frontiers | Research Topics

Mind the gap: To what extent do social, economic, and psychological factors explain underperformance in achievements assessments? Identifying interventions to narrow the gap

Edited by
Marcos Cupani, Ana E. Azpilicueta, Jose Juan Gongora and Sebastian Urquijo

Published in
Frontiers in Education
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-3554-7
DOI 10.3389/978-2-8325-3554-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Mind the gap: To what extent do social, economic, and psychological factors explain underperformance in achievements assessments? Identifying interventions to narrow the gap

Topic editors

Marcos Cupani — National Scientific and Technical Research Council (CONICET), Argentina

Ana E. Azpilicueta — National Scientific and Technical Research Council (CONICET), Argentina

Jose Juan Gongora — Escuela de Medicina y Ciencias de la Salud Tec Salud, Tecnológico de Monterrey, Mexico

Sebastian Urquijo — CONICET Institute of Basic, Applied and Technology Psychology (IPSIBAT), Argentina

Citation

Cupani, M., Azpilicueta, A. E., Gongora, J. J., Urquijo, S., eds. (2023). *Mind the gap: To what extent do social, economic, and psychological factors explain underperformance in achievements assessments? Identifying interventions to narrow the gap*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3554-7

Table of contents

frontiers in Education

ORIGINAL RESEARCH
published: 25 March 2021
doi: 10.3389/feduc.2021.632289

General Intelligence and Socioeconomic Status as Strong Predictors of Student Performance in Latin American Schools: Evidence From PISA Items

Carmen Flores-Mendoza[1]*, *Ruben Ardila*[2], *Miguel Gallegos*[3,4] and *Norma Reategui-Colareta*[5]

[1]Laboratory of Individual Differences Assessment, Post-Graduation Program in Neuroscience, Federal University of Minas Gerais, Belo Horizonte, Brazil, [2] Department of Psychology, National University of Colombia, Bogota, Colombia, [3]Universidad Católica del Maule, Maule, Chile, [4]Consejo Nacional de investigaciones Científicas y Técnicas, Buenos Aires, Argentina, [5]Facultad de Humanidades, Universidad San Ignacio de Loyola, Lima, Peru

Numerous technical—scientific reports have demonstrated that student performance variability is linked to several factors, especially socioeconomic factors. For a century, differential psychology has shown that students' socioeconomic level has little or no relevance in the explanation of student performance variation when the intellectual factor is considered. Here we present a study on a student samples (N = 1264) aged 13 to 16 yrs, enrolled in 32 schools from five Latin American countries (*Argentina*, Brazil, Chile, Colombia, and Peru). A short version of the PISA test (composed by 16 items) and five cognitive measures were administered, in addition to a socioeconomic questionnaire. Multilevel analysis (marginal models) indicated that general intelligence (*g*-factor) and socioeconomic school status were robust predictors, and the students' socioeconomic status very little accounted for the variation in the PISA test. This study concludes that education policy must incorporate individual differences in intelligence, beyond socioeconomic variables, as an important predictor variable in student performance studies.

Keywords: intelligence, g factor, PISA, latin america, school performance

OPEN ACCESS

Edited by:
Jose Juan Gongora, Tecnológico de Monterrey, Mexico

Reviewed by:
Giray Berberoglu, Başkent University, Turkey
H. Cigdem Yavuz, Çukurova University, Turkey

***Correspondence:**
Carmen Flores-Mendoza
carmencita@fafich.ufmg.br

Specialty section:
This article was submitted to Assessment, Testing and Applied Measurement, a section of the journal Frontiers in Education

Received: *22 November 2020*
Accepted: *03 February 2021*
Published: *25 March 2021*

Citation:
Flores-Mendoza C, Ardila R, Gallegos M and Reategui-Colareta N (2021) General Intelligence and Socioeconomic Status as Strong Predictors of Student Performance in Latin American Schools: Evidence From PISA Items. Front. Educ. 6:632289. doi: 10.3389/feduc.2021.632289

INTRODUCTION

At the end of the 1990s, the Organization for Economic Co-operation and Development (OECD) envisaged the increasing importance of education in the development of skills that would allow citizens to adapt and absorb rapid changes in technology. From this, the OECD developed and promoted in 2000 a large-scale assessment of 15-year-old students through a test termed PISA (The Programme for International Student Assessment). The PISA test is an assessment tool, conducting three-yearly surveys, that scores reading, mathematic and scientific literacies. The focus of this assessment is not surveying memorization or simple knowledge. The PISA test items focus on how well students apply knowledge to solve real-world problems (OECD, 2001). In the first survey (2000–2001) 43 countries participated in the PISA assessment, which increased to 79 countries in the last survey, conducted in 2018. After seven PISA surveys, the result has been consistent, where students from developed countries present better performance than students from developing

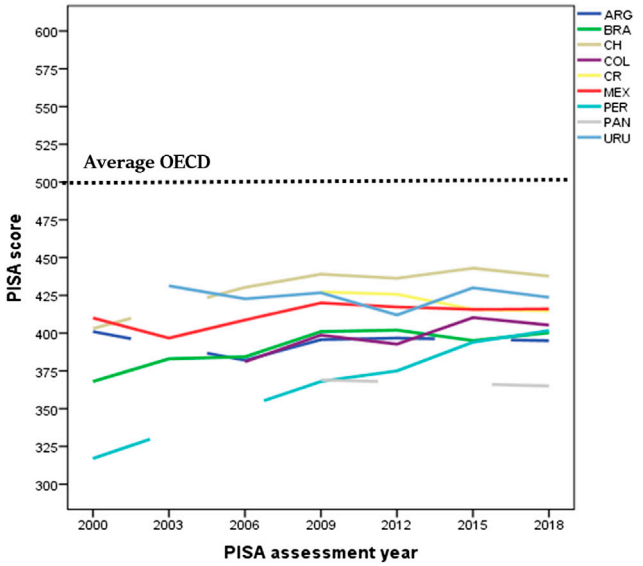


FIGURE 1 | Average PISA score of each Latin American country over time (except the Dominic Republic which only participated in the 2015 PISA assessment).

countries. Presented in this way, this result over time suggest the hypothesis that education drive national economies forward. In this regards, studies estimated that an increase of 0.5 standard deviations in PISA scores, would lead to an increase in national Gross Domestic Product per capita of up to 5% (Hanushek and Woessmann, 2007; Hanushek and Woessmann, 2015). Thus, it is not surprising that most nations are focused on these findings regarding the basic or fundamental skills for the development of their citizens and socioeconomic impact.

Latin American Region

Education foster national economic growth, and this evidence has been accepted by some Latin American governments. Despite an expected unsatisfactory result, five Latin American countries participated in the first PISA survey (2000–2001), and nine countries (*Argentina*, Brazil, Chile, Costa Rica, Colombia, México, Peru, Panama, and Uruguay) participated in the most recent PISA survey (2018). Only two countries (Brazil and Mexico) participated in all surveys. Considering the average of all assessments, Chile and Uruguay have had the highest mean score in the PISA test, while the Dominican Republic, Panama, Peru, and Brazil have had the lowest. In general, all participating Latin American countries performed below the OECD average, which formed a relative cluster within the general picture of the PISA assessment (**Figure 1**).

The most worrying result was that a large share of Latin American students underperformed in level 2 (out of six levels). For instance, in the first 2000 PISA survey, which emphasized reading skills, the percentage of students that performed at level 1 (students that show basic skills) plus below level 1 (students that are not able to show most basic skills), varied between 44 and 80%. This considerable percentage did not change in the next assessments (48–75% in 2003- mathematic emphasized; 40 to 60% in 2006- sciences emphasized; 40 to 64% in 2009- reading emphasized; 51.5 to 74.6% in 2012-mathematic emphasized; 23.3 to 46.7% in 2015-science emphasized, and 35 to 79% in 2018-reading emphasized). Translated to years of schooling, these results are equivalent to a gap of 1.7 years of schooling for the Latin American country with the best performance (Chile), and a gap of 3.1 years of schooling for the Latin American countries with the lowest performance (Peru and Colombia), compared to the OECD countries (OECD, 2016). Note that the estimation of the schooling gap is independent of the subject assessed (Math, Reading or Science), given the high correlation (above 0.80) among them. In 2018, the proportion of Latin American human capital capable of understanding complex situations and provide innovative solutions (top performers) varied between 0.1% (Dominican) to 3.5% (Chile) compared to 15.7% from the OECD average (OECD, 2019a). The Latin American results dramatically contrasted those observed in some Asian countries (e.g. China, Singapore, Korea) and some European countries (e.g. Ireland, Finland, Poland, Estonia), where the majority of students (60%) perform at level 3 above, and there is a proportion of top performers, ranging from around 20% (Finland, Ireland) to 45% (ex. China, Singapore).

Factors Pointed Out as Predictors of the Performance in the PISA Test

Since the first PISA assessment, a significant number of publications have been produced. For instance, between 1999–2015, a thousand documents (Hopfenbeck et al., 2018) analyzed several factors that could explain the variation in student performance. Among these factors were educational features (e.g. repetition rates, enrollment rates in tertiary education, attending pre-primary school, financial capacity to provide quality education services, spending per students, number of teachers per student, teachers' salaries, percentage of teachers with at least a master's degree), gender differences, family background (parental occupational status, parental education, family wealth, parents' expectations for their child's future), school's socio-economic composition, characteristics of high learners (motivation, attitudes, self-related beliefs, anxiety, learning habits, life satisfaction), exposure to bullying, learning engagement, student truancy, immigrant status, access to internet, spending time online outside of school, spending time playing videogame, or school and classroom climate (OECD, 2001; OECD, 2004; OECD, 2005; OECD, 2007; OECD, 2010; OECD, 2013; OECD, 2016; OECD, 2018; OECD, 2019a).

From all factors analyzed, those related to socioeconomic background variation have received considerable attention of education policy makers and researchers (Coleman et al., 1966; Avvisati, 2020). In the 2018 PISA survey, the 10% most socioeconomically advantaged students outscored their 10% most disadvantaged counterparts in reading by 1.5 standard deviation (150 points or three years of schooling). This gap in school performance has persisted over the last decade, despite a 15% increase in education spending (Schleicher, 2019). Regarding

Latin American countries, the OECD (2016) has identified that students from some countries (e.g. Brazil and Argentina) underperformed students from countries with the same level of economic development (e.g. Thailand and Bulgaria). Additionally, simulations showed that even if Latin American students had the OECD average socio-economic status, there would be an increase of 28 points on PISA average scores, but there would not be changes in the general ranking. On the other hand, between 6% to 20% of the PISA variation was explained by the socio-economic status of Latin-American students, proportions that are not so different to the OECD average (15%), however, when the inter-school socioeconomic status (between-school) is taken into account instead inter-student socio-economic status (within-school), a strong association with student performance is revealed. For instance, in México, a one-unit increase in the socio-economic status of students is associated with an increase of five points in mathematics, but a one-unit increase in the school socio-economic status is associated with 30 points in mathematics. This kind of results was observed in all PISA surveys.

Despite the gathering of information and analysis of a wide range of psychosocial variables, no PISA survey considered the administration of intelligence measures. Historical and cultural reasons may underlie why education policies take no notice of the concept of intelligence (Maranto and Wai, 2020). As this study will demonstrate, intelligence exerts a strong influence on student performance beyond socioeconomic factors, a critical point that has been ignored in the educational field.

Intelligence and Student Performance

It is not our intention to elaborate on the history of differential psychology, but it is worth remembering that the Stanford-Binet scale was the first intelligence test created in the beginning of 20th century for educational purpose (Terman, 1916). A century has passed since the creation and massive use of intelligence tests, and countless studies have indicated that, independent of the applied cognitive measure, it correlates significantly with student performance and, consequently, explains a significant part of the student performance variation (between 20 and 40%) (Roth et al., 2015). For example, Strenze (2007) conducted a meta-analysis of 85 longitudinal datasets, where predictors (intelligence, parental SES, and student performance) were surveyed at an earlier time and the dependent variable career success (composed by education, occupation, and income) at a later time, minimum three years between the surveys. Regarding education, intelligence was the stronger predictor than the other two predictors. Although other psychological factors (e.g. motivation, self-control, personality) also correlate with any aspect of education, intelligence is the best single predictor (Kuncel et al., 2004; Leeson et al., 2008) which has been recognized by the world's most influential intelligence researchers (Neisser, Boodoo, Bouchard Jr., Neisser et al., 1996; Gottfredson, 1997; Hunt, 2010).

However, there is a particular issue in the literature about the relationship between intelligence and its correlates, especially those of education. The correlation between education and intelligence is stronger when intelligence is represented at general level instead measured by a score on a specific ability test (Coyle, 2015; Cucina et al., 2016; Costa et al., 2018). General intelligence is represented by a g-factor, which refers to the broader mental ability extracted from a correlation matrix of a battery of diverse and reliable cognitive tests. According to Jensen (1998), independently of the specificity of the information content, skill, or strategy of the mental tests, the *g*-factor is the source of variation associated with the efficiency of neural processes that affect cognitive behavior. If the *g*-factor is the best estimate of intelligence, stronger correlations are expected between this level of cognitive generality and student performance, than with specific abilities.

On the other hand, the significant relationship between student performance and intelligence has been verified through studies that use individual–level data designs. From the present millennium, the same relationship was identified in studies working at national-level data (Lynn and Vanhanen, 2002; Lynn and Becker, 2019). According to these studies, the intelligence of the nation relates to several educational outcomes such as technological achievement over a millennium (from 0.42 for 1000BC to 0.75 for 2000 AD; Lynn and Becker, 2019), adult literacy (r = 0.64; Lynn and Becker, 2019), patents indexes (r = 0.51; Gelade, 2008); Nobel prize in science (r = 0.34; Rindermann et al., 2009); technology exports (r = 0.38; Rindermann et al., 2009). Moreover, the correlation between intelligence of nations and international student assessment such as TIMMS (Third International Math and Science Study), PISA, PIRLS (Progress in International Reading), IEA-Reading (International Association for the Evaluation of Educational Achievement), IAEP-II (International Assessment of Educational Progress) was not less than 0.80 (also see in Rindermann, 2007; Rindermann, 2018; Lynn and Becker, 2019), a value that is much higher than what is obtained in studies that use data at the individual level. Not surprisingly, the strong correlation between intelligence and education assessment at national level led some differential psychologists to asserts that, empirically and theoretically, there is no significant differences between them (Rindermann, 2007). However, cross-nation estimates rely on aggregated data., i.e., multiple sources of school and intelligence assessments compiled into data summaries. Data of international school assessment usually are reliable and use representative sample, while cognitive data of nations usually come from small studies that use unrepresentative samples, present insufficient information regarding the quality of the tests, and were administered at different years of the XX century. Thus, the conclusion that intelligence and student performance is the same phenomenon has been built on fragile data sources. Furthermore, if the relationship between these two variables is almost perfect, it would be expected that the factors influencing intelligence must influence student performance in the same intensity. However, there is reasonable evidence that it does not happen. For instance, there is a certain consensus that student performance is sensitive to socioeconomic factors (e.g., high SES students outperform low SES students) (Daniele, 2021), while intelligence seems to be less affected by socioeconomic differences (O'Connell and Marks, 2021). Strong evidence that genetic components of intelligence are not moderated by socioeconomic factors (SES) is the study

TABLE 1 | Descriptive analysis of number of students per school and country.

Variables	N	Mean	S.D.	Min.	1°Q	2°Q	3°Q	Max.
School	32	40.72	18.83	15.00	29.00	33.50	50.50	111.00
Country	5	260.60	113.52	168.00	186.00	199.00	314.00	436.00

conducted by Hanscombe et al. (2012), where 8,716 twin pairs clustered in eight ages (from infancy through adolescence) were analyzed. The genetic effect on intelligence did not differ for low and high SES groups; however, a shared environment (e.g., parental education, family income, occupation) influenced a little more the low SES families than high SES families. This influence decreased with age, meaning that intelligence is influenced differently by the shared environment and genetic factors throughout the life cycle. On the other hand, age affects intelligence sooner than it affects education (Lenehan et al., 2015). In the case of fluid intelligence, which matches the *g*-factor, it reaches a plateau between the end of adolescence and early adulthood (Hartshorne and Germine, 2015), i.e., there is no significant increase in performance on non-verbal cognitive measures after 18–20 years of age. That is one reason why intelligence (or *g*-factor) is considered a biological phenomenon with ontogenetic characteristics (Jensen, 1998). Regarding education performance, age can act negatively through the distortion age-grade, which can be related to individual cognitive differences (promotion-delay) or the delay that students enter the school system (a phenomenon named RAE-relative age effect; Juan-Jose et al., 2015). Sex is another variable that may affect intelligence and student performance differently. For instance, there is controversy about whether sex affects specific cognitive abilities or affects the *g*-factor (Halpern et al., 2020). The sex effect on education is only on specific domains such as math, favoring males, and reading, favoring females (see Trucco, 2014 for data from Latin American countries). Hence, intelligence and student performance are not twin constructs, but it is recognized that they may exhibit a strong dependence on each other.

Moreover, the dependence degree between student performance and intelligence seems to vary according to the development degree of nations. Since the famous Coleman report (Coleman et al., 1966), countless studies have confirmed the main result of that report that 80 to 90% of the total variance in student performance was due to students' characteristics, and between 10 to 20% was due to characteristics of schools. However, these results fit well in developed countries, not in developing countries. The review of the Coleman report (as it was known) after 40 years conducted by Gamoran and Long (2006) with data of developed and developing countries indicated that schools might account for 57% of the student performance variance in the Latin American region.

In this sense, to verify whether intelligence and student performance reveal a strong dependency regardless of cultural settings, it requires overcoming the use of compiled data. To our knowledge, no cross—national empirical study has been conducted using simultaneously an international school test such as PISA test and cognitive measures. Moreover, no cross-national study used individual-level data and analyzed the influence of intelligence at the latent level (*g*-factor). From this, the SLATINT project (Study of Latin American Intelligence) came to be developed. We considered that the obtained results are pertinent to the proposal of the present edition.

The SLATINT Project

In 2007, a group of Latin American researchers from six Latin American countries (Argentina, Brazil, Chile, Colombia, Mexico, and Peru) designed a large-scale assessment using the PISA test and several cognitive measures in each Latin American country. The project was termed "Study of Latin-American Intelligence" (SLATINT). To that end, measures, questionnaires, general and specific instructions for data collection, logistics for sending the material to the participating countries, receipt, examination, and codification of the protocols were planned by researchers in face-to-face and virtual meetings (video call). The project was conducted between 2007 and 2010, which included a total sample of 4,074 students. The SLATINT results can be summarized in three points: 1) positive relationship between the PISA test and cognitive measures, although a stronger correlation was observed as aggregated, rather than when individual scores were used; 2) after controlling social variables, the PISA scores could present stronger variability due to the variations in cognitive scores; 3) the socioeconomic status of schools had a greater influence on PISA scores than the socioeconomic status of students, and 4) Sex and age differences did not affect cognitive measures, but slightly affected the PISA test (Flores-Mendoza, et al., 2015; Flores-Mendoza et al., 2017). However, the obtained results were based on the administration of just one cognitive measure. Recently (Flores-Mendoza et al., 2018), it was analyzed the relationship between the PISA test score and intelligence differences at the latent level using a generalized linear mixed model, where the individual is the target (subject-specific model) of inference. The obtained results were similar to previous studies.

Propose of this Study

This paper aims to present the results based on a population-averaged model, also named marginal model, regarding the influence of a set of predictors (sex, age, kind of school, SES of schools, SES of students, g-factor) on the PISA test using the SLATINT data. Marginal models are robust and less susceptible to biases from misspecification of random effects (Heagerty and Kurland, 2001). Unfortunately, the Mexican sample was small (N = 66), and recruited only in a private and high SES school. Thus, without variation in SES school, data from Mexico was not included in the analysis.

TABLE 2 | Sociodemographic characteristics of the studied samples.

Participant characteristics	Total (*N* = 1303)	Arg (*N* = 436)	Bra (*N* = 186)	Ch (*N* = 168)	Co (*N* = 199)	Pe (*N* = 314)
Sex	%	%	%	%	%	%
Female	50.5	52.3	53.7	48.8	44.2	51.0
Male	49.5	47.7	46.3	51.2	55.8	49.0
Age						
13	3.2	0.7	9.1	5.3	0.0	3.8
14	55.7	51.1	60.8	77.4	59.3	45.3
15	37.7	44.3	23.1	16.7	39.7	47.1
16	3.4	3.9	7.0	0.6	1.0	3.8
School characteristics						
Private	52.3	48.9	36.0	47.0*	16.6	92.0
Public	45.4	51.1	64.0	34.5	83.4	8.0
SES school						
Low	33.7	29.6	32.8	34.5	83.4	8.0
Middle	28.8	37.8	15.6	34.5	16.6	29.0
High	37.5	32.6	51.6	31.0	0.0	63.0
Parents education						
Father						
College	51.6	42.7	45.2	62.4	11.2	83.8
High school	31.2	33.0	31.2	26.1	58.0	14.7
Primary school	17.2	24.3	23.6	11.5	30.8	1.4
Mother						
College	71.5	42.1	42.4	55.8	11.0	78.4
High school	14.0	38.2	35.1	37.0	59.7	19.9
Primary school	14.5	19.6	21.5	7.2	28.3	1.7

Note: Arg = Argentina. Bra = Brazil. Ch = Chile. Co = Colombia. Pe = Peru.

