational Psychology, 90*(2), 282–300. https://doi.org/10.1111/bjep.12281

Gupta, A. (2021). Social legitimacy of private tutoring: an investigation of institutional and affective educational practices in India. *Discourse, 43*(4), 571–584. https://doi.org/10.1080/01596306.2020.1868978

Hajar, A., & Karakus, M. (2022). A bibliometric mapping of shadow education research: achievements, limitations, and the future. *Asia Pacific Education Review, 23*(2), 341–359. https://doi.org/10.1007/s12564-022-09759-4

Hajar, A., & Karakus, M. (2023). Throwing light on fee-charging tutoring during the global pandemic in Kazakhstan: implications for the future of higher education. *Asia Pacific Education Review*. https://doi.org/10.1007/s12564-023-09831-7

Hajar, A., & Karakus, M. (2023b). Private supplementary tutoring and educational inequality in secondary education in Kazakhstan. *Central Asian Survey, 42*(3), 617–636. https://doi.org/10.1080/02634937.2023.2234420

Hajar, A., & Karakus, M. (2024). An inevitable phenomenon: a bibliometric analysis of shadow education research (1982–2022). In A. Wiseman (Ed.), *Annual review of comparative and international education 2022* (pp. 193–213). Emerald Groups Publishing Limited.

Hajar, A., & Tabaeva, A. (2024). Being participatory: employing geographic lenses to understand young people's experiences of private supplementary tutoring in Uzbekistan. *Compare: A Journal of Comparative and International Education*. https://doi.org/10.1080/03057925.2024.2333520

Han, S., & Suh, H. (2020). The effects of shadow education on high school students' creative thinking and academic achievement in mathematics: the case of the Republic of Korea. *Educational Studies, 49*(2), 314–333. https://doi.org/10.1080/03055698.2020.1850427

Harwell, M., Maeda, Y., Bishop, K., & Xie, A. (2016). The surprisingly modest relationship between SES and educational achievement. *Journal of Experimental Education, 85*(2), 197–214. https://doi.org/10.1080/00220973.2015.1123668

Hau, K. T., & Ho, I. T. (2010). Chinese students' motivation and achievement. In M. H. Bond (Ed.), *The oxford handbook of chinese psychology* (pp. 187–204). Oxford University Press.

Holloway, S. L., & Kirby, P. (2019). Neoliberalising education: new geographies of private tuition, class privilege, and minority ethnic advancement. *Antipode, 52*(1), 164–184. https://doi.org/10.1111/anti.12587

Hox, J., & McNeish, D. (2020). Small samples in multilevel modeling. In R. V. D. Schoot & M. Miočevic (Eds.), *Small sample size solutions: a guide for applied researchers and practitioners* (pp. 215–225). Routledge.

Huang, F. L. (2022). *Practical multilevel modeling using R*. Sage Publications.

Huang, F. L. (2024). Using plausible values when fitting multilevel models with large-scale assessment data using R. *Large-scale Assessments in Education*. https://doi.org/10.1186/s40536-024-00192-0