METHODS

Participants

1303 students enrolled between grade 8 and 10 (73% ninth-grade) from 32 schools and five Latin American cities (Rosario-Argentina, Belo Horizonte-Brazil, Santiago-Chile, Bogota-Colombia and Lima-Peru) participated in this study, which it was conducted between 2007 and 2011 (80% in 2008–2009). **Table 1** shows the average number of students per school (mean = 40.72; min = 15 and max = 111). The average number of students per country was 260.6 (min = 168 and max = 436).

The sociodemographic characteristics of the sample by country are shown in the **Table 2**. As it can be seen, the samples are not representative of their countries. For instance, according to the statistics of the Economic Commission for Latin America and The Caribbean (https://www.cepal.org/en), the percentage of population aged 25 and 59 years with schooling above high school does not exceed 30%. However, in our samples, except for Colombia, there was a high percentage of parents (especially Peruvian mothers) with tertiary education. Additionally, there was a percentage of private schools that participated which was not expected for some countries. For example, 20% of all Brazilian students are enrolled in private schools, however in our sample, 46% of the Brazilian students in this study were enrolled in private schools. This occurrence was even more pronounced when considering Peru, where almost all schools were private (92%). In Colombia, 60% of students are enrolled in public schools, however, in our study 83.4% of students were studying in public schools. In Argentina, 70% of students are enrolled in public schools (Vior and Rodríguez, 2012), however, in our study 51% of students were enrolled in public schools. The Chilean sample was more representative of Chile, where 45% of students are enrolled in private schools (Bellei, 2008), almost the same proportion found in our study. In general, with the exception of the Colombian sample, students came from families and educational backgrounds with better resources than the average of the Latin American population.

Measures

PISA 2003–short version. The PISA 2003, complete version, contained 85 items distributed in four clusters of mathematical areas (Space and shape, Change and relationships, Quantity, and Uncertainly), which required to activate three cognitive skills groups (Reproduction–simple mathematical operation; Connection–bringing together ideas to solve straightforward problems, and Reflection–wider mathematical thinking) (OECD, 2004; page 24). A short version was available on the website of the Brazilian Ministry of Education. This version contained 29 items, which were in a mixture of multiple-choice and constructed-response formats. Despite their format, all the items requested only one right answer. A pilot study with 181 Brazilian students indicated an alpha coefficient of 0.906, and it took, on average, 2 h. A reduction version was necessary due to the limited time offered by the schools for administering all the instruments proposed by the project. However, the shorter version had to preserve the accuracy and validity of the previous version. To accomplish such requirements, the item response theory (IRT) was used. IRT is a model that assumes that each item within a scale is a measure of some underlying construct, and the latent variable causes the observed item responses. This model detects the error variance (measurement errors) and provides a test of overall model fit and model fit indices. We conducted a Rasch analysis (a special case of IRT) for dichotomous items using the software WINSTEPS 3.63.2 (Linacre, 2007). It was detected that by deleting a maximum of 13 questions, the person separation reliability (used to classify people) of the new version of the PISA test (16 items) was 0.875, a value considered acceptable. All 16 items showed fit indices between 0.50 and 1.50, meaning good fit indices. Rasch factor analysis indicated a 62.1% of the variance explained, which supported the hypothesis of unidimensionality and an eigenvalue of 2.4 (or 3.2% of the variance) explained by the first contrast. This last result indicated a minimal deviation from de unidimensionality, but it was not considered a threat to the short PISA test version's validity. The set of 16 items were representative of Space and Shape (n = 3), Change and Relationship (n = 5), Quantity (n = 4), and Uncertainly (n = 4), and they demanded Reproduction (n = 8), Connection (n = 6), and Reflection (n = 2) skills. Example of item of each area, kind of skills demanded by each item (according to OECD, 2013), and results from Rasch model are in Supplemental Material. In order to extend the validity of this version, a second pilot study with PISA-16 items was conducted in a sample of 167 Brazilian students. The new version took, on average, 1 h and 15 min. The reliability of the 16-item version (Cronbach's alpha) was

0.844, and it was associated to the Raven test at 0.650. The correlation between the 19-item and 16-item version was 0.970. Thus, the shorter version of the PISA test preserved its reliability and validity. Native Portuguese and Spanish speakers conducted double-check translation of the PISA test (Portuguese to Spanish language). In the present study, the Cronbach Alpha (reliability) for the total sample was 0.807, varying from a minimum of 0.706 (Colombian sample) to a maximum 0.835 (Brazilian sample).

Standard Matrices Progressives of Raven (SPM). The SPM was the cognitive measure used in this study. This non-verbal exam is the most frequently used test to study cognitive differences at the individual, as well as at the national level (Lynn and Vanhanen, 2012). Additionally, the SPM is considered a good measure of basic cognitive functioning (Raven et al., 2000; Jensen. 1998). In the present study, the Cronbach Alpha (coefficient of reliability) of SPM was 0.885, varying from a minimum 0.859 (Colombian sample) to a maximum of 0.916 (Argentina sample). Test takers were allowed 45 min to complete the SPM test.

Berlin Intelligence Structure Model (BIS tasks). Four subtests of the BIS battery (Rosas, 1996), which took between 1 to 2 min to complete, were administered to the samples. These were: BIS_MF (a figural short term memory test), BIS_PN3 (a numerical reasoning test), BIS_RN3 (a numerical reasoning test), and BIS_RN1 (a numerical simple mental speed test). The Cronbach Alphas were 0.870 (BIS_MF), 0.647 (BIS_PN3), 0.905 (BIS_RN3), and 0.812 (BIS_RN1).

Socio-economic questionnaire for students. There was no standardized Latin American approach to measure socioeconomic status. For this reason, the Latin American researcher team defined that the estimation of the SES student would be based on available resources found at their home (e.g. cable TV, MP3Player, Phone, Computer, Internet, Videogames, and Weekend Magazine), and parents level of education (mother and father). Each item of available resources in home represented one point. Regarding education of parents, the lowest level of schooling was equivalent to primary school and the higher level was college.

Socio-economic classification and questionnaire for schools. Schools were classified as low, middle, and high SES in each country. At least two representative schools from each socioeconomic stratum were required. Samples of schools from Peru and Brazil were randomly selected, however the collection data for all cognitive measures in Peru was not attained in low SES schools. School samples from Chile, Argentina and Colombia were non-probability samples. In these cases, researchers selected schools based on their available knowledge about school infrastructure and socioeconomic characteristics of the community where the schools were located. In order to validate their subjective appreciation, researchers responded to a questionnaire regarding sanitary and urban conditions (e.g. waste collection system, drainage system, public street lighting, etc.), and items regarding school environment (e.g. school instruction time, class size, mathematic instruction time, presence of computers). The points accumulated in this questionnaire were correlated to the SES school classification performed by the researchers. The result was a r of 0.72 (p = 0.05) for Chile and 0.63 (p = 0.03) for Colombia. The Argentinean researcher could not collect information related to this questionnaire. So in this case, we correlated the Argentina classification of SES school with SES of students (r = 0.610), education level of father (r = 0.641) and mother (r = 0.671). We considered all these results as evidence of validity of the SES classification of schools.

Analysis

The dataset was composed by 32 schools and samples from five countries. There were 39 missing values for PISA score, thus all variables related to these cases were eliminated of the dataset.

Absolute and relative frequencies were used for qualitative variables, and measures of central tendency and dispersion were used for quantitative variables, and Eta correlation in cases of nonlinear relationship.

Intelligence was represented at the latent level, using the five cognitive measures (described in Instruments). These cognitive measures were subjected to principal axis factoring (PAF), which analyzes only the common factor variance of the tests. Inspection of the correlation matrix (**Table 3**) revealed the presence of coefficients of 0.3 and above. The Kaiser-Meyer-Olkin value was 0.802, and the Barlett's Test of Sphericity reached statistical significance, supporting the factorability of the correlation matrix. PAF analysis revealed the presence of only one factor explaining 40% of the variance. We used the factor score as a representative of intelligence at the latent level (g-factor).

Our dataset can be considered a clustered data with hierarchic structure: students within schools, and school within country. The statistical approaches that address the description of systematic variation in the mean response as well as associations among observations within clusters include marginal models fit with generalized estimating equations (GEE). Our interest was the estimation of overall population average relationships between independent variables (e.g. g factor, sex, age, SES) and dependent variables (e.g. PISA score) across all of the different clusters. The term 'marginal mean' refers to the averaging over both measurement errors and random interindividual heterogeneity. This model does offer advantages over other approaches for dependent data. First, GEE has been popularly applied because it is the easiest to understand and it is more relaxed when considering distribution suppositions or when there are variables that are not continuous. Second, marginal models allow inferences about overall marginal relationships and permit calculate robust standard errors that reflect the sampling variance in the estimated parameters that arises from the clustered study design. Marginal model is considered a population-level approach and it provides the population-averaged estimates of the parameters. Thus, the target inference is the population (Liang and Zeger, 1986).

A symmetric working correlation structure was inserted in the population-averaged model to account for the correlation among students from the same school. To estimate the parameters of the population-average model, estimation equations proposed by Prentice, 1998) with the 'geese.fit' function of the 'geepack' package of R software (version 3.3.1) were used.

TABLE 3 | Correlation matrix with socioeconomic variables, cognitive and PISA measures.

	PISA	SES student	Sex	SPM	BIS MF	BIS PN3	BIS RN3	BIS RN1	*g*
PISA	1	0.410**	−0.061*	0.575**	0.331**	0.511**	0.408**	0.485**	0.643**
SES student	0.429**	1	−0.074*	0.370**	0.178**	0.318**	0.239**	0.265**	0.289**
SPM	0.614**	0.419**	−0.041	1	0.372**	0.470**	0.349**	0.409**	0.653**
BIS MF	0.321**	0.186**	0.022	0.392**	1	0.319**	0.302**	0.329**	0.476**
BIS PN3	0.518**	0.332**	−0.108**	0.507**	0.311**	1	0.450**	0.448**	0.749**
BIS RN3	0.420**	0.273**	−0.219**	0.387**	0.302**	0.451**	1	0.498**	0.609**
BIS RN1	0.489**	0.295**	−0.024	0.439**	0.328**	0.443**	0.476**	1	0.701**
g	0.649**	0.321**	−0.124**	0.639**	0.468**	0.750**	0.597**	0.704**	1

Note: SPM = Standard Progressive Matrices of Raven; BIS MF = figural short term memory test; BIS PN3 = numerical reasoning test; BIS RN3 = numerical reasoning test; BIS RN1 = numerical simple mental speed test; g = g-*factor (or general intelligence).*

TABLE 4 | PISA results according sociodemographic variables.

Variables		N	Mean	S.D.	Min.	1°Q	2°Q	3°Q	Max.
Country	Argentina	435	7.26	3.77	0.00	4.50	7.00	10.00	15.00
	Brazil	152	8.04	3.75	1.00	5.00	8.00	11.00	16.00
	Chile	167	6.56	3.93	0.00	3.00	6.00	10.00	15.00
	Colombia	196	5.85	3.00	0.00	3.50	6.00	8.00	14.00
	Peru	314	8.36	3.84	0.00	6.00	8.00	11.00	16.00
Sex	Female	642	7.09	3.78	0.00	4.00	7.00	10.00	16.00
	Male	622	7.55	3.80	0.00	5.00	8.00	10.00	16.00
Age	13	40	7.28	3.94	0.00	4.00	8.00	11.00	14.00
	14	698	7.40	3.88	0.00	4.00	7.00	10.00	16.00
	15	486	7.25	3.71	0.00	5.00	7.00	10.00	16.00
	16	40	6.78	3.03	0.00	4.00	7.00	9.00	14.00
Kind school	Others*	31	7.97	3.21	2.00	6.00	8.00	9.50	15.00
	Private	678	8.92	3.45	0.00	7.00	9.00	12.00	16.00
	Public	555	5.32	3.24	0.00	3.00	5.00	8.00	14.00
SES school	Low	407	4.45	2.88	0.00	2.00	4.00	6.00	14.00
	Middle	375	7.51	3.21	0.00	5.00	7.00	10.00	15.00
	High	482	9.58	3.26	0.00	7.00	10.00	12.00	16.00
Father educational level	College	639	8.77	3.57	0.00	6.00	9.00	11.00	16.00
	High school	402	6.22	3.49	0.00	3.00	6.00	9.00	15.00
	Primary	223	5.12	3.21	0.00	2.00	5.00	7.00	14.00
Mother educational level	College	893	8.27	3.63	0.00	6.00	8.00	11.00	16.00
	High school	182	5.46	3.24	0.00	3.00	5.00	8.00	14.00
	Primary	189	4.59	2.99	0.00	2.00	4.00	7.00	12.00

**mix school (public and private).*

The model was initially adjusted with all the explanatory variables of interest and, later, the Backward method (Efroymson, 1960) was applied for the final selection of the variables. The Backward method is the procedure of removing the variable with the highest *p*-value, and the analysis is repeated until only significant variables remain in the model. In the Backward method, a level of significance of 5% was adopted.

Specification of the Marginal Model

Considering y_{ijk} the value of response (Pisa or *g*-factor) for i-country, j-school, and k-student, the μ_{ijk} denotes the mean value expected for the response of the i-country, in the j-school and for the k-student. Hence, considering P explanatory variables $x_1, x_2, \ldots, x_p$, we have the following model for the mean:

$$\log\left(\mu_{ijk}\right) = \sum_{p=1}^{P} \beta_p x_{pijk}.$$

The correlation among students from the same school was computed by symmetric working correlation structure:

$$\mathrm{Corr}\left(Y_{ijk,}\, Y_{ijl}\right) = \begin{matrix} 1, & \text{se } k = l, \\ \alpha, & \text{se } k \neq l. \end{matrix}$$

Considering discrete data and the possibility of over or under dispersion of the data, the variance was computed by:

$$\mathrm{Var}\left(Y_{ijk}\right) = \varnothing \mu_{ijk},$$

Where $\varnothing$ is a common scale parameter and μ_{ijk} is a known variance function.

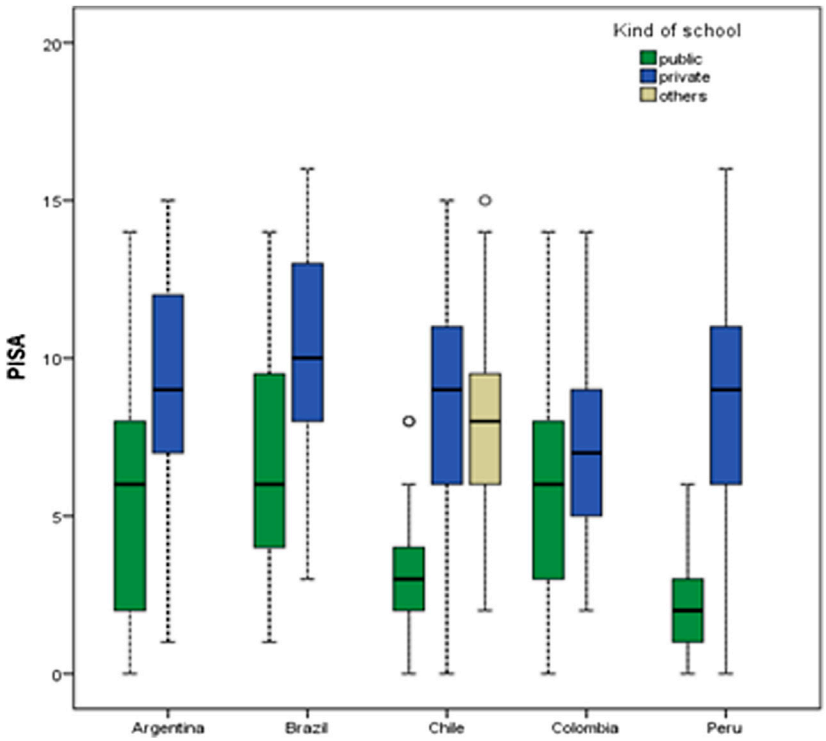


FIGURE 2 | Boxplot of PISA score according kind of school and country.

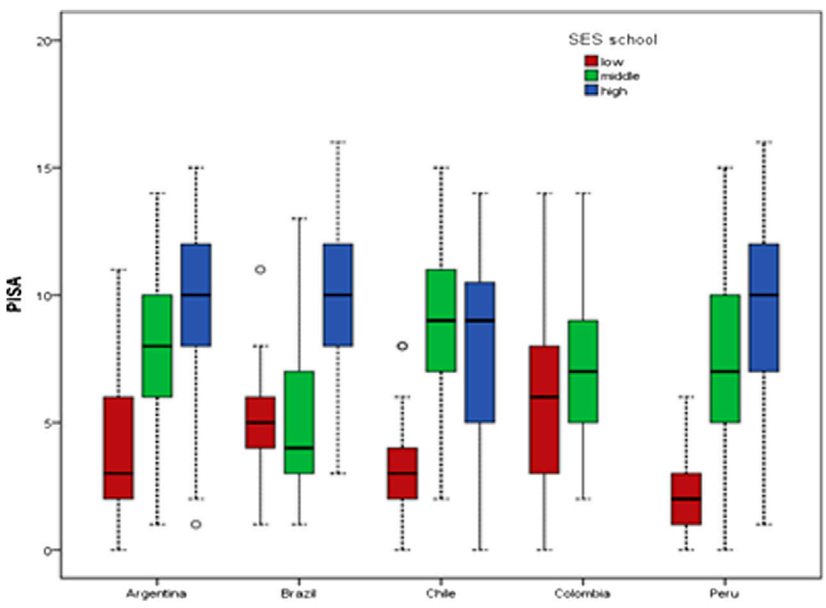


FIGURE 3 | Boxplot of PISA score according SES school and country.

RESULTS

Descriptive Statistic

Table 4 indicates that the Peru and Brazil samples had highest PISA test mean scores, while the Colombian sample had the lowest mean score. The highest PISA test score was presented by the Chilean sample. Males had higher score than females, and 16-yrs old students had a lower score than 14-yrs old students. Students whose parents had a high level of education outscored students whose parents had a low level of education.

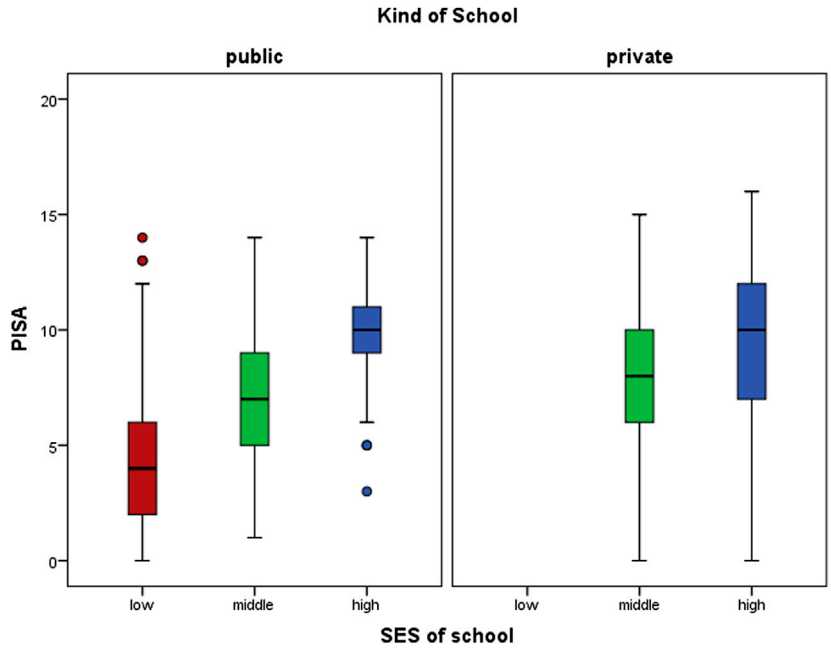


FIGURE 4 | Boxplot of PISA score according SES school and kind of school.

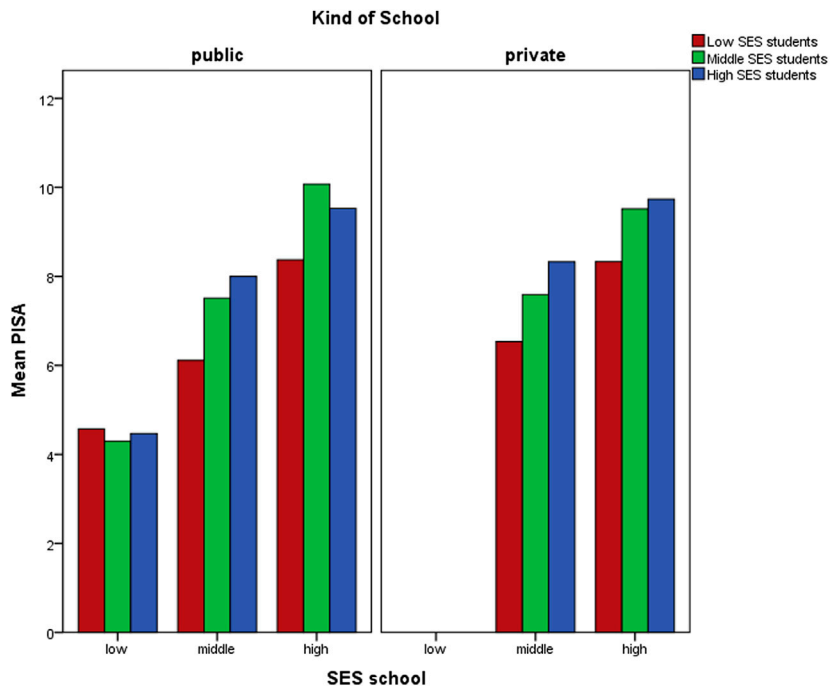


FIGURE 5 | Distribution of mean PISA score in each SES school, according to SES student classification, and kind of school.

Figure 2 shows differences between kind of schools. For all samples, private schools outscored public schools. The between-schools gap was more pronounced in the Chilean and Peruvian samples and less pronounced in the Colombian sample.

Figure 3 shows that high SES schools had higher PISA test mean score than middle SES schools, and middle SES schools had higher PISA test mean score than low SES school. Note that there was no high SES school in the Colombian sample.

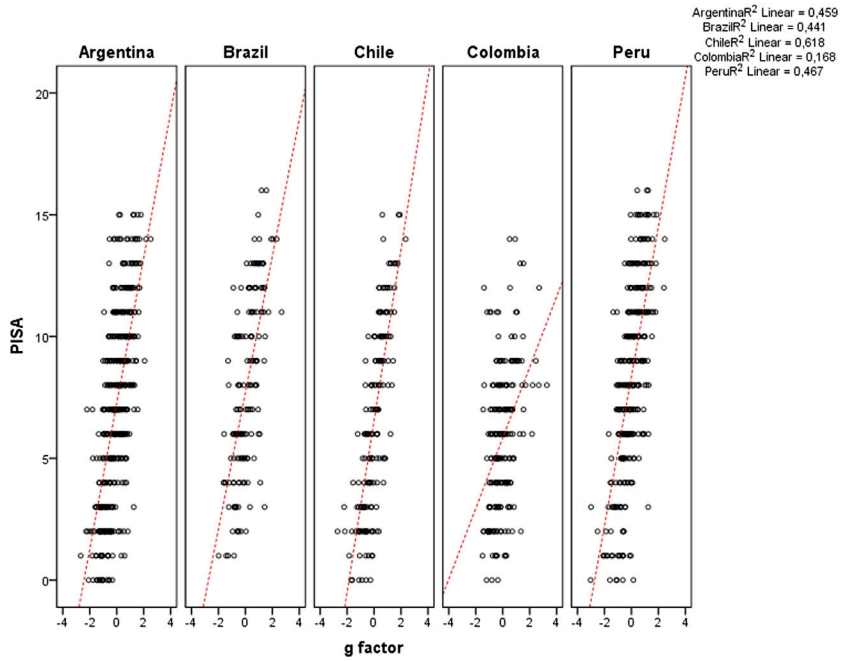


FIGURE 6 | PISA score vs. *g* factor for each country.

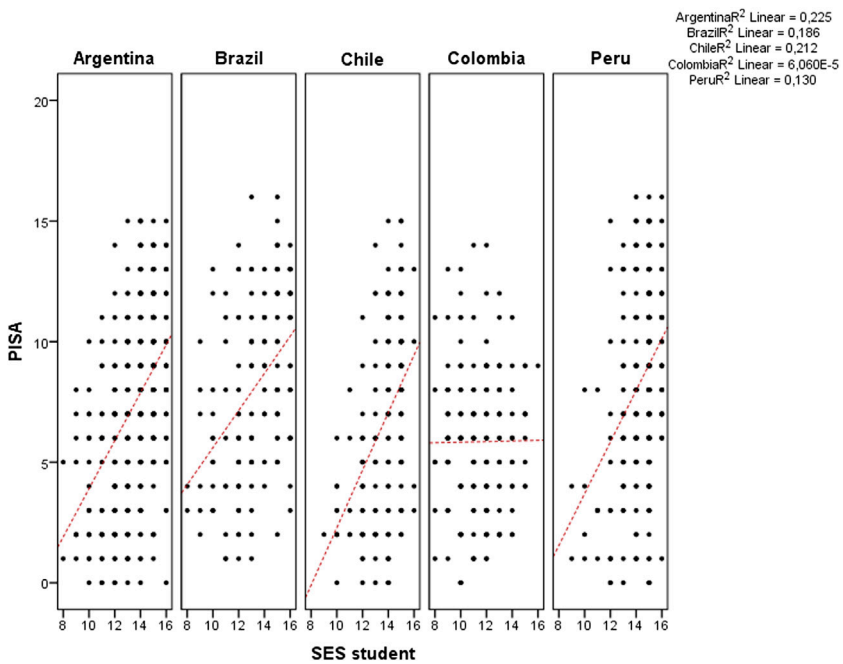


FIGURE 7 | PISA score vs. SES of students for each country.

Figure 4 shows that high SES schools presented higher PISA test score than middle SES schools, independent of being private or public. Middle SES school (public and private schools) showed higher PISA test score than public SES school.

SES student scores were converted to percentiles ($p < 25$, P50, and $p > 75$). **Figure 5** shows that independent of individual SES and kind of school, students who were enrolled in high SES schools outperformed students who were enrolled in low SES schools.

TABLE 5 | Marginal effects for log-linear regression–Final model for PISA.

Variables	β	s.e (β)	*p*-value	C.I.-95%
Intercept	–0.79	0.16	0.000	–
g factor	0.57	0.04	0.000	[0.480; 0.65 0]
SES student	0.02	0.01	0.020	[0.000; .04 0]
SES school = low	—	—	—	—
SES school = middle	0.51	0.13	0.000	[0.250; 0.77 0]
SES_school = high	0.82	0.14	0.000	[0.550; 1.10 0]

α = 0.122 (p-value=0.006).

TABLE 6 | Marginal effects for log-linear regression–Final model for PISA with country.

Variables	β	s.e (β)	*p*-value	C.I.-95%
Intercept	–0.76	0.16	0.000	—
g factor	0.57	0.04	0.000	[0.48; 0.65]
SES student	0.02	0.01	0.021	[0.00; 0.05]
Country = Brazil	—	—	—	—
Country = Argentina	–0.07	0.12	0.580	[–0.30; 0.17]
Country = Chile	–0.20	0.15	0.191	[–0.50; 0.10]
Country = Colombia	0.03	0.27	0.903	[–0.50; 0.56]
Country = Peru	0.04	0.13	0.776	[–0.22; 0.30]
SES school = low	—	—	—	—
SES school = middle	0.52	0.12	0.000	[0.28; 0.76]
SES_school = high	0.82	0.13	0.000	[0.57; 1,06]

α = 0.123 (p-value = 0.014).

The correlation matrix with ordinal variables is presented in **Table 3**. The PISA test and g-factor correlated at 0.643 (Pearson coefficient)/0.649 (Spearman coefficient). This values corroborate the values obtained in traditional studies regarding school performance and intelligence. The correlation between the PISA test and SES students was also significant, but lower than *g*-factor.

Considering sample from each country, **Figure 5** shows the scatter plot of PISA test results and *g-factor* score. The correlation was from a minimum 0.409 (p = 0.000) from the Colombian sample, to a maximum 0.786 (p = 0.000) from the Chilean sample.

Figure 7 shows the scatter plot of PISA test results and SES of students. The correlation was from a minimum 0.008 (r non-significant) from the Colombian sample to 0.470 (p = 0.000) from the Argentina sample.

Eta, the coefficient of nonlinear association, indicated a value of 0.100 (weak association), 0.101 (weak association), 0.444 (medium association), and 0.571 (medium association) between PISA test score and age, sex, kind of school, and SES school respectively.

Population-Averaged Model

The complete population-averaged model with all variables of reference (*g*-factor, SES of student, country, kind of school, SES of school, sex, and age) indicated that only *g*-factor, SES of student and SES of the school were important contributor to explain the PISA test score. After the application of the backward algorithm to select significant variables, we arrived at this model presented in **Table 5**.

The final model (**Table 5**) indicated that the *g*-factor influenced the PISA test score (p-value = 0.000). For each additional standard deviation to the mean *g* score, an average increase of 0.57 units [0.48; 0.65] in the mean PISA test score could be expected. In the same direction, SES of students influenced the PISA test score (p-value = 0.020). For each additional unit in SES of students, an average increase of 0.02 units [0.00;.04] in the mean PISA test score could be expected. On the other hand, students enrolled in middle SES outscored students from low SES schools (p-value = 0.000). They had an average value of 0.51 units [0.25; 0.77] higher than students enrolled in low SES schools. Differences were more accentuated with students enrolled in high SES schools. They had an average value of 0.82 units [0.55; 0.1.10] higher than students enrolled in low SES schools (p-value = 0.000). Note that α parameter quantifies the correlation of the PISA test score among students from the same school. In our study the α was 0.122 and significant (p-value = 0.006), i.e., there was homogeneity among students from the same school.

TABLE 7 | Marginal effects for log-linear regression–Final model for *g*.

Variables	β	s.e (β)	*p*-value	C.I.-95%
Intercept	0.11	0.08	0.162	
PISA	0.51	0.03	0.000	[0.45; 0.57]
Kind school = public	—	—	—	—
Kind school = private	0.19	0.10	0.058	[–0.01; 0.38]
Sex = male	—	—	—	—
Sex = female	–0.19	0.03	0.000	[–0.26; –0.13]
Age = 13	—	—	—	—
Age = 14	–0.09	0.04	0.046	[–0.17; 0.00]
Age = 15	–0.18	0.07	0.011	[–0.32; –0.04]
Age = 16	–0.42	0.11	0.000	[–0.64; –0.20]

α = 0.188 (p-value=0.000).

To allow for comparison between countries, the PISA's final model was adjusted with the country variable, with Brazil as a reference, as shown in **Table 6**. No significant difference (p-value > 0.05) between Brazil and the other countries concerning PISA test results were observed.

The same procedure was conducted with *g*-factor, as a dependent variable. The complete population-averaged model with all reference variables (PISA test score, SES of student, country, kind of school, SES of school, sex, and age) indicated that only the PISA test score, sex (female), and age were important contributors to explain the *g*-factor variability. After applying the backward algorithm to select significant variables, we arrived at the model presented in **Table 7**.

The final model (**Table 7**) indicated that the PISA test performance influenced *g* (p-value = 0.000). For each additional standard deviation to the mean of the PISA test, an average increase of 0.51 units [0.45; 0.57] in the mean *g* value could be expected. There was influence of sex on *g* (p-value = 0.000). Females had a lower mean value of *g* [-0.19 units; C.I. 95% = -0.26 to -0.13] than males. Similarly, age influenced *g*. Students at 14, 15, and 16 years old underperformed significantly 13-years-old students.

DISCUSSION

Here we presented results from the SLATINT project based on simultaneous administration of a short version of 2003 PISA and

cognitive measures, to students from five Latin American countries. To the best of our knowledge, there are no other studies that have presented date of this kind. The project was designed to answer the extent to which cognitive ability and social variables influence Latin American students' academic performance. Three important results are discussed:

g-factor and School Performance

The influence of intelligence at the latent-level (or *g*-factor) on student performance was higher (57%; **Table 3**) than the influence of intelligence measured by a single test (35%) (see this last result in Flores-Mendoza et al., 2015). This result was expected as student performance shares a strong common factor, which is indistinguishable from *g*, if compared to the influence of a cognitive ability measured by just one test (Jensen, 1998). In others words, if the PISA test requires several domains (e.g. reading, mathematics, science) it is assumed that the activation of general intelligence is greater than the activation of specific cognitive skills. For instance, one item of the PISA test was related to the internet chat between Mark (from Sydney, Australia) and Hans (from Berlin, Germany). The world time table indicates that Mark and Hans were not allowed to chat from 9:00 am to 4:30 pm due to their respective time zones (school time hours) or between 11:00 pm and 7:00 am (bedtime/late hours). Thus, what time would be a good time for Mark and Hans to have a chat? According to the PISA test developers, this item is representative of Changes and Relationship (math area) and it demands a cognitive skill named as Reflection (see classification of all items in supplemental material A). However, our interpretation is that this item does not require just a cognitive skill, it requires a good understanding of reading, mathematics, and science (time zones between countries), and concomitantly, it requires the joint work of various mental skills (for example, verbal, mathematical, spatial reasoning). Thus and considering the high internal consistency of the PISA test-short version test used in the present study ($\alpha = 0.807$) we inferred that the PISA test required more *g* than a specific cognitive ability, which explain the higher correlation obtained in the present study compared to the previous studies.

The Relationship Between School Performance and G-Factor is Strong, but, as Expected, it is Not Perfect as Aggregate Data Analyses Would Suggest

General (using total sample) and the within-countries correlations indicated values between a minimum of 0.409 and a maximum of 0.786. None correlation in the present study reached the values of aggregated data analysis (above 0.80) presented in studies as those of Rindermann (2018) or Lynn and Becker (2019). Generally, aggregate analyses present a bias leading to inflated estimates above the corresponding values from micro-level data. The origin of this bias is in the error variance and measurement error (Ostroff, 1993). Additionally, aggregate data are assigned equal weight to different sample sizes, which affect the resulting mean effect (Volken, 2007). Therefore, we maintain our assertion made in 2015 (Flores-Mendoza et al., 2015) that school performance is strongly associated to general intelligence (or *g*-factor), but both are not perfectly associated, thus both are not the same construct. Moreover, factors that affected the PISA test score (**Table 6**) were not the same that affected *g* variation (**Table 7**). For instance, SES-school affected the PISA test score, but not *g*; sex (female) did not affect the PISA test, but affected *g*; age did not affect the PISA test, but negatively affected *g*; SES-student slightly affected the PISA test, but not *g*. Therefore, both constructs, despite their strong association, were influenced differently by the variables proposed by the study design. The results seem to indicate that student performance is more sensitive than *g* to the socioeconomic influence, and *g* is sensitive to biological factors, such as sex (see specialized discussion about it in Halpern et al., 2020) and age (distortion age-grade due to individual differences in intelligence, i.e., students at a lower age were more intelligent than older students in the same grade).

g-Factor was Not the Only Source of PISA Variation

The another strong predictor was related to socioeconomic differences between schools, and less to socioeconomic status of students or kind of school. Note in **Figure 3** that even in the public educational system, schools with better socioeconomic status scored better on the PISA test. Thus, low SES-students could benefit from studying in high SES schools. However, what is considered to classify a SES of schools in the present study? In general, several indices can be part of the school composition. For instance, the PISA assessment uses the index termed as ESCS (economic, social, and cultural status), a composition of dimensions which includes: level of education of parents, family wealth, home educational resources, and holding possessions. Psychometric studies have identified limitations of this index for some countries (Rutkowski and Rutkowski, 2013). For instance, while the dimension family wealth fit well in Chile, Argentina, Brazil, Panamá, and Uruguay, it did not fit in Mexico, Colombia, Peru. Home educational resources dimension did not fit well in any Latin American participant country, and the cultural possessions dimension showed reliabilities below of 0.70 (the cutoff criterion for internal reliability). Perhaps, this the reason why the index ESCS has changed somewhat over PISA assessment's cycle. For our study two environmental indices, one related to household possessions and parents' education, and the other related to resources available at the school (inside and outside) were considered. The SES of students implied resources within home (e.g. TV, computer, internet, etc.), and parents' educational. SES school implied conditions out of home. SES school covered school conditions (e.g. class size, presence of computers, etc.), and community resources where the school was located (e.g. waste collection system, drainage system, etc.). The correlation between SES of students and SES school was 0.641 (p-value = 0.000), indicating some independence between both socioeconomic components. The reader can see this certain independence in **Figures 4**, **5**. Students of any socioeconomic status, any kind of school, but enrolled in high SES schools outperformed students enrolled in low SES schools. Moreover,

our multilevel modeling indicated that students enrolled in high SES schools had an average value of 0.51 units higher than students enrolled in low SES schools. In contrast, for each additional unit in students' SES, the increase in mean of the PISA test score was only 0.02 units. Therefore, SES school was a stronger predictor of the PISA test score than SES of students. This result was observed in several independent studies or reported by the PISA assessments (Sirin, 2005; Liu et al.,2014; Perry and McConney, 2010; OECD, 2004; OECD, 2005), some of them including Latin American samples (Duarte et al., 2010). Hence, school and community environments may exert more significant influence than the home environment. The effect of inequalities in the neighborhood or community on school performance has been investigated. Children in poor SES-school, located in a vulnerable neighborhood, tend to experience less social support, fewer school activities, more noise, dangerous and greater physical deterioration environments, which affect educational outcomes (Catsambis and Beveridge, 2001; Evans, 2004; Otero et al., 2017). A meta-analysis based on 88 studies conducted by Nieuwenhuis and Hooimeijer (2016) indicated that among environmental variables the neighborhood poverty, the neighborhood's educational climate, the proportion of ethnic/migrant groups, and social disorganization in the neighborhood affect educational outcomes. We did not include a refined assessment of the community environment; instead, we used a global criterion related to infrastructure and sanitary conditions. Thus, better community assessment is required in future studies.

On the other hand, there is still an open question: why the SES of students and kind of school, traditional predictors, had a weak contribution to student performance? Particular characteristics of some Latino American countries may have contributed to such results. For instance, in Lima city, the capital of Peru, since 2014, private schools' performance decreases to the point that they had the same reading performance and lower performance in mathematics than free public schools in 2016. The reason is that most of the private schools in Lima are low cost (62.5%) and located in poor neighborhoods Ministerio (de Educación, 2018). It is an example that the school type may contribute less than other social variables to the PISA test's performance variation.

We are aware that our results are not new, and they corroborate previous findings in psychology, economy, and sociology (Colom and Flores-Mendoza, 2007). However, as far as we know, our study is one of the first to present the contribution of schools' socioeconomic level to student performance in samples from different cultural contexts. In other words, our study indicated that students benefited from environments/neighborhoods that offered more educational stimuli, good community services, and facilities, despite their cognitive ability, type of school, or socioeconomic level of their families.

Other environmental factors can contribute to student performance, such as educational practices and kind of curricula, as well pointed out by reviewers of the present paper. There is a generalized recognition that high-order thinking skills must be developed in students in the current global knowledge society, however practices school varies widely within and across school systems. Additionally, educational practices vary in uses of time, space, and roles in the interest of more engaging and successful learning. Its effect on the PISA test performance is not clear. For instance, regarding teacher support, the PISA 2018 report (OECD, 2019b) informed that above 80% of students from low perform PISA countries (including Latin American countries) reported that their teachers help with their learning until they understand, while less of 70% of students from high perform PISA countries stated their teacher help them in their learning. Moreover, the OECD reported that, on average across OECD countries, students enrolled in socio-economically disadvantaged schools were more likely than students in advantaged schools to report that they had supportive teachers. Teacher support had positive and moderate relationship with other educational practices (e.g. $r = 0.060$ with Teacher-directed instruction), meaning that any other kind of educational practice could show the same diversity of results showed by teacher support. Additionally, educational practices have to adapt to the levels of cognitive ability and prior knowledge that students bring with them, which would render high complexity to the statistical model proposed by the design of the present study. Also, there were practical reasons (related to the time limit allowed by school principals) that did not allow survey information regarding school teaching characteristics; thus, educational practices' predictive power on the student performance beyond individual differences in intelligence is unknown, and it deserves a special research design.

In general, the implications of our results directly address educational public policies, demonstrating the need to raise the cognitive ability and socioeconomic condition of schools. While it is certain that significantly increasing intelligence within a generation is still an open discussion (Haier, 2014), improving the SES of schools depends exclusively on government decisions. To this regards, our study strongly emphasizes that (high SES) schools can offer resources to low SES-students, in order to achieve improved learning opportunities, and this support is independent of the individual students' abilities.

Note the reader that despite the effort that our samples parallel key variables and characteristics of the Latin American cities under examination (e.g., age, sex, socioeconomic status, kind of schools), our samples were not random samples. Our samples were composed of schools that allowed the study. In other words, our samples were not chosen in a random manner that allows for each variable/member of their original population to have an equal chance of being chosen. Thus, caution is required.

CONCLUSION

This paper presents results of the SLATINT project, a Latin America initiative that verified the human capital present in the region through assessment of student and intelligence performance. This paper was written in a context of coronavirus pandemic. We do not know the impact of the long term absence of schools in 2020 due to pandemic, particularly when considering the psychological development of our children. To this regard, the next PISA survey, scheduled to be conducted in 2022, may very well bring valuable results.

Our results refer to the pre-pandemic social context, and it revealed that general intelligence (or *g*-factor), and SES of

schools, predicted the variation of the PISA test score. SES of students had very small contribution to this prediction. However, the present study faced several limitations regarding the use of non-representative samples from the different countries. Our intention was not to rank countries. We intended to verify the impact of social factors and intelligence on school performance using samples from diverse cultural settings, specifically samples from the Latin American region. Nevertheless, extreme position that favor intelligence and SES school as the only predictors of student performance is not possible due to non-random sampling used in the present study, and the absence of other potential predictors such as quality of school education. On the other hand, we recognize that the cognitive measures used in the present study tended to emphasize spatial reasoning and numerical domains. Thus, a greater variation in the cognitive domains measured would be essential to verify the reliability of our main results. Additionally, the reader may have noticed that our dataset showed some low SES students enrolled in high SES schools, and some high SES students enrolled in low SES schools. The reasons for this SES school-SES student distortion has not been explored. This should be taken into consideration in future studies. Considering all these limitations, our study can be seen as a preliminary investigation about the influence of the schools' resources and the students ability on student performance, and it deserves attention from Latin American educational public policies.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research Ethical Committee of the Federal University of Minas Ferais [N. ETIC 263/07].

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This study was funded by the National Council for Scientific and Technological Development (CNPq/N. 472181/2013-0), and Fundação de Amparo à Pesquisa do Estado e Minas Gerias (FAPEMIG-PPM 03/2014).

ACKNOWLEDGMENTS

The research behind this paper would not have been possible without the Brazilian support from Editora VETOR.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2021.632289/full#supplementary-material.