Jansen, D., Elffers, L., & Jak, S. (2021). A cross-national exploration of shadow education use by high and low SES families. *International Studies in Sociology of Education, 32*(3), 653–674. https://doi.org/10.1080/09620214.2021.1880332
Jerrim, J., López-Agudo, L. A., Gutiérrez, Ó. D. M., & Shure, N. (2017). What happens when econometrics and psychometrics collide? An example using the PISA data. *Economics of Education Review, 61*, 51–58. https://doi.org/10.1016/j.econedurev.2017.09.007
Junaidi, S. (1977). Empirical study of private tutoring in Saudi Arabia', summarised in Saudi Arabia, ministry of education (1978): introducing some educational documents. *Journal of Educational Documentation, 15*, 86–104.
Kim, S., & Lee, J. H. (2010). Private tutoring and demand for education in South Korea. *Economic Development and Cultural Change, 58*(2), 259–296. https://doi.org/10.1086/648186
Kim, Y. C. (2016). *Shadow education and the curriculum and culture of schooling in South Korea*. Palgrave Macmillan. https://doi.org/10.1057/978-1-137-51324-3
Kim, Y. C., & Jung, J. H. (2022). *Theorizing shadow education and academic success in East Asia: understanding the meaning, value, and use of shadow education by East Asian students*. Routledge.
Kirby, P. (2016). Shadow schooling: Private tuition and social mobility in the UK. The Sutton Trust. http://www.suttontrust.com/wp-content/uploads/2016/09/Shadow-Schoolingformattedreport_FINAL.pdf.
Kobakhidze, N., & Suter, L. (2020). The global diversity of shadow education. *European Journal of Education, 55*(3), 316–321. https://doi.org/10.1111/ejed.12411
Ku, I., Lee, H., & Kim, J. E. (2022). Does shadow education contribute to inequality? *Asia Pacific Journal of Education*. https://doi.org/10.1080/02188791.2022.2114424
Kuan, P.-Y. (2011). Effects of cram schooling on mathematics performance: evidence from junior high students in Taiwan. *Comparative Education Review, 55*(3), 342–368. https://doi.org/10.1086/659142
Kuwait, Ministry of Education (1962). *Private Tutoring. Circular No.2/100/14740* (1 December). Kuwait City: Ministry of Education [in Arabic].
Lee, J. C., Quadlin, N., & Ambriz, D. (2023). Shadow education, pandemic style: social class, race, and supplemental education during COVID-19. *Research in Social Stratification and Mobility, 83*, 100755. https://doi.org/10.1016/j.rssm.2022.100755
Lee, S., & Shouse, R. C. (2011). The impact of prestige orientation on shadow education in South Korea. *Sociology of Education, 84*(3), 212–224. https://doi.org/10.1177/0038040711411278
Liang, H., Wang, Z., & Wu, W. (2022). The effect of shadow education on Hong Kong student wellbeing: evidence from PISA 2018. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2022.860179
Liao, X., & Huang, X. (2018). Who is more likely to participate in private tutoring and does it work?: Evidence from PISA (2015). *ECNU Review of Education, 1*(3), 69–95. https://doi.org/10.3092/ecnuroe2018010304
Liu, J. (2012). Does cram schooling matter? Who goes to cram schools? Evidence from Taiwan. *International Journal of Educational Development, 32*(1), 46–52. https://doi.org/10.1016/j.ijedudev.2011.01.014
Marimuthu, T., Singh, J. S., Ahmad, K., Lim, H. K., Mukherjee, H., Osman, S., & Jamaluddin, W. (1991). *Extra-school instruction, social equity and education quality [in Malaysia]*. International Development Research Centre.
Marshall, J. H., & Fukao, T. (2019). Shadow education and inequality in lower secondary schooling in Cambodia: understanding the dynamics of private tutoring participation and provision. *Comparative Education Review, 63*(1), 98–120. https://doi.org/10.1086/701064
Merlo, J., Wagner, P., Austin, P. C., Subramanian, S., & Leckie, G. (2018). General and specific contextual effects in multilevel regression analyses and their paradoxical relationship: a conceptual tutorial. *SSM, Population Health, 5*, 33–37. https://doi.org/10.1016/j.ssmph.2018.05.006
Mikhaylova, T. (2019). Shadow education—Russian edition. Presentation to the First Meeting of Asian Scandinavian Network on Shadow Education, Danish School of Education, Aarhus University, Copenhagen.
Nash, J. C. (2014). On best practice optimization methods in R. *Journal of Statistical Software, 60*(2), 1–14. https://doi.org/10.1863/jss.v060.i02
OECDa. (2022). *Chapter 6: Sample Design*. https://www.oecd.org/pisa/data/pisa2022technicalreport/
OECDb. (2022). *PISA 2022 Database*. https://www.oecd.org/pisa/data/2022database/
Park, H. (2013). *Re-evaluating education in Japan and Korea: De-mystifying stereotypes*. Routledge.
Park, H, Buchmann, C, Choi, J & Merry, J. J. (2016). Learning beyond the school walls: Trends and implications. *Annual Review of Sociology*, *42*(1), 231–252. https://doi.org/10.1146/annurev-soc-081715-074341
R Core Team. (2023). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing.
Robitzsch, A., Kiefer, T., & Wu, M. (2022). *TAM: Test Analysis Modules.* R package version 4.1–4. https://CRAN.R-project.org/package=TAM
Rosseel, Y. (2012). Lavaan: an R package for structural equation modelling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.1863/jss.v048.i02
Ryu, D., & Kang, C. (2013). Do private tutoring expenditures raise academic performance? Evidence from Middle School Students in South Korea. *Asian Economic Journal, 27*(1), 59–83. https://doi.org/10.1111/asej.12002
Sato, Y. (Ed.). (2012). *History: the 100 years of Juku and 50 years of Juku associations*. Private Tutoring Federation.
Silova, I. (2009). *Private supplementary tutoring in Central Asia*. UNESCO International Institute for Educational Planning.
Singh, K. (2015). *Report of the special rapporteur on the right to education, Kishore Singh: Protecting the right to education against commercialization.* United Nations General Assembly, Human Rights Council, 29th session, Agenda item 3. https://digitallibrary.un.org/ record/797838
Stevenson, D. L., & Baker, D. P. (1992). Shadow education and allocation in formal schooling: transition to university in Japan. *American Journal of Sociology, 97*(6), 1639–1657. https://doi.org/10.1086/229942
Su, M., & Lee, J. (2023). Eastern versus Western gaps in whole child education: how school climate works for academic achievement and well-being across China, Korea, Finland, and the United States. *Educational Research for Policy and Practice, 22*(1), 117–138. https://doi.org/10.1007/s10671-022-09322-3
Sung, K. S., & Kim, J. Y. (2010). A comparative study on the relationship between private tutoring experiences and high school students' academic science achievement: Korea, Finland and Japan. *Korean Journal of Sociology of Education, 20*(1), 103–126. https://doi.org/10.3246/ksocio.2010.20.1.005