REFERENCE

Avvisati, F. (2020). The measure of socio-economic status in PISA: a review and some suggested improvements. *Large-scale Assess. Edu.* 8, 8. doi:10.1186/s40536-020-00086-x

Bellei, C. (2008). "The public-private school controversy in Chile," in *School choice international: exploring public-private partnerships*. Editors R. Chakrabarti and P. E. Peterson (chile: MIT Press Scholarship Online).

Catsambis, S., and Beveridge, A. A. (2001). Does neighborhood matter? Family, neighborhood, and school influences on eighth-grade mathematics achievement. *Sociological Focus.* 34 (4), 435–457. doi:10.1080/00380237.2001.10571212

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, F., Mood, A. M., Weinfeld, G. D., et al. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

Colom, R., and Flores-Mendoza, C. (2007). Intelligence predicts scholastic achievement irrespective of SES factors: evidence from Brazil. *Intelligence.* 35, 243–251. doi:10.1016/j.intell.2006.07.008

Costa, H. M., Nicholson, B., Donlan, C., and Van Herwegen, J. (2018). Low performance on mathematical tasks in preschoolers: the importance of domain-general and domain-specific abilities. *J. Intellect. Disabil. Res.* 62, 292–302. doi:10.1111/jir.12465

Coyle, T. R. (2015). Relations among general intelligence (g), aptitude tests, and GPA: linear effects dominate. *Intelligence.* 53, 16–22. doi:10.1016/j.intell.2015.08.005

Cucina, J. M., Peyton, S. T., Su, C., and Byle, K. A. (2016). Role of mental abilities and mental tests in explaining high-school grades. *Intelligence.* 54, 90–104. doi:10.1016/j.intell.2015.11.007

Daniele, V. (2021). Socioeconomic inequality and regional disparities in educational achievement: the role of relative poverty. *Intelligence* 84, 101515. doi:10.1016/j.intell.2020.101515

de Educación, M. (2018). *Tipología y caracterización de las escuelas privadas en el Perú (Estudios Breves N° 3)*. Lima: Oficina de Medición de la Calidad de los Aprendizajes.

Duarte, J., Bos, M. S., and Moreno, M. (2010). Inequity in School Achievement in Latin America: multilevel Analysis of SERCE results according to the socioeconomic status of students. Available at https://publications.iadb.org/es/publicacion/inequity-school-achievement-latin-america-multilevel-analysis-serce-results-according.

Efroymson, M. A. (1960). Multiple regression analysis. *Math. Methods digital Comput.* 1, 191–203. 10.1145/363567.363580.

Evans, G. W. (2004). The environment of childhood poverty. *Am. Psychol.* 59, 77–92. doi:10.1037/0003-066X.59.2.77

Flores-Mendoza, C., Mansur-Alves, M., Ardila, R., Rosas, R., Guerrero-Leiva, K., Maqueo, M. E. G., et al. (2015). Fluid intelligence and school performance and its relationship with social variables in Latin American samples. *Intelligence* 49, 66–83. doi:10.1016/j.intell.2014.12.005

Flores-Mendoza, C., Saraiva, R. B., Câmara, G. C. V., Lopes, W. M. G., Passos, A. P. C. P., Gama, A. M. V. P., et al. (2017). Socioeconomic environment effect on inferential reasoning of Latin American students. *Salud Ment.* 40 (5), 183–189. doi:10.17711/SM.0185-3325.2017.024

Flores-Mendoza, C., Ardila, R., Rosas, R., Lucio, M. E., Gallegos, M., and Reategui-Colareta, N. (2018). *Intelligence measurement and school performance in Latin America. A report of the study of Latin American intelligence project*. Springer International Publishing.

Gamoran, A., and Long, D. A. (2006). *Equality of educational opportunity: a 40-year retrospectiveWCER Working Paper No. 2006–9*. Madison, WI: University of Wisconsin–Madison, Wisconsin Center for Education Research.

Gelade, G. A. (2008). IQ, cultural values, and the technological achievement of nations. *Intelligence* 36, 711–718.

Gottfredson, L. (1997). Why g matters: the complexity of every life. *Intelligence* 24 (1), 79–132.

Haier, R. J. (2014). Increased intelligence is a myth (so far). *Front. Syst. Neurosci.* 8, 34. doi:10.3389/fnsys.2014.00034

Halpern, D., Flores-Mendoza, C., and Rindermann, H. (2020). Sex, gender, and intelligence: does XX = XY for intelligence? in *The cambridge Handbook of the international Psychology of women*. Editors F. Cheung and D. Halpern (Cambridge: Cambridge University Press), 139–152.

Hanscombe, K. B., Trzaskowski, M., Haworth, C. M., Davis, O. S., Dale, P. S., and Plomin, R. (2012). Socioeconomic status (SES) and children's intelligence (IQ): in a UK-representative sample SES moderates the environmental, not genetic, effect on IQ. *PLoS ONE* 7, e30320. doi:10.1371/journal.pone.0030320

Hanushek, E. A., and Woessmann, L. (2007). *Education quality and economic growth*. Washington, DC: The World Bank.

Hanushek, E. A., and Woessmann, L. (2015). *The knowledge capital of nations: education and the economics of growth*. Cambridge, MA: MIT Press.

Hartshorne, J. K., and Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychol. Sci.* 26, 433–443. doi:10.1177/0956797614567339

Heagerty, P. J., and Kurland, B. F. (2001). Misspecified maximum likelihood estimates and generalized linear mixed models. *Biometrika.* 88, 973–985. doi:10.1093/biomet/88.4.973

Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., and Baird, J.-A. (2018). Lessons learned from PISA: a systematic review of peer-reviewed articles on the Programme for International Student Assessment. *Scand. J. Educ. Res.* 62, 333–353. doi:10.1080/00313831.2016.1258726

Hunt, E. (2010). *Human intelligence*. Cambridge: Cambridge University Press.

Jensen, A. (1998). *The g factor: the science of mental ability*. Westport, CT: Praeger.

Juan-Jose, N., Garcia-Rubio, J., and Olivares, P. R. (2015). The relative age effect and its influence on academic performance. *PlosOne.* 10 (10), e0141895. doi:10.1371/journal.pone.0141895

Kuncel, N. R., Hezlett, S. A., and Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: can one construct predict them all? *J. Pers Soc. Psychol.* 86 (1), 148–161. doi:10.1037/0022-3514.86.1.148

Leeson, P., Ciarrochi, J., and Heaven, P. C. L. (2008). Cognitive ability, personality, and academic performance in adolescence. *Personal. Individual Differences.* 45 (7), 630–635. doi:10.1016/j.paid.2008.07.006

Lenehan, M. E., Summers, M. J., Saunders, N. L., Summers, J. J., and Vickers, J. C. (2015). Relationship between education and age-related cognitive decline: a review of recent research. *Psychogeriatrics.* 15, 154–162. doi:10.1111/psyg.12083

Liang, K., and Zeger, S. (1986). Longitudinal data analysis using generalized linear models. *Biometrika.* 73 (1), 13–22. doi:10.2307/233626

Linacre, J. M. (2007). *FACETS Rasch measurement computer program* Chicago: Winsteps.com.

Liu, H., Damme, J. V., Gielen, S., and Noortgate, W. V. D. (2014). School processes mediate school compositional effects: model specification and estimation. *Br. Educ. Res. J.* 41, 423–447. doi:10.1002/berj.3147

Lynn, R., and Becker, D. (2019). *The intelligence of nations*. London: Ulster Institute for Social Research.

Lynn, R., and Vanhanen, T. (2002). *IQ and the wealth of nations*. Westport, CT: Praeger.

Lynn, R., and Vanhanen, T. (2012). *Intelligence: a unifying construct for the social sciences*. London: Ulster Institute for Social Research.

Maranto, R., and Wai, J. (2020). Why intelligence is missing from American education policy and practice, and what can Be done about it. *J. Intelligence.* 8 (1), 2. doi:10.3390/jintelligence8010002

Neisser, U., Boodoo, G., Bouchard, T., Boykin, A., Brody, N., Ceci, S., et al. (1996). Intelligence: knowns and unknowns. *Am. Psychol.* 51 (2), 77–101.

Nieuwenhuis, J., and Hooimeijer, P. (2016). The association between neighbourhoods and educational achievement, a systematic review and meta-analysis. *J. Hous Built Environ.* 31, 321–347. doi:10.1007/s10901-015-9460-7

O'Connell, M., and Marks, G. N. (2021). Are the effects of intelligence on student achievement and well-being largely functions of family income and social class? Evidence from a longitudinal study of Irish adolescents. *Intelligence* 84, 101511. doi:10.1016/j.intell.2020.101511

OECD (2001). *Knowledge and skills for life. First results from the OECD Programme for international student assessment (PISA) 2000*. Paris: OECD.

OECD (2004). *Learning for tomorrow's world: first results from PISA 2003*. Paris: OECD.

OECD (2005). *School factors related to quality and equity: results from PISA 2000*. Paris: OECD.

OECD (2007). PISA*2006 Science competencies for tomorrow's world, Volume 1*. Paris: Analysis.

OECD (2010). *PISA 2009 results: learning trends: changes in student performance since 2000*. Paris: OECD. doi:10.1787/19963777

OECD (2013). *PISA 2012 Results in Focus. What 15-years-olds know and what they can do with they know?* Paris: OECD.

OECD (2016). *Skills in ibero-America. Insights from PISA 2012*. Paris: OECD.

OECD (2018). *PISA 2015 results in focus*. Paris: OECD.

OECD (2019a). *PISA 2018 results (volume I): what students Know and can do*. Paris: OECD. doi:10.1787/35665b60-en

OECD (2019b). *PISA 2018 results (Volume III): What School Life Means for Students' Lives*. Paris: OECD. doi:10.1787/acd78851-en

Ostroff, C. (1993). Comparing correlations based on individual-level and aggregated data. *J. Appl. Psychol.* 78 (4), 569–582. doi:10.1037/0021-9010.78.4.569

Otero, G., Carranza, R., and Contreras, D. (2017). 'Neighbourhood effects' on children's educational achievement in Chile: the effects of inequality and polarization. *Environ. Plann. A: Economy Space* 49 (11), 2595–2618. doi:10.1177/0308518X17731780

Perry, L., and McConney, A. (2010). School socio-economic composition and student outcomes in Australia: implications for educational policy. *Aust. J. Edu.* 54 (1), 72–85. doi:10.1177/000494411005400106

Prentice, R. L. (1988). Correlated binary regression with covariates specific to each binary observation. *Biometrics* 44 (4), 1033–1048.

Raven, J., Raven, J. C., and Court, J. H. (2000). *The standard progressive Matrices, 3*. Oxford, EnglandSan Antonio, TX: Oxford Psychologists PressThe Psychological Corporation.

Rindermann, H., Sailer, M., and Thompson, J. (2009). The impact of smart fractions, cognitive ability of politicians and average competence of peoples on social development. *Talent Develop. Excell.* 1, 3–25. doi:10.1177/0956797611407207

Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: the homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *Eur. J. Personal.* 21, 667–706. 10.1002/per.634.

Rindermann, H. (2018). *Cognitive capitalism: human capital and the wellbeing of nations*. Cambridge: Cambridge University Press.

Rosas, R. (1996). Replicación del Modelo de Estructura de Inteligencia de Berlín en una Muestra de Estudiantes Chilenos. *PSYKHE* 5 (1), 39–56.

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., and Spinath, F. M. (2015). Intelligence and school grades: a meta-analysis. *Intelligence* 53, 118–137. doi:10.1016/j.intell.2015.09.002

Rutkowski, D., and Rutkowski, L. (2013). Measuring Socioeconomic Background in PISA:one size might not fit all. *Res. Comp. Int. Edu.*, 8, 259–278. doi:10.2304/rcie.2013.8.3.259

Schleicher, A. (2019). *PISA 2018. Insights and interpretations*. Paris: OECD.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: a meta-analytic review of research. *Rev. Educ. Res.* 75, 417–453. 10.3102/003465430.

Strenze, T. (2007). Intelligence and socioeconomic success: a meta-analytic review of longitudinal research. *Intelligence* 35, 401–426. doi:10.1016/j.intell.2006.09.004

Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton, Mifflin and Company.

Trucco, D. (2014). *Educación y desigualdad en América Latina*. Santiago, Chile: Naciones Unidas, CEPAL .

Vior, S. E., and Rodríguez, L. R. (2012). La privatización de la educación Argentina: un largo proceso de expansión y naturalización. *Pro-Posições.* 23 (2), 91–104.
Volken, T. (2007). Methodological and conceptual notes on national cognitive ability. *Eur. J. Personal.* 21 (5), 755–757.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 13 May 2021
doi: 10.3389/feduc.2021.658973

Teacher and Student Practices Associated with Performance in the PISA Reading Literacy Evaluation

Luis Rojas-Torres[1]*, *Graciela Ordóñez*[1] and *Karen Calvo*[1,2]

[1]Institute of Psychological Research, University of Costa Rica, San José, Costa Rica, [2]Faculty of Literature, National University of Costa Rica, San José, Costa Rica

This article aims at finding teacher's and student's practices that relate to performance in PISA reading literacy evaluations and that are feasible to intervene in order to assist the improvement of reading competency. To achieve this purpose, the study was developed with data collected from the population of Costa Rica that took the PISA evaluation in 2018 (n = 4691, 2340 men, and 2351 women). A linear regression of the reading score was performed utilizing plausible values and sampling weights. The predictors of the regression were contextual factors, teacher practices, and student habits. Time spent and interest in reading showed a positive and relevant association with student's performance in reading, controlling important background aspects like economic resources and parents' education. Moreover, 28.19% to the obtained variance explanation of the reading literacy (27%) was only due to the teacher's and student's practices. These results provide favorable information to design interventions for the improvement of reading competency.

Keywords: reading literacy, sampling weights, teacher practices, student habits, time reading, plausible values

OPEN ACCESS

Edited by:
Marcos Cupani, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

Reviewed by:
Valeria Moran, Siglo 21 Business University, Argentina
Arturo Orozco-Vargas, Universidad Autónoma del Estado de México, Mexico

***Correspondence:**
Luis Rojas-Torres
luismiguel.rojas@ucr.ac.cr

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

Received: *26 January 2021*
Accepted: *28 April 2021*
Published: *13 May 2021*

Citation:
Rojas-Torres L, Ordóñez G and Calvo K (2021) Teacher and Student Practices Associated with Performance in the PISA Reading Literacy Evaluation. Front. Educ. 6:658973. doi: 10.3389/feduc.2021.658973

INTRODUCTION

In 2018, ten Latin American countries participated in the Program for International Student Assessment (PISA) tests: Argentina, Brazil, Chile, Costa Rica, Dominican Republic, Mexico, Panama, Peru, and Uruguay. In all countries of the region, except for Chile, it was observed that more than 40% of the student body presented a level 1 in reading literacy; while the countries of the Organization for Economic Cooperation and Development (OECD), on average, reached 22% of students at this level (OECD, 2019b). According to this result, almost half of 15-year-old students in Latin America have very low levels of reading literacy, that is, they barely understand the explicit or literal information in short texts.

Based on the above situation, we concluded the need to implement strategies to improve reading literacy performance in the region, as proposed in the present study which seeks to determine which teacher's and student's practices can have an impact on the improvement of reading literacy. To meet this objective, we first presented a set of variables that may be associated with reading literacy according to the literature and then we analyzed the effect of these variables on PISA 2018 reading literacy scores in a country in the region: Costa Rica.

According to PISA's own definition, reading literacy is one of the many communicative components that involves "the capacity to understand, use and reflect on written texts in order to achieve goals, develop knowledge and potential, and participate in society" (OECD, 2019a: 34). This conceptualization requires not only the act per se of reading, but also the replication that this

reading may have in the person who exercises it from the generation of new meaning. It therefore includes cognitive skills beyond decoding, such as, mastery of grammatical and linguistic structures, as well as contextual knowledge. The tasks considered by PISA for the evaluation of reading comprehension also assess the applicability of the reading exercise, which translates into the establishment of purposes that serve as a stimulus for reading itself or with the development of writing.

The notion of reading literacy has been differentiated from that of reading competence (Jiménez, 2013), from the pragmatic position by emphasizing a particular use that motivates the act of reading (Solé 2012) or applies it when seeking to solve a particular problem or situation with reading (Diez and Egío, 2017). Hence, the purpose is fundamental for reading comprehension and meaning-making to achieve the expected results both in the assessment and even outside of it.

Academic achievement or student's performance, understood as the result in the measurement of a skill or knowledge that in turn implies the performance on an assessment (Edel, 2003; Lamas, 2015) is subjected to and influenced by several variables that could have a negative or positive impact on that student's performance (Bormuth, 1973; Wilkinson, 1998; Artelt, Schiefele, and Schneider, 2001; Shiel and Cosgrove, 2002; Rasmussen, 2003; Connor, Son, Hitman and Morrison, 2005; Morrison, Bachman, and Connor, 2005; Brozo et al., 2014). These variables could include the practices that teachers have selected to teach their subject, (OECD, 2005; Guo et al., 2012; Meroni et al., 2015; Duke, Cerveti and Wise, 2016), student's habits (Bormuth, 1973; Wilkinson, 1998; Artelt, Schiefele, and Schneider, 2001; Shiel and Cosgrove, 2002; Rasmussen, 2003; OECD, 2010; Brozo et al., 2014; Brenes, 2019), and the factors that come from family context, such as the level of education of both the father and the mother, and their socioeconomic status (Montero et al., 2012; Hernández-Padilla and Bazán-Ramírez, 2016; Alves et al., 2017; García et al., 2018; Brenes, 2019).

Research concerning teacher's practices finds that teachers with positive attitudes toward their subject matter or student's learning are associated with students with high academic achievement (OECD, 2005; Guo et al., 2012; Meroni et al., 2015; Duke et al., 2016). Thus, one of the teacher's practices associated with student's performance, particularly in reading, is that the language instructor expresses interest in their subject (Wray and Medwell, 2000; Guo et al., 2012). Furthermore, Wray and Medwell (2000) indicating that when investigating instructors' teachings in reading literacy, they demonstrated their own appreciation for writing and reading, which generated better results in their students' learning, and therefore, in their academic performance.

Another relevant teacher's practice is for the language teacher to take an interest in the students' learning, providing them with more opportunities for understanding the subject matter. Thus, Vidal-Moscoso and Manriquez-López (2016) stated in their study that teachers should assume the commitment to teach reading adequately to guide and support in the formation of reading literacy, due to the fact that there is a significant influence on student's performance when teachers express interest in their students' learning.

On the other hand, relevant teacher practices include the classroom environment where disciplinary control is crucial for teaching, learning and, therefore, for student academic performance. From this perspective in a number of studies (Omoteso and Semudara, 2011; Akiri, 2013; Duke et al., 2016), it is stated that the environment generated by the instructors during class influences significantly the students' academic achievement in public high schools. In addition, a teacher's ability to effectively manage a classroom also depends on the mode of training and the work experience they have; generally, more experienced teachers tend to have better disciplinary control in their classrooms (Omoteso and Semudara, 2011).

Regarding the variables associated to the students' habits, studies show that absenteeism, daily dedicated time to read for pleasure (reading time), and the interest the student has for reading are factors associated with students' academic performance. From this perspective, in a number of studies (Romer, 1993; Chen and Lin, 2008; Schmulian and Coetzee, 2011; López-Bonilla and López-Bonilla, 2013; Teixeira, 2016) it is affirmed that there is a negative association between students' absenteeism and academic performance. Moreover, researches have hypothesized that class attendance should be positively correlated with academic achievement, consequently being beneficial in the development of reading literacy. For example, Schmulian and Coetzee (2011), utilizing simple correlation techniques and a sampling characterized by low levels of class absenteeism (less than 10%), observed that there is a positive and significant correlation between class attendance and academic performance. Nonetheless, Schmulian and Coetzee (2011) affirmed that such correlation is low.

On the other hand, López-Bonilla and López-Bonilla (2013) determined that absenteeism is a complex and multifactorial phenomenon. Their study showed that efficiency, teaching style, academic interest, content, teaching format, peer influence, and peer fears are determinant on absenteeism; however, absenteeism had an association with student's performance. In another line of research, Teixeira (2016) substantiates that class absenteeism weakens student's academic performance, substantially influenced by contextual factors such as attendance rules, perceived difficulty of the class, teacher's characteristics and access to online reading material. In addition, Teixeira (2016) affirmed that it is true that factors of individual difference such as motivation, conscience, and intelligence increase the probability of a student attending class.

Some other studies indicate that there is an impact on student's academic performance when subjects are absent from class. To illustrate, Romer (1993), Devadoss and Foltz (1996), Chen and Lin (2008) found a positive and significant relation between class attendance and the grades obtained by students in their tests. Thus, Chen and Lin (2008) in their study determined that 114 students who attended classes for an entire semester had a better grade and a positive impact on test scores. According to Chen and Lin (2008), the effect of attending class correlated with an improvement between the 9.4 and 18% in performance on the exams on those students who chose to attend every class. Meanwhile, Devadoss and Foltz (1996) found that a student

who attended every class had chances of achieving, on average, a grade 0.45 points higher than a student who only attended half of the classes, this on a sampling of 400 students.

In relation to the variable interest in reading, different studies (OECD, 2011; Schiefele et al., 2012; Brozo et al., 2014; Dezcallar et al., 2014) indicate that those students who read for pleasure have a better performance in the PISA evaluation, and they are more efficient in their learning, as well as in their scholarly achievements in general. For example, Shiel and Cosgrove (2002) conducted a study about the association between reading literacy and the variables: positive attitude towards reading, frequency of reading during free time, and the socioeconomic status of the student. These researchers determined that the students with higher achievement were those who kept a positive attitude toward reading, participated in a moderate amount of reading during leisure time, and read a vast range of texts. Similarly, in the studies of Montero et al. (2012), Dezcallar et al. (2014), Valdés (2013), and Castro (2014) it is stated that interest in reading contributes to student achievement. For example, Castro (2014) points out that students who show interest in reading, and who read to satisfy their curiosity and enjoyment, obtained 439 points in their performance level in the PISA 2009 reading literacy test, while those students who indicated disinterest in reading had a performance of 413 points, generating a difference of 26 points between both groups.

In this regard, it should be noted that the pleasure of reading is also associated with academic performance in general and not only with reading comprehension. In the study by Dezcallar et al. (2014), it is indicated that reading not only implies access to information, but also a mechanism for critical thinking and socialization. There is a relationship between knowing how to read and enjoying reading that positively favors the perception of learning. Today, however, reading rivals other forms of entertainment such as video games, television and digital media, the use of which diminishes the time spent on academic homework.

Despite the advantages of reading for pleasure, according to the OECD, the percentages of students who read daily for pleasure declined in most OECD countries between 2000 and 2009, and it is women and students with higher socioeconomic status who are more likely to read for pleasure.

Regarding time spent reading, Shiel and Cosgrove (2002) found that the frequency of reading during leisure time was one of the factors explaining reading literacy performance. Similarly, Yubero and Larrañaga (2015) through a logistic regression analysis, in a sample of 2,745 students, found that part of the time devoted to leisure reading is defined by including reading in the student's lifestyle which allows a greater appropriation of vocabularies and an improvement in reading comprehension. Likewise, Guerra and Guevara's (2017) study reports that students who spend little time reading obtain lower scores in reading comprehension tests while those who invest more time obtain higher scores; however, Guerra and Guevara (2017) argue that the high scores were largely due to the use of metacognitive strategies and motivation towards reading. However, these variables: metacognitive strategies, motivation towards reading, and time spent reading, presented a high positive association. Moreover, Valdés (2013) in his correlational study found that reading is an activity seldom performed in the spare time of pre-adolescents and adolescents whose positive disposition towards reading decreases as age increases; this is because said population reads due to the demands imposed by the school which prevents the development of skills that forge the competent reader.

Studies concerning contextual factors and their relationship with academic performance have shown that parental education, dependence on the school, and resources available for their performance in the school environment are determinants of student's performance (Hernández-Padilla and Bazán-Ramírez, 2016; García et al., 2018; Brenes, 2019). This way, research indicates students have more probabilities of staying in school and performing better if they have the support of their families, both in affective and economic terms (availability of economic resources). It has also been observed that students' academic achievement is lower if the household has a precarious socioeconomic state, due to youngsters having to find a job to support their households (Trejos, 2010; Montero et al., 2012; Brenes, 2019).

From another perspective, in several studies it is stated that the level of education of legal guardians, fathers, and mothers, as well as their socioeconomic status, represent a positive relation with student's performance (Trejos, 2010; Montero et al., 2012; Brenes, 2019). To illustrate, in the study conducted by Alves et al. (2017) on the latent variable family in which the association between the fathers' and mothers' levels of education, socioeconomic level, students' cognitive performance, and academic achievement were considered; it was found that the family variable has an important association with the students' cognitive and academic performance, the fathers and mothers with higher academic and socioeconomic levels are associated with higher academic achievement of their children.

This section has showed that teacher practices, student's habits and contextual factors are related to the reading literacy. These variables were selecting because they are showed a relevant relationship with reading literacy in several studies. The importance of this study lies in recognizing which of these variables are more related to the reading literacy in order to give information to researchers to create new strategies to improve this competency. The hypothesis is that the contextual variables will be the more relevant predictors, besides that the rest of variables will have relevant associations too.

MATERIALS AND METHODS

Participants

The participants of this study are a sampling of 15-year-old students enrolled in a Costa Rican educational institution at grade 7 or higher (Schleicher, 2019). The sampling was of a probabilistic type in two stages, whose objective was to obtain a representative sample. In the first one, the institutions of

education of 5 established strata (technical rural, technical urban, academic rural, academic urban, and private) were randomly selected. In the second stage, a number of students from these institutions were randomly selected too. The sampling collected by PISA was of 7,119 subjects.

The sample used in this study has a total of 4,691 people. We only considered people that completed every instrument used in this study (2,351 women and 2,340 men; 417, 701, 718, 2,222, and 633 from the strata technical rural, technical urban, academic rural, academic urban, and private, respectively).

Measures

Teacher Practices

In this study, three practices of Spanish teachers (the language of the PISA reading test considered) were taken into account who are also in charge of promoting reading comprehension in the secondary education setting in Costa Rica: interest in the subject matter, interest in student's learning (interest in learning) and disciplinary control of the group (disciplinary control).

The variables were measured from the students' perception regarding the practices of their Spanish teacher. For the measurement, four-point Likert scales were used. Every scale consisted of four items, except for the one regarding disciplinary control of the class which consisted of five. The Cronbach's alfas of the scales were 0.88, 0.86, and 0.79 for interest in the subject, interest in students' learning, and disciplinary control.

The measures of each variable of teacher practices were the average score of the items; therefore, the potential ranges of variation were of 1–4. In all three variables, the value 4 indicated the highest positive perception of the students regarding the practices of the Spanish teacher.

Student Habits

The three student's habits analyzed were the amount of lessons from which the student was absent in the last two weeks (absenteeism), daily dedicated time to reading for pleasure (reading time), and interest in reading. The data of the first two variables was collected via direct questions. For the first variable, 4 answer options were available (0 lessons, 1 or 2 lessons, 3 or 4 lessons, 5 or more lessons); for the second variable there were 5 answer options (I do not read for pleasure, 30 min or less per day, more than 30 min but less than 60 min per day, from 1 to 2 h per day, more than 2 h per day). The third variable, interest in reading, was collected through a Likert scale with 5 four-point items. The scale had 5 items about behaviors related to reading (e.g., I like to talk about books).

For the variables 'Absenteeism' and 'Reading time' an ordinal score was considered to which a series of consecutive numbers were assigned to the categories of the variables. In the first variable, a score from 0 to 3 was created (0 = no absenteeism; 3 = high absenteeism) and in the second, a score from 0 to 4 (0 = no reading time; 4 = high reading time). This means that the values of these variables represent levels. Lastly, the unit of measure of interest in reading was the average score of the scale, therefore, its potential range of variation was of 1 to 4.

Contextual Factors

The contextual factors considered were the mother's level of education, the father's level of education, and the household resources. As with the student's variables, in the levels of education of the mother and father ordinal measures were created. These measures varied from 0 to 4 (0 = incomplete primary education; 1 = complete primary education; 2 = complete middle school; 3 = complete secondary education, 4 = university degree).

To create a summary measure of the household resources, the reports of the amount of televisions, automobiles, computers, cellphones with internet connection, tablets, bedrooms with private restrooms, or electronic book readers, and musical instruments were considered. Thus, with this data a principal component analysis was developed to obtain the linear combination that captures the higher percentage of variance of these variables. This linear combination is the measure of household resources utilized in this study and this measure explained a 38% of the variables' variance. This index is presented in standardized units.

Reading Literacy

The reading evaluation conducted by PISA is computer based. The test was divided in three sets of questions (core, stage 1, and stage 2). Each set included a reading describing a real-life situation, as well as a number of multiple-choice questions or short answer questions. In addition, the test was conducted with an adaptive approach, meaning the sets of questions were assigned utilizing the information obtained in previous sets. The questions of the reading test were organized within a 1-h period.

For the reading literacy construct a single measure is not generated, instead, ten indicators called 'plausible values' are utilized. These values are obtained in the following way: a) A distribution of ability in reading literacy for each individual evaluated in PISA is generated based on the answers submitted in the test and other collected measures. Then, b) random values of this distribution are generated which are the plausible values. The use of plausible values is due to the total scores reflecting grades in a specific set of items, instead of the entire potential universe of items (OECD, 2019a; OECD, 2019b).

Procedure

The collection of data was executed by the OECD, as it is known. The selected students in the sampling completed the cognitive evaluation in their high schools' computer laboratories (the cognitive evaluation was composed of the reading evaluation already mentioned, a mathematics questionnaire, as well as a science questionnaire, the latter two are organized in a 1-h period). Then, the students completed a background questionnaire, which contained the information used in the creation of the variables considered in this study. This questionnaire was completed in 35 min, approximately, and was taken in a computer too (OECD, 2019a; OECD, 2019b).

The authorization to collect the students' information was provided to OECD by Education Department of Costa Rica.

Students were not obligated to complete the instruments. After the data recollection, all the variables related to the students' identification were dropped.

Data Analysis

First, a descriptive analysis of the variables was performed with the purpose of analyzing the distribution of the variables of interest among the population. All the statistics of the predictor variables were calculated using a weighted estimation based on the sampling weights of the observations which allowed the creation of unbiased estimates of the intended parameters. The weights used in this study were those estimated by PISA which sought to control for differences in the probabilities of selection of examinees, representation of strata, and school participation rates (OECD, 2009). The sampling weights used for this estimation were called final weights.

To calculate the sample variance of the estimated statistic $(\widehat{\theta})$ other weights which were plausible in the population were considered and called replications. The replications were generated by the OECD by means of a Balanced Repeated Replication with Fay's modification with a factor of $k = 0.5$ and a total of 80 weight vectors were generated. Then, the statistic of interest was calculated with each of the weights given by the replicates, referred to as replicate statistics $(\widehat{\theta}_r)$. The formula for the sample variance of the statistic of interest, with the replicate construction method used, is

$$\sigma_s^2\left(\widehat{\theta}\right) = \frac{1}{R(1-k)^2}\sum_{r=1}^{R}\left(\widehat{\theta} - \widehat{\theta}_r\right)^2,\ R = \text{number of replications}$$

The square root from the formula above represents the estimated statistic sampling error and can be used for the calculation of the statistic *t*.

In the case of descriptive statistics of reading literacy, parameters were estimated based on plausible values and sampling weights. For this, with each of the 10 vectors of plausible values the statistic of interest was calculated and its respective sample variance, by means of the formulas indicated in the previous paragraph. These statistics were called plausible value statistics $(\widehat{\theta}_{pv})$. The final statistic $(\widehat{\theta})$ was the average of the plausible value statistics. The error variance of the statistic's estimate is a weighted sum of the sample variance and the imputation variance. The first is the average of the estimated sampling variances at each plausible value and the second is the variance of the plausible value statistics. The formula of the error variance is the following (OECD, 2009):

$$\sigma_e^2\left(\widehat{\theta}\right) = \frac{1}{M}\sum_{pv=1}^{M}\sigma_s^2\left(\widehat{\theta}_{pv}\right) + \left(1+\frac{1}{M}\right)\frac{1}{M-1}\sum_{pv=1}^{M}\left(\widehat{\theta} - \widehat{\theta}_{pv}\right)^2;$$
$$M = \text{number of } pv$$

Secondly, to determine if one of the variables of interest was associated with reading literacy performance, the correlation coefficient was estimated based on plausible values and sampling weights. The estimate was similar to that presented with the descriptive statistics. For a particular correlation of a variable of interest with reading literacy, the correlation with each plausible value was calculated, considering the final weights. The final correlation coefficient was the average of the coefficients obtained. On the other hand, for the calculation of the standard errors, the parameter estimates based on the weights of the replicas and the formula of the standard error mentioned previously were considered. A correlation was considered relevant if its value exceeded the threshold of 0.20 in absolute value. The analysis of the results was based on the relevance of the coefficients, rather than on the *p*-values because most of the coefficients were significantly different from 0; the latter is to be expected due to the size of the sampling used (Lin et al., 2013).

Lastly, to determine which variables were relevant in explaining the variance of reading literacy, a regression analysis was carried out based on the plausible values and sampling weights. The independent variables of this analysis were those described in teacher practices, student's habits and contextual factors. As with the correlation, a multiple regression was performed on each plausible value considering the final weights. The regression coefficient of a particular variable was the average of the regression coefficients obtained in these regressions. A regression coefficient was considered relevant if its standardized value exceeded the threshold of 0.20 in absolute value (Acock, 2014).

The estimation of all the models was done with the software R, in version 3.6.3.

RESULTS

Descriptive Statistics

Table 1 presents the descriptive statistics of the variables considered in the study. In the teacher's practices, it was obtained that the median interest in the subject (me = 3.00, se < 0.01) and the disciplinary control (me = 3.02, se = 0.06) were higher than the central point of the scale (2.5); in contrast, the median interest in students' learning was located in the center of the scale (me = 2.50, se = 0.33). In the student's habits, it was observed that the median absenteeism was 0 (se < 0.01) which indicated that at least 50% of students marked level 0 of absenteeism (that is, they were not absent to any class in the two weeks prior to taking the questionnaire); the median interest in reading (me = 2.40, se< 0.01) was lower than the central point of the scale, and the median reading time was 1.00 (se < 0.01) which indicated that at least 50% of the students barely reached the lowest reading level: less than half an hour of reading per day. In the contextual factors, the medians in father's education and mother's education were equal to 2 (se < 0.01 in both). In the case of the first variable, this result indicated that at least 50% of the students have a parent with an education level less than or equal to 2 (completed middle school); then, in the resource index a positive asymmetry was observed (me = −0.13< mean = 0.01) which suggests that there are fewer people in the higher levels of the index than those in the lower levels.

The predictor variables defined three groups of variables correlated with each other. As expected, the groups were

TABLE 1 | Descriptive statistics and correlations of the variables of the study.

	Mean	Med	Sd	1	2	3	4	5	6	7	8	9
1. Resources index	0.01	−0.13	1.03	1.00								
	(0.03)	(0.07)	(0.03)	(0.00)								
2. Mother's education	2.25	2.00	1.37	**0.47**	1.00							
	(0.04)	(0.00)	(0.01)	(0.02)	(0.00)							
3. Father's education	2.15	2.00	1.39	**0.48**	**0.49**	1.00						
	(0.04)	(0.00)	(0.01)	(0.02)	(0.02)	(0.00)						
4. Teacher's interest in the subject	3.01	3.00	0.74	0.01	0.00	0.01	1.00					
	(0.01)	(0.00)	(0.01)	(0.02)	(0.02)	(0.02)	(0.00)					
5. Teacher's interest in learning	2.66	2.50	0.81	0.07	0.04	0.05	**0.51**	1.00				
	(0.02)	(0.33)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.00)				
6. Teacher's group control	3.02	3.20	0.66	−0.01	0.00	−0.01	**0.24**	**0.21**	1.00			
	(0.01)	(0.06)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)			
7. Student absenteeism	0.52	0.00	0.74	0.00	0.00	0.00	−0.07	−0.05	−0.11	1.00		
	(0.01)	(0.00)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.00)		
8. Student's interest in reading	2.52	2.40	0.74	0.01	0.03	0.05	0.05	0.09	0.05	−0.05	1.00	
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.00)	
9. Student's reading time*		1.00		0.02	0.02	0.03	0.07	0.10	0.03	−0.02	**0.67**	1.00
		(0.00)		(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.00)
10. PISA reading Literacy	425.09	423.36	8.58	**0.39**	**0.31**	**0.31**	0.13	0.12	0.12	−0.10	**0.22**	0.16
	(2.73)	(2.99)	(1.48)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)

*The values in parentheses are the sampling errors of the estimated statistics. All coefficients were significant at 5% with the exception of the mean and median of Resources Index. Values in bold are the correlations greater than 0.20. *Ordinal variable: for that reason, we don't include its mean and standard deviation.*

TABLE 2 | Linear regression of the PISA reading literacy scores, using plausible values and sampling weights.

Variable	Coef	Std. coef	Samp. e	Imp. e	Se	*t*
Constant	296.75	0.00	8.10	3.79	9.02	32.90
Resources index	21.29	**0.27**	1.70	0.63	1.82	11.71
Mother's education	6.44	**0.11**	1.09	0.42	1.17	5.48
Father's education	6.97	**0.12**	1.01	0.30	1.06	6.60
Teacher's interest in the subject	7.12	0.07	1.93	0.69	2.07	3.45
Teacher's interest in stud. Learning	2.90	0.03	1.81	0.58	1.91	1.52
Teacher's group control	11.98	**0.10**	1.90	0.67	2.03	5.90
Student absenteeism	−6.80	−0.06	1.78	0.46	1.84	−3.69
Student's interest in reading	17.27	**0.17**	2.05	0.65	2.16	7.99
Student's reading time	2.98	0.05	1.16	0.31	1.20	2.48
R^2	0.27		0.02	0.01	0.02	13.11

Coef, coefficient; std. coef, standardized coefficient; samp. e, sampling error; imp. e, imputation error; se, standard error; t, t *value,* R^2*, determination coefficient. All coefficients were significant at 5%. Values in bold are the standardized coefficients greater than 0.10.*

determined by the divisions considered in the selection of the variables. In the group of contextual factors, the three variables had correlations between them of 0.47–0.49; in the group of teacher practices, the correlations were between 0.21 and 0.51; while in the group of student habits, the variable absenteeism was not associated significantly with the other two variables of the group, but these two variables, reading time and interest in reading, presented a correlation of 0.67.

As for the correlations with the variable reading literacy, it was obtained that the relevant variables in its prediction were the three contextual factors (r = 0.39, 0.31 and 0.31 for resource index, mother's education and father's education, respectively) and the interest in reading (r = 0.22, se = 0.01). The other two student's habits presented statistically significant correlations, but low in absolute value (absenteeism: r = −0.10; reading time: r = 0.16). Similarly, the teacher practices variables presented statistically significant correlations, but low, between 0.12 and 0.13.

Variables Relevance in the Reading Scores Explanation

Table 2 presents the results of the linear regression predicting reading literacy. It was obtained that all the variables included were statistically significant except for the teacher's variable: interest in learning. The set of variables used explained a 26.50% of the variance belonging to the reading literacy scores (se = 2.02%). The regression coefficient b of each variable indicated that an increase of one unit in its score, keeping the other variables constant, is associated with an average increase of b units in the reading score, for instance, the increase of one unit in student's interest in reading was associated with an average increase of 17.27 units in reading literacy.

As for the contextual variables, it was concluded that the three indicators considered are positively associated with reading literacy. Based on the criterion of the standardized coefficient (β), it is concluded that the variable of the resource index is the

variable considered with the greatest relative importance in the prediction proposed (β = 0.27). The other two contextual variables had a weak relevance in the prediction (mother's education: β = 0.11, father's education: β = 0.12). As for the teacher's practices, the interest in the learning, the interest in the subject and the disciplinary control were positively associated with the reading literacy scores, all of them presented weak coefficients (β = 0.03, 0.07 y 0.10, respectively).

In the student's habits, it was observed that reading time and interest in reading were positively associated with the variable studied while absenteeism was negatively associated. It is important to point out that within this group of variables the second most relevant factor in the prediction of reading literacy scores is found: interest in reading (β = 0.17); the other two variables presented low relative importance (absenteeism: β = −0.06; reading time: β = 0.05).

It is important to mention that the variables interest in reading and reading time, despite having a high correlation, did not show signs of collinearity in the model (reading time: vif = 1.86, se = 0.44; interest in reading: vif = 1.85, se = 0.46). However, two additional models were estimated considering only one of these variables within the group of predictors, to further analyze their associations. If only the variable interest in reading is considered, the coefficient of determination of the model is practically equal to that of the complete model (R^2 = 26.46, se = 2.01), but the standardized coefficient of the variable increases to 20.35 (se = 0.01). In the model that considers only the reading time, the determination coefficient decreases to 24.83 (se = 2.06) while the standardized coefficient of the variable increases considerably to 16.16 (se= 0.14). Therefore, both variables are almost relevant in the prediction of reading literacy (the coefficients were close to the threshold 0.20), but the prediction made by the interest in reading in the complete model absorbs part of the explanation that time for reading can offer.

Means of the two most relevant variables in the model were analyzed by strata using an ANOVA without replications. It was no found significant differences in the interest in reading variable (F(4,6777) = 1.03, p = 0.39) and all the means were close to 2.5. On the other hand, the resources index showed relevant differences between strata (F(4,6061) = 509.5, p < 0.001). The highest value was obtained by students from private high schools. The other stratas' means, in descendent order, were urban public academic, urban public technical and rural high schools (there was not statistical significance difference between academic and technical rural high schools).

Finally, the explained variance of reading literacy provided by the model with only contextual variables was 18.65%, (se = 1.90%); then, the increase in variance explained when including student and teacher elements was 7.47% (se =.68%). This implies that 28.19% of the explained variance of the reading literacy by the independent variables is due only to non-contextual factors.

DISCUSSION

In this study a set of variables were analyzed that, according to the literature (Bormuth, 1973; Wilkinson, 1998; Artelt et al., 2001; Shiel and Cosgrove, 2002; Rasmussen, 2003; Connor et al., 2005; Morrison et al., 2005; Brozo et al., 2014) could be associated with reading literacy. Individually, it was found that the contextual variables presented solid evidence of association with reading literacy, while from the teacher's and student's variables, only the interest in reading showed a relevant linear correlation as argued in the theory (Montero et al., 2012; Valdés, 2013; Castro, 2014; Dezcallar et al., 2014).

Another conclusion of the results obtained from the matrix correlations is the subsets of variables related between them. It was found that the variables of each group defined previously (contextual, student and teacher variables) presented high correlations between them, but low correlations with the rest of variables. The only exception was student absenteeism; this variable did not show relevant correlations with student's reading time nor with their interest in reading. This result suggests that students do not appreciate Spanish classes and reading in the same way, and maybe it is because reading it is not associate with a specific course. It should be remembered that in the Costa Rican context, literature is approached in the subject of the country's official language. On the other hand, it is expected that a student who loves reading, wants to go to a class about interesting things about books. This analysis implicates that students evaluate Spanish classes like a course far away from the hobby of reading.