UNESCO. (2017). *Unpacking sustainable development goal 4: education 2030*. UNESCO.
UNESCO (2021). *Non-state actors in education: Who chooses? Who loses*? Global Education Monitoring Report. https://doi.org/10.54676/XJFS2343
Verger, A., Fontdevila, C., & Zancajo, A. (2016). *The privatization of education: a political economy of global education reform*. Teachers College Press.
Warm, T. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*(3), 427–450. https://doi.org/10.1007/BF02294627
Wickham, H., Miller, E., & Smith, D. (2023). *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. R package version 2.5.2. https://CRAN.R-project.org/package=haven
Wiseman, A. W. (2021). Does inequality stay in the shadow? The cross-national effects of educational equity on private tutoring and student achievement. *International Journal of Educational Research, 109*, 101808. https://doi.org/10.1016/j.ijer.2021.101808
Yung, K. W. H. (2020). Problematising students' preference for video-recorded classes in shadow education. *Educational Studies, 48*(5), 719–726. https://doi.org/10.1080/03055698.2020.1814697
Yung, K. W. H., & Wong, S. F. P. (2024). L2 motivation in shadow education: analysing students' L2 learning experience in online English private tutoring. *Language Learning Journal*. https://doi.org/10.1080/09571736.2024.2305409
Zayet, T. M. A., Ismail, M. A., Almadi, S. H. S., Zawia, J. M. H., & Nor, A. M. (2023). What is needed to build a personalized recommender system for K-12 students' E-Learning? Recommendations for future systems and a conceptual framework. *Education and Information Technologies, 28*(6), 7487–7508. https://doi.org/10.1007/s10639-022-11489-4
Zhang, W. (2023). *Taming the wild horse of shadow education: the global expansion of private tutoring and regulatory responses*. Routledge.
Zhang, W., & Bray, M. (2020). Comparative research on shadow education: achievements, challenges, and the agenda ahead. *European Journal of Education, 55*(3), 322–341. https://doi.org/10.1111/ejed.12413
Zheng, L., Long, M., Zhong, L., & Gyasi, J. F. (2022). The effectiveness of technology-facilitated personalized learning on learning achievements and learning perceptions: a meta-analysis. *Education and Information Technologies, 27*(8), 11807–11830. https://doi.org/10.1007/s10639-022-11092-7
Zuo, M., Ma, Y., Hu, Y., & Luo, H. (2021). K-12 students' online learning experiences during COVID-19: lessons from China. *Frontiers of Education in China, 16*, 1–30. https://doi.org/10.1007/s11516-021-0001-8
Zwier, D., Geven, S., & Van De Werfhorst, H. G. (2020). Social inequality in shadow education: the role of high-stakes testing. *International Journal of Comparative Sociology, 61*(6), 412–440. https://doi.org/10.1177/0020715220984500

Publisher's Note

Mehmet Karakus is an Assistant Professor at the Centre for Global Learning, Coventry University, UK. Mehmet has 17 years of experience in higher education research and practice. He worked as an Associate Professor in the Department of Higher Education, Graduate School of Education, Nazarbayev University before this position. He has employed advanced statistical methods for data analysis and developed many original regression and structural equation models. His research focuses on emotions and emotionality in educational administration and leadership, teacher psychology, shadow education, and the well-being, personality, and job outcomes of educational leaders, teachers, and students.

Alisher Tlessov He is a PhD student at the Centre for Multilevel Modeling, University of Bristol, specializing in quantitative analysis and statistical modeling of educational data. His interests lie in applying statistical modeling techniques to analyze student achievement data and various student- and school-level factors that influence student outcomes in national assessments and international large-scale assessments (ILSAs). He has developed extensive expertise in the application of the R programming language and statistical modeling, including fitting Rasch modeling, multilevel modeling (MLM), value-added modeling (VAM), and structural equation modeling (SEM) techniques, and working with plausible values and weights in ILSAs. He has a solid background in classical test theory, item response theory (IRT), and factor analysis, demonstrating experience in these areas as well.

Anas Hajar is a graduate of Warwick University holding a PhD in Language Education. He won the Giving to Warwick Prize for his PhD thesis. He worked as a Postdoctoral Research and Teaching Fellow at Warwick, Coventry, and Christ Church Universities in the UK and as an Assistant Professor at the Chinese University of Hong Kong in Hong Kong SAR. He is currently an Associate Professor of Multilingual Education at Nazarbayev University in Kazakhstan. He is particularly interested in motivational issues in language learning and intercultural engagement. He also works in the areas of internationalization and education abroad, language learning strategies and shadow education.

Matthew Courtney is an Associate Professor of Educational Assessment and Student Achievement. He has a broad interest in student and teacher development and enjoys applying quantitative methods to answer questions about education and learning. Dr Courtney has publications in peer-reviewed journals in the fields of assessment, higher education, cyber behavior and psychology, youth academic engagement, and quantitative research methods. He has developed extensive skills and experience in the application of IRT, multilevel modelling, VAM, and SEM models.