These results show that the contextual factors present marked relations with reading literacy outcomes. The results support the hypothesis that students from households with high educational and economic climates have better reading literacy scores than those from households with less favorable conditions. This inequality is not due to the high values of these variables, per se, but to the large number of favorable conditions for the development of reading literacy that these variables entail: availability of books, early reading promotion, access to reading peers, investment of time in leisure and cultural recreation activities, among others.

The teacher variables did not present such marked associations with reading literacy as the contextual factors did. This result was expected because contextual variables have influenced students' lives throughout their lifespan while those of teachers only at specific times, being consistent with the studies conducted in Montero et al. (2012) and Dezcallar et al. (2014). On the other hand, the student's variables showed more relevant correlations than those observed in the teacher's factors. This may be because student's variables are highly determined by contextual factors, so that student's variables may reflect a part of the behavior of the contextual ones and, therefore, show some marked correlations.

The analysis of all the variables together in the linear regression with plausible values and sample weights showed that the variables with relevant individual correlations were the ones that presented the relevant regression coefficients. Nonetheless, the second most relevant variable was not a contextual one, but one associated with student's practices: interest in reading. As mentioned in the introduction, interest in reading has been linked to better performances in reading literacy (Montero et al., 2012; Valdés, 2013; Castro, 2014; Dezcallar et al., 2014). This result is due to the fact that people with a higher interest in reading develop better text comprehension skills and other areas, as they seek to understand in a deeper way the readings they undertake. Likewise, as mentioned before, those who have a better reading

comprehension develop an effective reading competence from which they draw a clear and motivational purpose for reading, ranging from the recreational enjoyment of the texts, as in the case of fiction books, to the acquisition of information and expansion of specific knowledge through scientific texts or specific subjects. Moreover, it was showed that interest in reading' means by stratas were similar and low, it implies that there is a wide margin to improve this variable in all this groups.

It is important to mention that the variable reading time presents a similar relevance to that of interest in reading when the latter variable is removed from the model; therefore, it is an important variable in the prediction of reading literacy (Shiel and Cosgrove, 2002; OECD, 2011; Schiefele et al., 2012; Valdés, 2013; Brozo et al., 2014; Dezcallar et al., 2014; Yubero and Larrañaga, 2015; Guerra and Guevara, 2017). This implies that for people whose environments have the same considered variables in contextual and teacher's work aspects, the variable interest in reading or reading time is positively and significantly associated with reading literacy. This result is encouraging, as it shows two variables that can be worked on improving reading literacy, despite the contexts students face.

As for the teacher variables, these had little relevance, but we can analyze the variable with the more relevant coefficient (teacher's group control) because it showed a weak effect that can be useful to an intervention. This result indicates that proper group management can contribute to increased reading literacy scores (Omoteso and Semudara, 2011; Akiri, 2013; Duke et al., 2016). This result is justified by the fact that effective language teaching requires a suitable classroom environment. When there is no proper group control, students face many distractions that get in the way of the development of the intended competencies.

The teacher's and student's variables added 28.19% of the explanation of the variance of reading literacy given by the model with only contextual variables. This implies that a portion of the variability of the grades is explained by teacher's and student's variables alone with no incidence of contextual elements. Based on this, it is concluded that there is a range of action that can contribute to the improvement of reading literacy, despite the contextual factors of the students.

Finally, it is important to mention that even though the teacher variables were the less relevant among the non-contextual variables (because they got the lower standardized coefficients), the role of the instructor in the improvement of the student's variables is crucial. Based on the results obtained in this article, the teacher should look for ways to increase students' interest in reading and encourage them to dedicate more time to it which would lead to better results in reading literacy. This task falls on teachers, for example, uneducated or overworked parents cannot be asked to help their children become involved in reading.

Among the activities that teachers can do so that students develop more interest in reading or devote more time to it are to provide reading material according to the student's preferences, without imposing a type of text, be it fiction or not; to establish meeting points or comparative exercises between reading and other texts such as filmic texts, paintings, comics, music, among others, and to promote collective readings with their consequent discussion.

Some of the limitations of the study were that from the 7,119 students initially chosen, only 4,691 completed all the instruments. Other limitation was the low level of explanation of the teacher's variables, one of the hypotheses of this study was that these variables had more impact in the reading literacy. We think that is necessary to study the validity of the teacher's scales because they showed unexpected results.

It is still pending to include new variables, for example, in contextual factors to which the support resources offered by the institution could be added which include time dedicated to Spanish or literature classes, technological stimuli and access to extra-class activities.

DATA AVAILABILITY STATEMENT

The datasets for this study can be found in the repository PISA 2018 database, available in http://www.oecd.org/pisa/data/2018database/

ETHICS STATEMENT

The Education Department of Costa Rica approved PISA test administration developed by OECD, along with the ethical aspects of the process.

AUTHOR CONTRIBUTIONS

LR was the primary author of the manuscript. All authors contributed to the writing and editing of the manuscript.

FUNDING

The authors received no specific funding for this work. However, University of Costa Rica has funded several authors' projects related with the topic studied in this article. We believe that the University of Costa Rica gave a relevant support to this article.

REFERENCES

Acock, A. C. (2014). *A Gentle Introduction to Stata*. 4th ed.. Texas: Stata Press.

Akiri, A. A. (2013). Effects of Teachers' Effectiveness on Students' Academic Performance in Public Secondary Schools; Delta State—Nigeria. *J. Educ. Sci. Res.* 3 (3), 105–111. doi:10.5901/jesr.2013.v3n3p105

Alves, A. F., Gomes, C. M. A., Martins, A., and Almeida, L. d. S. (2017). Cognitive Performance and Academic Achievement: How Do Family and School Converge? *Eur. J. Educ. Psychol.* 10, 49–56. doi:10.1016/j.ejeps.2017.07.001

Artelt, C., Schiefele, U., and Schneider, W. (2001). Predictors of Reading Literacy. *Eur. J. Psychol. Educ.* 16 (3), 363–383. doi:10.1007/BF03173188

Bormuth, J. R. (1973). Reading Literacy: Its Definition and Assessment. *Reading Res. Q.* 9 (1), 7–66. doi:10.2307/747227

Brenes, G. (2019). El peso de los factores contextuales en la asistencia a la secundaria entre los adolescentes costarricenses según la información censal. *Estudios Demográficos Y Urbanos* 34 (2), 395–426. doi:10.24201/edu.v34i2.1636

Brozo, W. G., Sulkunen, S., Shiel, G., Garbe, C., Pandian, A., and Valtin, R. (2014). Reading, Gender, and Engagement. *J. Adolesc. Adult Liter* 57 (7), 584–593. doi:10.1002/jaal.291

Castro, D. (2014). Motivaciones, actitudes y prácticas lectoras de los alumnos mexicanos y el nivel de desempeño en lectura de la prueba de pisa 2009. (Análisis del constructo "leer por placer"). *Tendencias Pedagógicas* 23, 301–318. doi:10.1787/888932510751

Chen, J., and Lin, T.-F. (2008). Class Attendance and Exam Performance: A Randomized Experiment. *J. Econ. Educ.* 39, 213–227. doi:10.3200/JECE.39.3.213-227

Connor, C. M., Son, S. H., Hindman, A. H., and Morrison, F. J. (2005). Teacher Qualifications, Classroom Practices, Family Characteristics, and Preschool Experience: Complex Effects on First Graders' Vocabulary and Early Reading Outcomes. *J. Sch. Psychol.* 43, 343–375. doi:10.1016/j.jsp.2005.06.001

Devadoss, S., and Foltz, J. (1996). Evaluation of Factors Influencing Student Class Attendance and Performance. *Am. J. Agric. Econ.* 78, 499–507. doi:10.2307/1243268

Dezcallar, T., Clariana, M., Cladelles, R., Badia, M., and Gotzens, C. (2014). La lectura por placer: su incidencia en el rendimiento académico, las horas de televisión y las horas de videojuegos. *Ocnos* 12, 107–116. doi:10.18239/ocnos_2014.12.05

Díez, A., and Egío, V. C. (2017). La competencia lectora. Una aproximación teórica y práctica para su evaluación en el aula. *Investigaciones sobre Lectura* 7, 22–35. doi:10.24310/revistaisl.vi7.10979

Duke, N. K., Cervetti, G. N., and Wise, C. N. (2016). The Teacher and the Classroom. *J. Educ.* 196 (3), 35–43. doi:10.2307/26612627 10.1177/002205741619600306

Edel, R. (2003). El Rendimiento Académico, Concepto, Investigación Y Desarrollo. *Revista Electrónica Iberoamericana de Calidad, Eficacia y Cambio en Educación* 1 (15), 114, 2003 . Available at: https://repositorio.uam.es/handle/10486/660693 doi:10.23913/ride.v11i21.719

García, M. A., Arévalo, M. A., and Hernández, C. A. (2018). La Comprensión Lectora Y El Rendimiento Escolar. *Cuadernos de Lingüística Hispánica* 32, 155–174. doi:10.19053/0121053X.n32

Guerra, J., and Guevara, C. Y. (2017). Variables académicas, comprensión lectora, estrategias y motivación en estudiantes universitarios. *Revista electrónica de investigación educativa* 19 (2), 78–90. doi:10.24320/redie.2017.19.2.1125

Guo, Y., Connor, C. M., Yang, Y., Roehrig, A. D., Morrison, F. J., and Morrison, F. (2012). The Effects of Teacher Qualification, Teacher Self-Efficacy, and Classroom Practices on Fifth Graders' Literacy Outcomes. *Elem. Sch. J.* 113 (1), 3–24. doi:10.1086/665816

Hernández Padilla, E., and Bazán Ramírez, A. (2016). Efectos contextuales, socioeconómicos y culturales, sobre los resultados de México en Lectura en PISA 2009. *Reice* 14 (2), 123. doi:10.15366/reice2016.14.2.005

Jiménez, E. (2013). Comprensión lectora vs competencia lectora: qué son y qué relación existe entre ellas. *Investigaciones sobre lectura* 1, 65–74. doi:10.24310/revistaisl.vi1.10943

Lamas, H. (2015). Sobre El Rendimiento Escolar. *Propósitos y representaciones* 3 (1), 313–386. doi:10.20511/pyr2015.v3n1.74

Lin, M., Lucas, H. C., and Shmueli, G. (2013). Research Commentary-Too Big to Fail: Large Samples and Thep-Value Problem. *Inf. Syst. Res.* 24 (4), 906–917. doi:10.1287/isre.2013.0480

López-Bonilla, J. M., and López-Bonilla, L. M. (2013). The Multidimensional Structure of University Absenteeism: an Exploratory Study. *Innov. Educ. Teach. Int.* 52 (2), 185–195. doi:10.1080/14703297.2013.847382

Meroni, E. C., Vera-Toscano, E., and Costa, P. (2015). Can Low Skill Teachers Make Good Students? Empirical Evidence from PIAAC and PISA. *J. Pol. Model.* 37 (2), 308–323. doi:10.1016/j.jpolmod.2015.02.006

Montero, E., Rojas, S., Zamora, E., and Rodino Pierri, A. M. (2012). Costa rica en las pruebas PISA 2009 de Competencia Lectora y Alfabetización Matemática. Ponencia presentada como investigación base para el Informe Estado de la Educación 2013 (no. 4). Available at: http://hdl.handle.net/20.500.12337/861

Morrison, F. J., Bachman, H. J., and Connor, C. M. (2005). *Improving Literacy in America: Guidelines from Research*. New Haven, CT: Yale University Press. doi:10.12987/yale/9780300106459.001.0001

OECD (2010). *PISA 2009 Results: Learning to Learn—Student Engagement, Strategies and Practices (Volume III)*. Berlin: Springer.

OECD (2019a). *PISA 2018 Results (Volume I): What Students Know and Can Do. PISA*. Paris: OECD Publishing.

OECD (2019b). *PISA 2018 Technical Report*. PISA, Paris: OECD Publishing.

OECD (2009). *PISA Data Analysis Manual. L: SPSS*. Second Edition. Paris: PISA, OECD Publishing.

OECD (2011). *PISA in Focus*. Paris: OECD Publishing.

OECD (2005). *Proyecto DESECO: La definición y selección de competencias clave. Resumen ejecutivo (versión electrónica)*. Paris: OECD Publishing.

Omoteso, B. A., and Semudara, A. (2011). The Relationship between Teachers' Effectiveness and Management of Classroom Misbehaviours in Secondary Schools. *Psych* 02, 902–908. doi:10.4236/psych.2011.29136

Rasmussen, J. B. (2003). Reading Literacy Performance in Norway: Current Practice and Critical Factors. *Eur. J. Educ.* 38 (4), 427–443. doi:10.1111/j.0141-8211.2003.00159.x

Romer, D. (1993). Do students Go to Class? Should They? *J. Econ. Perspect.* 7 (3), 167–174. doi:10.1257/jep.7.3.167

Schiefele, U., Schaffner, E., Möller, J., and Wigfiel, A. (2012). Dimensions of Reading Motivation and Their Relation to Reading Behavior and Competence. *Reading Res. Q.* 47 (4), 427–463. doi:10.1002/RRQ.030

Schleicher, A. (2019). *PISA 2018: Insights and Interpretations. PISA*. Paris: OECD Publishing.

Schmulian, A., and Coetzee, S. (2011). Class Absenteeism: Reasons for Non-attendance and the Effect on Academic Performance. *Account. Res. J.* 24 (2), 178–194. doi:10.1108/10309611111163718

Shiel, G., and Cosgrove, J. (2002). International Perspectives on Literacy: International Assessments of Reading Literacy. *Reading Teach.* 55 (7), 690–692. Retrieved from: http://www.jstor.org.ezproxy.sibdi.ucr.ac.cr:2048/stable/20205119

Solé, I. (2012). Competencia Lectora Y Aprendizaje. *Rieoei* 59, 43–61. doi:10.35362/rie590456

Teixeira, A. A. C. (2016). The Impact of Class Absenteeism on Undergraduates' Academic Performance: Evidence from an Elite Economics School in Portugal. *Innov. Educ. Teach. Int.* 53 (2), 230–242. doi:10.1080/14703297.2014.937730

Trejos, J. D. (2010). *Indicadores sobre equidad en la educación para Costa Rica*. Costa Rica: Tercer Informe Estado de la Educación. Informe final. San José, Costa Rica: Programa Estado de la Nación. Recuperado de: http://hdl.handle.net/20.500.12337/776

Valdés, M. (2013). ¿Leen en forma voluntaria y recreativa los niños que logran un buen nivel de Comprensión Lectora? *Ocnos* 10, 71–89. doi:10.18239/ocnos_2013.10.04

Vidal-Moscoso, D., and Manriquez-López, L. (2016). El docente como mediador de la comprensión lectora en universitarios. *Revista de la Educación Superior* 45 (177), 95–118. doi:10.1016/j.resu.2016.01.009

Wilkinson, I. (1998). Dealing with Diversity: Achievement Gaps in Reading Literacy Among New Zealand Students. *Rrq* 33 (2), 144–167. doi:10.1598/RRQ.33.2.1

Wray, D., Medwell, J., Fox, R., and Poulson, L. (2000). The Teaching Practices of Effective Teachers of Literacy. *Educ. Rev.* 52 (1), 75–84. doi:10.1080/00131910097432

Yubero, S., and Larrañaga, E. (2015). Lectura y universidad: hábitos lectores de los estudiantes universitarios de España y Portugal. *Epi* 24 (6), 717–723. doi:10.3145/epi.2015.nov.03

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 22 November 2021
doi: 10.3389/feduc.2021.731634

Socioeconomic Status as a Multidimensional Predictor of Student Achievement in 77 Societies

Kimmo Eriksson[1]*, *Jannika Lindvall*[1], *Ola Helenius*[2] and *Andreas Ryve*[1]

[1]School of Education, Culture and Communication, Mälardalen University, Västerås, Sweden, [2]NCM, University of Gothenburg, Göteborg, Sweden

We reassess the relation between students' socioeconomic status (SES) and their achievement by treating SES as multidimensional instead of unidimensional. We use data from almost 600,000 students in 77 countries participating in the 2018 PISA assessment of student achievement in math, science, and reading. The composite measure of SES that PISA uses can be broken down into six component variables that we here use as simultaneous predictors of achievement. This analysis yields several new insights. First, in the typical society, two predictors (books at home and parents' highest occupational status) clearly outperform the rest. Second, a new composite measure based only on these two components often reveals substantially larger achievement gaps than those reported by PISA. Third, the analysis revealed remarkable differences between societies in the relation between achievement and wealth possessions. In most societies, the independent effect of wealth possessions on student achievement was zero or even slightly negative—but in the least developed societies it was strongly positive. These findings have implications for how SES achievement gaps should be measured and interpreted.

Keywords: student achievement, achievement gap, socioeconomic status, human development, composite measure

OPEN ACCESS

Edited by:
Ana E. Azpilicueta, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

Reviewed by:
Kim Koh, University of Calgary, Canada
Lokman Akbay, Istanbul University-Cerrahpasa, Turkey

***Correspondence:**
Kimmo Eriksson
kimmo.eriksson@mdh.se

Specialty section:
This article was submitted to Assessment, Testing and Applied Measurement, a section of the journal Frontiers in Education

Received: *27 June 2021*
Accepted: *26 October 2021*
Published: *22 November 2021*

Citation:
Eriksson K, Lindvall J, Helenius O and Ryve A (2021) Socioeconomic Status as a Multidimensional Predictor of Student Achievement in 77 Societies. Front. Educ. 6:731634. doi: 10.3389/feduc.2021.731634

INTRODUCTION

For more than half a century, socioeconomic status (SES) has been recognized as a major influence on student achievement (White, 1982; Coleman et al., 1966; Sirin, 2005; Harwell et al., 2017). In this body of research, SES is typically treated as a unidimensional predictor that can be operationalized in many different ways, such as by parents' educational attainment, parents' occupations, or the family's economic resources, or by a composite of these factors (Cowan et al., 2012). Our research question here is what we can learn by instead treating SES as a multidimensional predictor of achievement.

Definitions of SES typically refer to an individual's or a family's position on a hierarchical social structure based on their control and/or access to resources like wealth, prestige, power, and social and cultural capital (Mueller and Parcel, 1981; Willms and Tramonte, 2019). The dominant perspective in this research area is that different socioeconomic factors are not important in their own right. Key socioeconomic factors such as parents' income and occupations are treated merely as different indicators of the presumed unidimensional hierarchical social structure that influences the achievement of students. In this perspective, the only relevant comparison between different socioeconomic is which of them serves best as an indicator of SES and the literature offers a range of views on this issue (e.g., Blau and Duncan, 1967; Lienet al., 2001; Schulz, 2005; Marks, 2011;

Jerrim et al., 2019). A common view is that a composite of several socioeconomic factors is preferred, both to reflect how the concept is defined and to reduce measurement error (Cowan et al., 2012). This is the view taken by the international large-scale assessment PISA, which uses a composite measure based on parental educational attainment, parental occupational status, and home possessions (Avvisati, 2020).

For the big picture about the relation between SES and student achievement, it does not really matter how SES is operationalized. A positive relation with achievement tends to be found whether SES is represented by a single indicator or a composite measure based on several indicators, although the effect size may vary across different operationalizations (White, 1982; Sirin, 2005; Harwell et al., 2017). However, important information may be lost when SES is treated as unidimensional. Our view is that SES should instead be regarded as a multidimensional predictor of achievement with different socioeconomic measures to be used as multiple separate variables. While similar views have been expressed in passing by others (e.g., Harwell et al., 2017; Willms and Tramonte, 2019), it is difficult to find any studies that have actually used multiple separate socioeconomic measures. Harwell and colleagues note that to do so would be at odds with recent recommendations. To defend our view, we must therefore examine the basis for these recommendations.

The NCES Report on the Measurement of Socioeconomic Status

The recommendations referred to by Harwell et al. (2017) are given in a report from a panel of experts convened by the National Center for Education Statistics in the United States (Cowan et al., 2012). The panel was tasked with providing recommendations concerning the definition and measurement of socioeconomic status. Their report explicitly declares an instrumental view of SES: "Researchers and policy makers are interested in SES as a contextual variable to study educational equity and fairness issues, as a covariate with achievement to examine the effects of other variables such as class size or school governance policies, and as a matching variable to ensure the equivalence of treatment and control groups in educational intervention studies" (Cowan et al., 2012, p. 7). The report's argument against treating SES as multiple separate variables is that doing so could lead to potentially conflicting results for different variables, thereby complicating interpretation. Instead, the report recommends the use of a composite variable to combine information from multiple variables while "avoiding conflicting stories about relationships to achievement" (Cowan et al., 2012, p. 22). In line with the report's instrumental view of SES, this argument amounts to a preference for simplicity over complexity. However, to the extent that a phenomenon is in fact complex, a simplistic approach may stand in the way of deeper understanding.

Advantages of Treating SES as a Multidimensional Predictor of Achievement

Treating SES as multidimensional instead of unidimensional has several advantages. One advantage is the possibility of gaining a more detailed understanding of the phenomenon. Conspicuously absent in most empirical research on the link between SES and student achievement is any detailed consideration of what mechanisms cause this link in the first place. Yet, detailed knowledge of the pathways through which a socioeconomic advantage turns into an achievement advantage should be of great value to researchers and policy makers interested in leveling the playing field. Potential mechanisms suggested in the literature include genetic transfer of skills across generations, nonfinancial inputs into children's development (e.g., reading stories and helping with homework), monetary inputs into children's development (e.g., tuition fees and paying for private tuition), and the negative effects of high stress levels caused by economic hardship (Jerrim and Macmillan, 2015; Rözer and van de Werfhorst, 2019). Note that none of these mechanisms refer to an abstract social hierarchy. Rather it seems that what matters is a variety of more concrete things like genes, skills, money, time, etc. It should therefore be possible to gain a richer understanding by disentangling the separate effects of these things. For instance, consider the possession of wealth. People may acquire wealth in various ways, not all of them related to long education or a high-status occupation. To the extent that wealth has a direct effect on student achievement (the aforementioned "monetary input" pathway), it should be largely independent of where the money comes from and thus separable from other effects. By including parents' wealth, education, and occupation as simultaneous separate predictors of achievement we can examine whether wealth in fact has any independent effect.

A second advantage of treating SES as multidimensional arises when estimating the total amount of variation in achievement that is accounted for by socioeconomic factors. This can be thought of as the strength of the relation between SES and achievement, or simply the "socioeconomic achievement gap." Much research has focused on the size of the achievement gap and meta-studies find very different estimates across studies (White, 1982; Sirin, 2005; Harwell et al., 2017). In particular, there are strong indications that different choices of SES components yield different estimates of the achievement gap (Sirin, 2005). If different socioeconomic factors have independent effects on achievement, the use of any single factor will necessarily underestimate the total amount of variation in achievement that is accounted for by socioeconomic factors. An advantage of treating SES as multidimensional is that it helps avoid such underestimation.

Underestimation of the total effect of SES on achievement could also be achieved by use of a composite SES measure that is constructed through an optimal choice of component weights. A third advantage of using SES components as multiple separate predictors is that the results of such analyses provide the optimal weights for a composite measure to avoid underestimation of the SES effect. Extant composite measures are usually constructed based on other principles (Avvisati, 2020), hence do not avoid the underestimation problem.

Finally, estimations of the effect of SES on achievement vary considerably in magnitude across different societies and tend to be lower in developing countries (e.g., OECD, 2018; Kim et al., 2019). This cross-societal variation in SES effects is something of

a puzzle, because it is not accounted for by factors such as the number of teaching hours, reduced class size, and teacher quality (Strietholt et al., 2019; Rözer and van de Werfhorst, 2019). However, the cross-societal variation in SES effects may be mysterious in part because it is not well-defined. Prior research suggests that the relative relevance of different socioeconomic factors in a society depends on its development level (Kim et al., 2019). It is therefore possible that country differences in the effect of SES on achievement look completely different if SES is operationalized by, say, wealth or parents' occupational status. A fourth advantage of treating SES as a multidimensional predictor, specifically in a multi-society study, is that it enables an examination of how different SES measures interact with societal factors. Such analyses could provide crucial insights into the reasons why SES achievement gaps vary with the development level of countries.

Aims of the Current Study

Above we discussed four potential advantages of treating SES as a multidimensional predictor of achievement. The current study aims at empirically demonstrating these advantages using data from the international large-scale assessment PISA. By estimating the independent effects on achievement of the different components of SES, separately in each participating society, we 1) examine which components tend to have the largest independent effects, 2) assess how much the composite measure underestimates the SES achievement gap compared to multiple components, 3) propose an alternative composite measure with more desirable properties, and 4) examine how the independent SES component effects vary with the development level of countries.

METHODS

PISA is an international assessment of 15-year-old students' achievement in math, reading, and science, conducted by the Organization for Economic Cooperation and Development (OECD). PISA uses a representative sample of students from each participating country, with sample sizes usually around 5,000 per country but sometimes substantially larger, see the official report for details on the sampling strategy (OECD, 2019). In the present study we use data from 2018, the most recent wave of PISA for which data is available at this time. We use data from 72 participating countries plus 5 additional participating entities (Hong Kong, Taiwan, Macau, Beijing-Shanghai-Jiangsu-Zhejiang, and Baku in Azerbaijan), for a total of 77 societies with almost 600,000 participating students, generally born in 2002. See **Supplementary Table S1** for the full list of societies and sample sizes.

Achievement Measures

Full descriptions of the math, science, and reading skills assessed in PISA are available in the official report (OECD, 2019). To assess a broader range of topics at the country level, PISA only tests each student on a subset of the complete tests. Based on their test results, 10 "plausible values" are imputed for students' achievement in each domain. Special software is available for analyses of plausible values, see below.

Measures of Socioeconomic Status

To obtain multiple measures of SES we started with PISA′ composite measure and decomposed using subdivisions provided by PISA. As described below, this yielded six component measures: parents' highest occupational status, parents' highest educational level, wealth possessions, cultural possessions, home educational resources, and books at home. Below we describe these measures (with the name of the variable in the PISA dataset within parentheses). For further details, see Chapter 16 of the PISA 2018 Technical Report,[1] especially table 16.4 in that chapter.

PISA's Composite Measure of Socioeconomic Status

PISA offers the ESCS composite measure of socioeconomic status. It is based on three variables—parents' highest occupational status, parents' highest educational level, and home possessions—which are standardized and then averaged to an index. Home possessions are based on a set of 25 items which PISA subdivides further into four variables: wealth possessions, cultural possessions, home educational resources, and the number of books at home.

Parents' Highest Occupational Status

PISA asked open-ended questions to students on the occupations of their mother and father. Responses were coded and mapped to an international socioeconomic index of occupational status (Ganzeboom, 2010). HISEI is the higher score of either parent or the only available parent's score.

Parents' Highest Educational Level

PISA asked students about the educational level of their mother and father, ranging from primary education to post-graduate education. PAREDINT is the higher level of either parent, transformed into years of education based on an international standard.

Wealth Possessions

The wealth possessions variable is based on 12 items, such as possession of cars and a room of your own. There are also some country-specific wealth items.

Cultural Possessions (CULTPOSS)

The cultural possessions variable is based on 5 items relating to literature, art, and music.

Home Educational Resources

The home educational resources variable is based on 7 items relating to studies at home, such as a desk to study at, a computer to use, a dictionary, etc.

[1]https://www.oecd.org/pisa/data/pisa2018technicalreport/PISA2018_Technical-Report-Chapter-16-Background-Questionnaires.pdf

TABLE 1 | Mean (SD) within-society mean values, standard deviations, and Pearson intercorrelations among SES components in 77 societies.

	Mean value	Std dev	Correl. w. (1)	Correl. w. (2)	Correl. w. (3)	Correl. w. (4)	Correl. w. (5)
(1) Parents' highest occupational status	50.0 (21.7)	6.8 (1.7)	-				
(2) Parents' highest educational level	13.4 (2.6)	1.1 (0.7)	0.46 (0.07)	-			
(3) Wealth possessions	–0.5 (0.9)	0.7 (0.2)	0.30 (0.11)	0.28 (0.10)	-		
(4) Cultural possessions	–0.2 (0.9)	0.3 (0.1)	0.26 (0.06)	0.25 (0.05)	0.30 (0.07)	-	
(5) Home educational resources	–0.2 (1.0)	0.4 (0.1)	0.23 (0.07)	0.23 (0.07)	0.38 (0.12)	0.39 (0.05)	-
(6) Books at home	2.9 (1.3)	0.5 (0.1)	0.32 (0.08)	0.28 (0.06)	0.27 (0.07)	0.46 (0.08)	0.29 (0.05)

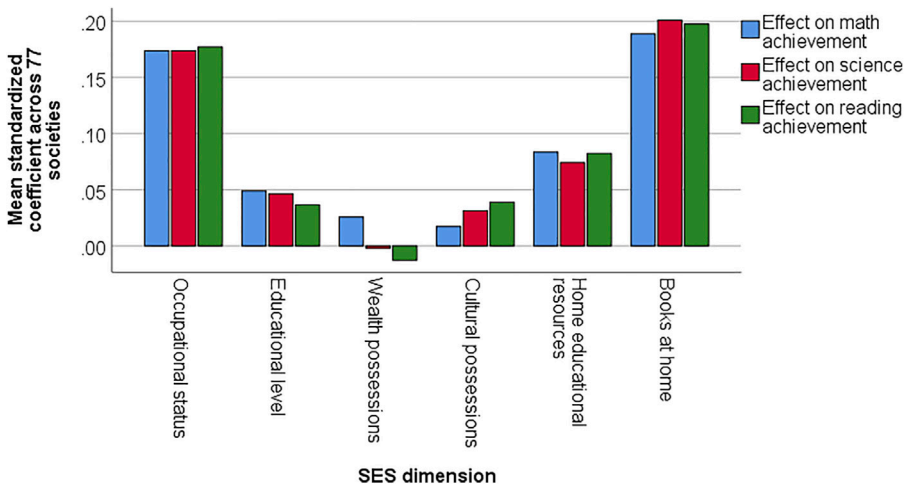


FIGURE 1 | Estimates of the independent effects on achievement of six SES components. Bars show mean values of standardized coefficients across 77 societies. Different colors denote different achievement domains: math (blue), science (red), and reading (green).

Books at Home (ST013Q01TA)

Students were asked to estimate the number of books at home on a six-step scale (1 = "0–10 books," 2 = "11–25 books," 3 = "26–100 books," 4 = "101–200 books," 5 = "201–500 books," 6 = "More than 500 books"). Note that the books at home variable is sometimes used on its own as a single-item measure of SES (e.g., Blömeke et al., 2016; Eriksson et al., 2019).

Analysis

We used the IDB Analyzer provided by the International Association for the Evaluation of Educational Achievement (IEA, 2017). The IDB Analyzer creates SPSS syntax to analyze PISA data in such a way that standard errors correctly reflect the complex design of the study (e.g., the use of plausible values). Using the IDB Analyzer we calculated the correlations, multiple linear regressions, and quartile means described below.

RESULTS

Intercorrelations of SES Components

Table 1 reports descriptive statistics for within-society mean values, standard deviations, and intercorrelations among the six SES components. Note that the correlations between different SES components were not very strong. Averaged across societies, the largest correlations were well below 0.50 and most correlations were below 0.30. In other words, different SES components are not very closely related to each other. There is therefore good reason to conceive of them as distinct dimensions and to examine their independent effects on student achievement.

Six SES Components' Independent Effects on Student Achievement

In each society we performed multiple linear regression analyses of student achievement in three different domains, using all six SES components as predictors. For comparability across SES components, we focus on the standardized coefficients. These coefficients estimate the standardized increase in achievement from an increase of the predictor by one standard deviation.

Domain Generality of SES Component Effects on Achievement

Figure 1 shows the mean value of the coefficient (i.e., averaged across societies) for each SES component, separately for the three academic domains. Note that the pattern of results was almost identical across different academic domains. For example, the effect of books at home was as large in the domains of science and math as in the domain of reading. We may therefore disregard

TABLE 2 | Descriptive statistics of the standardized coefficients of different SES components, averaged across three academic domains, in 77 societies.

	M	SD	Min	Max
(1) Parents' highest occupational status	0.17	0.05	0.06	0.27
(2) Parents' highest educational level	0.04	0.05	–0.07	0.14
(3) Wealth possessions	0.00	0.10	–0.16	0.34
(4) Cultural possessions	0.03	0.05	–0.11	0.14
(5) Home educational resources	0.08	0.06	–0.03	0.27
(6) Books at home	0.20	0.08	0.03	0.34

domain in analyses. For each SES component we averaged the effect across the three domains (Cronbach's $\alpha > 0.98$). Descriptive statistics of these average effect measures are reported in **Table 2**.

Special Importance of Books at Home and Parents' Occupations Status

A second thing to note in **Figure 1** is that books at home and parents' occupational status had much larger average effects on student achievement than the other SES components. This is an important finding. Among other things, it suggests that a useful composite SES index could be based on only these two components. We turn to this topic next.

Alternative Measures of the SES Achievement Gap

In their executive report, PISA reports two different measures of the SES achievement gap in a society. One measure is the mean difference in achievement between advantaged students and disadvantaged students, operationalized as the highest and lowest quarter of students on the ESCS index, respectively. Another measure is the proportion of variance in achievement (R^2) explained by the ESCS index. The ESCS index adds together the different SES components in a way that does not reflect their relative effects on achievement. Estimates of the SES achievement gap that are based on ESCS index will therefore underestimate the SES achievement gap. To illustrate this underestimation, we construct an alternative SES index based on the two SES components that had the largest effects on student achievement: books at home and parents' occupational status. After standardization (across the whole dataset) of these two SES components, we average them into a *two-item SES index*. We compare the estimate of SES achievement gaps when SES was measured by the ESCS vs. the two-item index.

The Two-Item Index Yields Larger Estimates of the Achievement Gap

Figure 2 shows that when using the two-item SES index instead of ESCS, the mean difference in achievement between advantaged students and disadvantaged students increased by about 11% (or 9 points on the test score) in the average society. The largest achievement gap in any society increased even more, by about 19% (or 23 points). Results were similar across all academic domains.

The Two-Item Index Explains More Variance in Achievement

In **Table 3** we report the proportion of variance in student achievement explained by the ESCS index and the two-item SES index, respectively. Consistent with the previous analysis, the two-item SES index explained more variance than the ESCS index in the average society (15 percent vs. 12-13 percent), and the difference was even larger in societies with the largest SES achievement gap (29-31 vs. 21-24 percent variance explained).

By using all six components as multiple predictors instead we will inevitably be able to explain even more variance. This analysis will produce the optimal weighting of all six components. Because we run separate analyses for each society, weightings will be optimized specifically for every society. The results of this analysis are reported in the last column of **Table 3**. Despite the inherent advantage of this method in accounting for variance, the proportion of variance explained only increased marginally compared to the simple two-item index. This was particularly

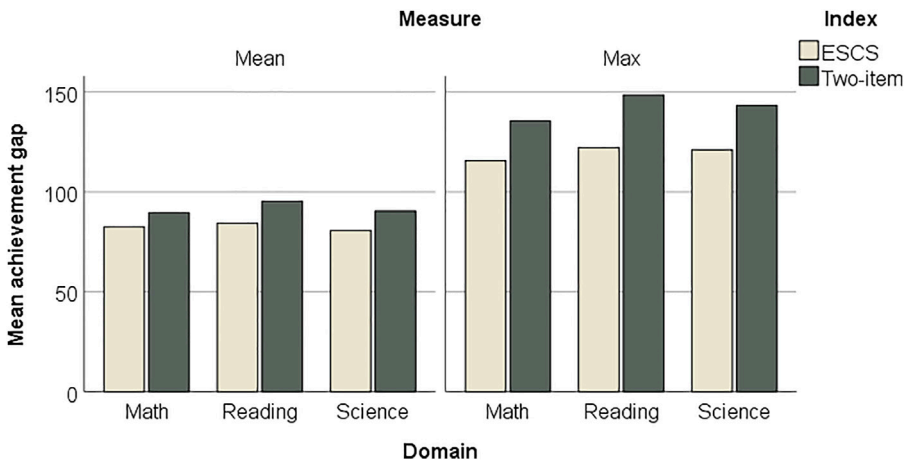


FIGURE 2 | SES achievement gaps for math, reading, and science, estimated either using the composite measure provided by PISA (light bars) or using the new two-item SES index (dark bars). The latter index yielded larger gap estimates, whether looking at mean values across 77 societies **(left panel)** or the maximal value among all 77 societies **(right panel)**.

TABLE 3 | The proportion of variance (R^2) in achievement explained by socioeconomic status when SES is operationalized either by PISA's composite measure (ESCS), or by the new two-item index, or by multiple predictors.

	R^2 (ESCS)		R^2 (two-item SES index)		R^2 (multiple predictors)	
	M	Max	M	Max	M	Max
Math	0.13	0.24	0.15	0.31	0.17	0.32
Science	0.12	0.21	0.15	0.31	0.17	0.32
Reading	0.12	0.21	0.15	0.29	0.17	0.30

TABLE 4 | Mean (SD) Pearson correlations among the standardized coefficients (averaged across academic domains) of six SES components and the Human Development Index in 77 societies.

	Correl. With (1)	Correl. With (2)	Correl. With (3)	Correl. With (4)	Correl. With (5)	Correl. With (6)
(1) Std. coeff. of parents' highest occupational status	–					
(2) Std. coeff. of parents' highest educational level	–0.30**	–				
(3) Std. coeff. of wealth possessions	–0.16	–0.11	–			
(4) Std. coeff. of cultural possessions	0.15	0.15	–0.52***	–		
(5) Std. coeff. of home educational resources	–0.14	–0.05	–0.08	–0.39**	–	
(6) Std. coeff. of books at home	0.30**	–0.10	–0.52***	0.21	–0.25*	–
(7) Human Development Index	0.09	0.15	–0.62***	0.25*	–0.20	0.71***

*** p < *0.001*, ** p < *0.01*, * p < *0.05 (two-sided).*

evident in the societies with the largest SES achievement gap, where the two-item index accounted for 31 percent of the variance and the six components together accounted for 32 percent, a negligible difference. As usual, similar results were obtained across all three academic domains.

The Multi-Predictor Model is Preferred by the Bayesian Information Criterion

Although the use of multiple predictors allows more variance to be explained, this gain will to some extent reflect overfitting. To assess whether the more complex model is in fact warranted, researchers use model selection criteria such as the Bayesian Information Criterion (BIC). The model that yields a smaller BIC value is preferred. BIC values are not provided in analyses produced by the IDB Analyzer. However, BIC values are provided by SPSS for standard linear regressions. We therefore used SPSS to perform linear regressions corresponding to those presented in **Table 3**. These alternative analyses replicated the pattern of results in **Table 3**, that is, ESCS explained less variance than the two-item index, which in turn explained slightly less variance than the multi-predictor model. Moreover, they showed that the same pattern holds for BIC, that is, the ESCS model had a higher BIC value (M = 42,729 across countries and domains) than the two-item index model (M = 39,767), which in turn had a slightly higher BIC value than the multi-predictor model (M = 38,423). We conclude that using SES components as multiple predictors is warranted.

Cross-Societal Variation in the Effects of SES on Achievement

We now turn to how the specific effects of different SES components on student achievement varies across societies. The spread between the minimum and maximum effects among the 77 societies in the study (**Table 2**) indicates that the cross-societal variation in the size of SES effects is substantial. If we want to order societies on the size of the SES achievement gap, does it matter which SES component we use? To answer this question, we calculated the correlations between the effects of different SES components. See **Table 4**. Note that these correlations are often negative, such as between the effect of books at home and the effect of wealth possessions. This means that we get completely different lists if we order societies by the size of the achievement gap between students with few vs. many books at home or by the size of the achievement gap between students with few vs. many wealth possessions.

The Effects of Books at Home and Wealth Vary in Opposite Ways With Development Level

To better understand this discrepancy, we consider the development level of societies, operationalized by the Human Development Index (HDI). We use the latest values available from the United Nations Development Programme (http://hdr.undp.org/) and the Subnational Human Development Database (Smits and Permanyer, 2019). The last row of **Table 4** reports how the effects of the six SES components correlate with the HDI. In **Table 4** we see that the effect of books at home exhibited a very strong positive correlation with HDI whereas the exact opposite held for the effect of wealth possessions. This means that in societies with a lower level of human development, student achievement is less strongly associated with books at home but more strongly associated with wealth possessions. To

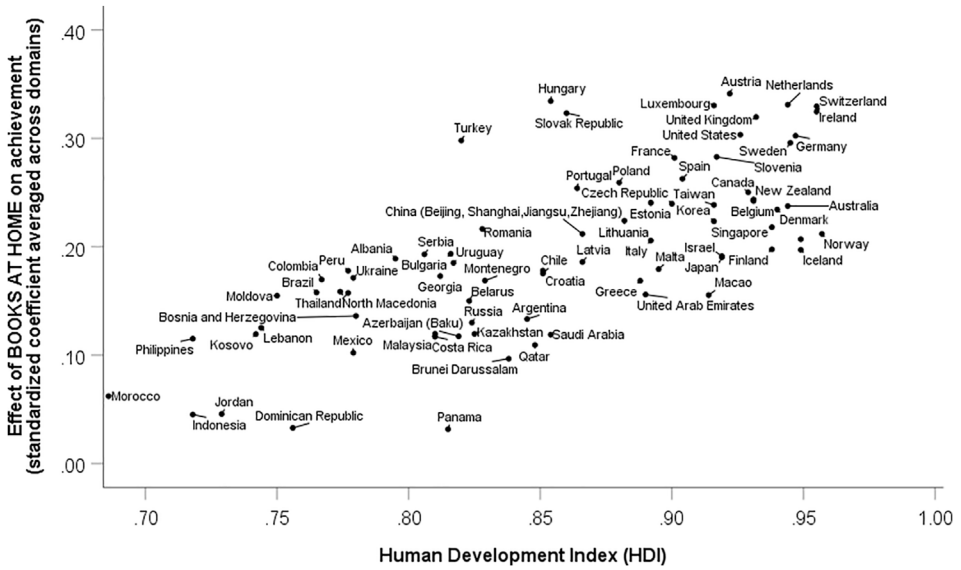


FIGURE 3 | Societies' HDI plotted against the independent effect of books at home on achievement as measured by standardized coefficients averaged across the domains of math, science, and reading.

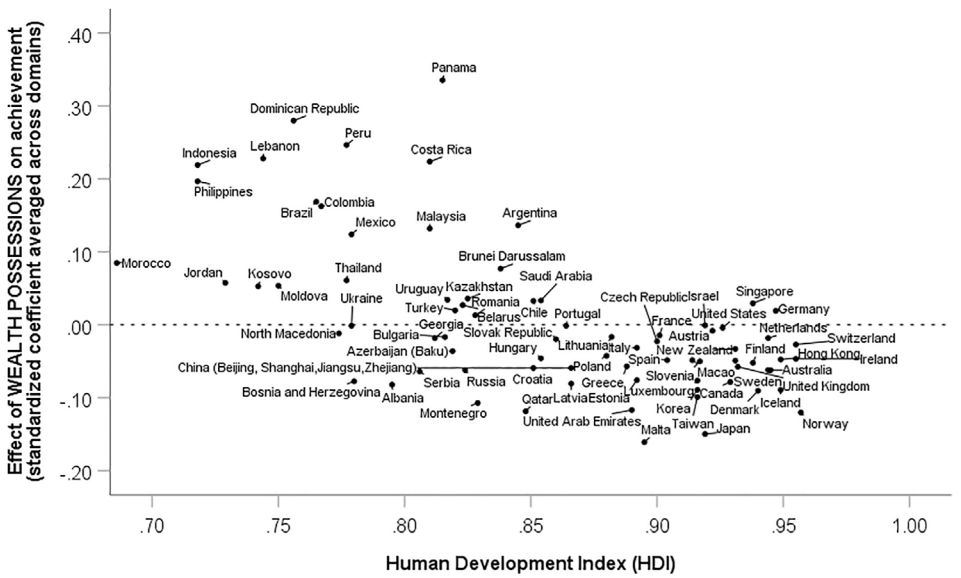


FIGURE 4 | Societies' HDI plotted against the independent effect of wealth possessions on achievement as measured by standardized coefficients averaged across the domains of math, science, and reading.

illustrate these findings, **Figures 3**, **4** present scatterplots of societies' HDI plotted against the effects on achievement (as measured by the average standardized coefficient) of books at home and wealth possessions, respectively. **Figure 3** shows that the effect of books at home on achievement was positive everywhere but it was weaker in low-developed societies. **Figure 4** shows that while the effect of wealth possessions on achievement was negative in most of the societies in this study, it was nonetheless strongly positive in several low-developed societies.

DISCUSSION

An influential recommendation has been that a multidimensional approach to socioeconomic status should be avoided in education

research because it complicates interpretation of results (Cowan et al., 2012). The working assumption behind the current study was that a multidimensional approach can nonetheless be valuable, and that the complexity of the results thus obtained will be a source of insight. We used the six components of PISA's composite measure of socioeconomic status as multiple predictors of achievement. Analyses of data for almost 600,000 students in 77 societies yielded several interesting findings that crucially rely on the multidimensional approach.

First and foremost, we found very clear results with respect to the relative importance of different SES components. In the average society, the single item on the number of books at home was the strongest predictor of achievement, closely followed by the parents' highest occupational status. The other four components (parents' highest educational attainment, home educational resources, cultural possessions, and wealth possessions) tended to contribute little, if at all, to prediction of student achievement. These findings offer a novel way of assessing the validity of theories about the effect of SES on student achievement: Can they account for the primacy of books at home and occupational status? We return to this question below.

Our second finding concerned the size of the SES achievement gap, a topic to which PISA devotes considerable attention. PISA reports achievement gaps estimated using a composite SES measure that is based on the six SES components. Compared to an analysis based on multiple separate components, the use of a composite measure will always underestimate the achievement gap (except when the component weights of the composite measure are chosen to exactly match the multiple regression coefficients). Indeed, by using multiple predictors instead of PISA's composite measure, the proportion of variance in achievement (R squared) explained by SES increased by 40% in the average society and even more in the society with the largest achievement gap. This finding illustrates that estimates of achievement gaps, whether based on composite measures or single measures, are likely to substantially underestimate the total effect of socioeconomic factors on student achievement. This is important to be aware of when interpreting meta-analyses of such estimates (e.g., Sirin, 2005; Harwell et al., 2017). In the literature there appear to be some misunderstandings around this issue, including unwarranted warnings that a use of a single SES component may somehow overestimate the effect of SES on achievement (Sirin, 2005).

To improve estimations of the achievement gap, it is not necessary to use multiple predictors. It is sufficient to improve the composite measure by adjusting the weights of components to reflect their relative importance. We illustrated this by replacing PISA's composite measure by a two-item composite measure based only on the two most important components (number of books at home and parents' highest occupational status). This simple composite measure performed almost as well as the multidimensional approach, especially in societies with large achievement gaps. Compared to using multiple predictors, an advantage of using a single composite measure is that it allows the SES achievement gap to be illustrated in more intuitive ways than as the proportion of variance explained. Following the PISA reports, we illustrated achievement gaps by the mean difference in achievement between the highest and lowest quarter of students on the SES measure. We observed a substantial increase in this gap when it was estimated using the two-item index instead of PISA's composite measure. We conclude that to grasp the extent of the SES achievement gap, the two-item index does a better job. It may also serve the purpose of making researchers aware that books at home and parents' occupational status are especially important predictors of achievement in the average society. Thus, for researchers who require a single measure of SES that is relevant across many countries (though not all, see below), we recommend the two-item index over the index provided by PISA.

Last, we examined how results varied across societies. These analyses showed a large systematic influence of societies' level of human development. Specifically, books at home had a much more positive effect on achievement in the most developed societies than in the least developed societies, where instead wealth possessions had a substantial positive effect on achievement (while having no positive effect at all in the most developed societies). An important conclusion is that the special importance of books at home and parents' occupational status is not universal. Researchers should be aware that cross-societal comparisons of the SES achievement gap may yield completely different results depending on how SES is operationalized. This underscores the value of taking the multidimensional approach to SES and applying it separately to each country. We recommend that researchers take the multidimensional approach using those SES components that are available to them, which ideally would include occupational status, books at home, and economic resources. In this study we have used the SES components available in PISA, which were limited in that there was only a proxy measure of economic resources (wealth possessions) instead of a direct measure such as household disposable income.

Theories About SES Effects: Direct Causation vs. Trait Transfer

So far, we have discussed our findings from a methodological perspective. However, as we mentioned, they also have implications for theories about *why* SES is related to achievement. We shall consider two broad classes of potential mechanisms: direct causation and trait transfer. By direct causation we mean that the parental possessions that SES measures are used to directly benefit children's achievement in school. For example, more wealth allows more monetary input into children's education, while more education may allow more non-financial input, such as quality help with schoolwork (Jerrim and Macmillan, 2015; Rözer and van de Werfhorst, 2019). If direct causation is an important mechanism, student achievement could be raised by giving parents more money and more education. However, direct causation does not seem to account for our finding that, in most countries, the SES effect is not attributable to parents' educational attainment and wealth possessions but mainly to their occupational status and the number of books at home. It is difficult to see how parents' occupational status could directly cause higher achievement. A

direct effect of the number of books at home on reading achievement could arise if children tend to read the books they find at home, but this hypothesis does not account for our finding of an equally strong effect of books at home on achievement in mathematics. For these reasons, our findings suggest that direct causation is not the main reason behind the SES effect. This conclusion is in line with studies of adopted children finding no clear influence of socioeconomic factors among adoptive parents, such as their education, on children's educational attainment (Kendler et al., 2015; Ludeke et al., 2021).

An alternative theory is trait transfer. This theory builds on the combination of two assumptions that are well supported by studies. The first assumption is that achievement in school and achievement of high socioeconomic status (in terms of educational attainment and high-status and high-paying jobs) partly rely on a common set of traits, such as intelligence, self-efficacy, and a conscientious personality (Briley et al., 2014; Krapohl et al., 2014). The second assumption is that these achievement-promoting traits are, to a large extent, genetically transferred from parents to children (Krapohl et al., 2014; Ayorech et al., 2017; Garon-Carrier et al., 2017). Transfer of achievement-promoting traits would readily account for the observed relation between children's achievement in school and parents' occupational status. That the independent effect of parents' educational attainment was much smaller is consistent with educational attainment being a less reliable indicator of achievement-promoting traits, especially in societies where most people get a long education (Chmielewski, 2019).

Parents' wealth possessions (e.g., cars and mobile phones) were found to have a positive relation to children's achievement in school only in countries with low levels of development. To see how trait transfer may account for this finding, consider that high economic development makes wealth possessions affordable for most people (Pokropek et al., 2017) and appears to lead to post-materialist values where wealth possessions are no longer what people strive for (Ahuvia and Wong, 2002). For these reasons, wealth possessions would be a poor indicator of achievement-promoting traits specifically in societies with high economic development. In addition, it is plausible that the direct effects of monetary input and economic hardship are more substantial in societies with low economic development.

A remaining puzzle is why books at home is so strongly related to student achievement, and especially so in countries with high levels of development. Here we speculate that, to the extent that people can afford buying books they desire, the number of books at home indicates parents' general interest in, and enjoyment of, reading. It is plausible that these traits facilitate schoolwork and that they are subject to genetic transfer. This would account for the observed main effect. In poorer countries, those who would like to read books may not afford to buy them, however. This would make the number of books at home a poorer indicator of parents' reading enjoyment, consistent with the decrease of the observed effect of books at home at lower levels of economic development. Moreover, the reported number of books at home may be less accurate in countries with lower development level (Rutkowski and Rutkowski, 2010), which would also contribute to the decrease of the observed effect.

In sum, although the main aim of this research was to contribute to the methodological debate, we believe that our empirical findings also may inform theories about the pathways of the SES effect on achievement. This important issue requires much more research.

CONCLUSION

To conclude, this study took a multidimensional approach to socioeconomic status, thereby revealing a striking variation in the effect of socioeconomic factors on achievement. The effect varies both across different factors and across societies, so that higher development is associated with increased importance of some factors and decreased importance of other factors. These findings have implications for how the SES achievement gap should be measured as well as for how it may be explained.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: osf.io/69tpu/.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

KE conceived of the study, performed the analyses, and wrote the paper. JL, OH, and AR assisted with the survey of the literature and the interpretation of results. All authors read and approved the submitted version.

FUNDING

This work was supported by the Swedish Research Council under Grants 2014-2008 and 2014-2468.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2021.731634/full#supplementary-material

REFERENCES

Ahuvia, A. C., and Wong, N. Y. (2002). Personality and Values Based Materialism: Their Relationship and Origins. *J. Consumer Psychol.* 12 (4), 389–402. doi:10.1016/S1057-7408(16)30089-4

Avvisati, F. (2020). The Measure of Socio-Economic Status in PISA: a Review and Some Suggested Improvements. *Large-scale Assess. Educ.* 8, 1–37. doi:10.1186/s40536-020-00086-x

Ayorech, Z., Krapohl, E., Plomin, R., and von Stumm, S. (2017). Genetic Influence on Intergenerational Educational Attainment. *Psychol. Sci.* 28 (9), 1302–1310. doi:10.1177/0956797617707270

Blau, P. M., and Duncan, O. D. (1967). *The American Occupational Structure*. New York, NY: Wiley.

Blömeke, S., Olsen, R. V., and Suhl, U. (2016). "Relation of Student Achievement to the Quality of Their Teachers and Instructional Quality," in *Teacher Quality, Instructional Quality and Student Outcomes*. Editors T. Nilsen and J. E. Gustafsson (Berlin: Springer), 21–50. doi:10.1007/978-3-319-41252-8_2

Briley, D. A., Domiteaux, M., and Tucker-Drob, E. M. (2014). Achievement-relevant Personality: Relations with the Big Five and Validation of an Efficient Instrument. *Learn. Individ Differ.* 32, 26–39. doi:10.1016/j.lindif.2014.03.010

Chmielewski, A. K. (2019). The Global Increase in the Socioeconomic Achievement gap, 1964 to 2015. *Am. Sociol. Rev.* 84 (3), 517–544. doi:10.1177/0003122419847165

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of Educational Opportunity*. Washington, DC: U.S. Department of Health Education and Welfare, U.S. Government Printing Office.

Cowan, C. R., Hauser, R. M., Kominski, R. A., Levin, H. M., Lucas, S. R., Morgan, S. L., et al. (2012). *Improving the Measurement of Socioeconomic Status for the National Assessment of Educational Progress: A Theoretical Foundation*. Washington, DC: National Center for Education Statistics.

Eriksson, K., Helenius, O., and Ryve, A. (2019). Using TIMSS Items to Evaluate the Effectiveness of Different Instructional Practices. *Instr. Sci.* 47 (1), 1–18. doi:10.1007/s11251-018-9473-1

Ganzeboom, H. B. G. (2010). "A New International Socio-Economic Index [ISEI] of Occupational Status for the International Standard Classification of Occupation 2008 [ISCO-08] Constructed with Data from the ISSP 2002-2007," in In paper presented at annual conference of international social survey programme, Lisbon, Portugal. Retrieved June 2021. Available at: http://www.harryganzeboom.nl/pdf/2010-Ganzeboom-ISEI08-ISSP-Lisbon-%28paper%29.

Garon-Carrier, G., Boivin, M., Kovas, Y., Feng, B., Brendgen, M., Vitaro, F., et al. (2017). Persistent Genetic and Family-wide Environmental Contributions to Early Number Knowledge and Later Achievement in Mathematics. *Psychol. Sci.* 28 (12), 1707–1718. doi:10.1177/0956797617721480

Harwell, M., Maeda, Y., Bishop, K., and Xie, A. (2017). The Surprisingly Modest Relationship between SES and Educational Achievement. *J. Exp. Educ.* 85 (2), 197–214. doi:10.1080/00220973.2015.1123668

Jerrim, J., and Macmillan, L. (2015). Income Inequality, Intergenerational Mobility, and the Great Gatsby Curve: Is Education the Key? *Social Forces* 94 (2), 505–533. doi:10.1093/sf/sov075

Jerrim, J., Volante, L., Klinger, D. A., and Schnepf, S. V. (2019). "Socioeconomic Inequality and Student Outcomes across Education Systems," in *Socioeconomic Inequality and Student Outcomes. Education Policy & Social Inequality*. Editors L. Volante, S. Schnepf, J. Jerrim, and D. Klinger (Singapore: Springer), Vol. 4, 3–16. doi:10.1007/978-981-13-9863-6_1

Kendler, K. S., Turkheimer, E., Ohlsson, H., Sundquist, J., and Sundquist, K. (2015). Family Environment and the Malleability of Cognitive Ability: A Swedish National home-reared and Adopted-Away Cosibling Control Study. *Proc. Natl. Acad. Sci. U S A.* 112 (15), 4612–4617. doi:10.1073/pnas.1417106112

Kim, S. W., Cho, H., and Kim, L. Y. (2019). Socioeconomic Status and Academic Outcomes in Developing Countries: A Meta-Analysis. *Rev. Educ. Res.* 89 (6), 875–916. doi:10.3102/0034654319877155

Krapohl, E., Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J. B., et al. (2014). The High Heritability of Educational Achievement Reflects many Genetically Influenced Traits, Not Just Intelligence. *Proc. Natl. Acad. Sci. U S A.* 111 (42), 15273–15278. doi:10.1073/pnas.1408777111

Lien, N., Friestad, C., and Klepp, K. I. (2001). Adolescents' Proxy Reports of Parents' Socioeconomic Status: How Valid Are They? *J. Epidemiol. Community Health* 55 (10), 731–737. doi:10.1136/jech.55.10.731

Ludeke, S. G., Gensowski, M., Junge, S. Y., Kirkpatrick, R. M., John, O. P., and Andersen, S. C. (2021). Does Parental Education Influence Child Educational Outcomes? A Developmental Analysis in a Full-Population Sample and Adoptee Design. *J. Pers Soc. Psychol.* 120 (4), 1074–1090. doi:10.1037/pspp0000314

Marks, G. N. (2011). Issues in the Conceptualisation and Measurement of Socioeconomic Background: Do Different Measures Generate Different Conclusions? *Soc. Indic Res.* 104 (2), 225–251. doi:10.1007/s11205-010-9741-1

Mueller, C. W., and Parcel, T. L. (1981). Measures of Socioeconomic Status: Alternatives and Recommendations. *Child. Develop.* 52 (1), 13–20. doi:10.2307/1129211

OECD (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*. Paris: PISA, OECD Publishing.

OECD (2019). *PISA 2018 Assessment and Analytical Framework*. Paris: PISA, OECD Publishing.

Pokropek, A., Borgonovi, F., and McCormick, C. (2017). On the Cross-Country Comparability of Indicators of Socioeconomic Resources in PISA. *Appl. Meas. Educ.* 30 (4), 243–258. doi:10.1080/08957347.2017.1353985

Rözer, J. J., and van de Werfhorst, H. G. (2019). "Achievement Inequalities and the Impact of Educational Institutions," in *ISOTIS Report (D 1.4a)* (Amsterdam: University of Amsterdam).

Rutkowski, L., and Rutkowski, D. (2010). Getting it 'better': the Importance of Improving Background Questionnaires in International Large-scale Assessment. *J. Curriculum Stud.* 42 (3), 411–430. doi:10.1080/00220272.2010.487546

Schulz, W. (2005). *Measuring the Socio-Economic Background of Students and its Effect on Achievement in PISA 2000 and PISA 2003*. San Francisco, CA: Paper presented at the annual meeting of the American Educational Research Association.

Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Rev. Educ. Res.* 75 (3), 417–453. doi:10.3102/00346543075003417

Smits, J., and Permanyer, I. (2019). The Subnational Human Development Database. *Sci. Data* 6 (1), 190038. doi:10.1038/sdata.2019.38

Strietholt, R., Gustafsson, J.-E., Hogrebe, N., Rolfe, V., Rosén, M., Steinmann, I., et al. (2019). "The Impact of Education Policies on Socioeconomic Inequality in Student Achievement: A Review of Comparative Studies," in *Socioeconomic Inequality and Student Outcomes: Cross National Trends, Policies and Practices*. Editors L. Volante, S. V. Schnepf, J. Jerrim, and D. A. Klinger (Singapore: Springer), 17–38. doi:10.1007/978-981-13-9863-6_2

White, K. R. (1982). The Relation between Socioeconomic Status and Academic Achievement. *Psychol. Bull.* 91 (3), 461–481. doi:10.1037/0033-2909.91.3.461

Willms, J. D., and Tramonte, L. (2019). "The Measurement and Use of Socioeconomic Status in Educational Research," in *The SAGE Handbook of Comparative Studies in Education*. Editors L. E. Suter, B. Denman, and E. Smith (London: Sage).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 23 December 2021
doi: 10.3389/feduc.2021.793790

The Impact of National and School Contextual Factors on the Academic Performance of Immigrant Students

Holmes Finch[1]*, *Maria E. Hernández Finch*[2] *and Brooke Avery*[2]

[1]Educational Psychology, Ball State University, Muncie, IN, United States, [2]Ball State University, Muncie, IN, United States

The issue of immigration has become central to the politics of nations across the world, impacting many aspects of life over the last decade. Researches investigating educational achievement through a cross-national lens have found that immigrant children tend to exhibit lower academic achievement than their native born peers, and that these differences are exacerbated by both family level variables (e.g., socioeconomic status) as well as the school climate. The goal of the current study was to build on earlier work in this area by investigating the nature and degree to which national attitudes towards immigration have changed over time, and whether any such changes were associated with academic achievement for immigrant and native born students. In particular, the relationship between changing attitudes towards immigration and the achievement gap between native and immigrant students. Results of the study demonstrated that nations with more negative attitudes towards refugees in general, and those for which these attitudes became more negative over time had greater achievement gaps than did those nations with more positive attitudes. In addition, these change trajectories moderated relationships between teacher attitudes towards multiculturalism and academic achievement.

Keywords: PISA, immigration, achievement gap, fragile states index, multicultualism

OPEN ACCESS

Edited by:
Jose Juan Gongora, Tecnológico de Monterrey, Mexico

Reviewed by:
Alberto Crescentini, University of Applied Sciences and Arts of Southern Switzerland (SUPSI), Switzerland
Arturo Orozco-Vargas, Universidad Autónoma del Estado de México, Mexico

***Correspondence:**
Holmes Finch
whfinch@bsu.edu

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

Received: *12 October 2021*
Accepted: *06 December 2021*
Published: *23 December 2021*

Citation:
Finch H, Hernández Finch ME and Avery B (2021) The Impact of National and School Contextual Factors on the Academic Performance of Immigrant Students. Front. Educ. 6:793790. doi: 10.3389/feduc.2021.793790

INTRODUCTION

Immigration has become an increasingly key issue across many nations during the second decade of the 21st century. The movement of individuals from the Middle East and Africa has been associated with political backlash in Europe, the United States, Australia among other regions (The Economist, 2019). Of course, such movements of people are not new to the last decade, as many nations have seen consistent immigration for many years, with some attendant issues associated with acculturation by both the new residents and those who were born within the host countries. A key component in the progress of individuals in the United States, both native born and immigrants, is access to and the ability to take advantage of quality educational opportunities (Pivovarova and Powers, 2019). As is discussed in more detail below, much research has examined the impact of immigration status on the academic performance of immigrants to various nations across the world.

In order to assess the impacts of immigration on both the nation to which individuals move and the nation that they left, a variety of tools have been developed. One of these is a component in the fragile states index (FSI), which is developed and maintained by the Fund for Peace (FFP). The FSI (The Fund for Peace, 2020) measures the overall fragility of nations across the world using a variety of metrics, one of which assesses the status of immigrants. The primary goal of the current study was to

ascertain the extent to which a nation's treatment of immigrants (as reflected in the refugees and internally displaced persons index score) was associated with the academic performance of immigrant children, after both individual and school level contextual effects (e.g., socioeconomic status) were taken into account. Following is a discussion of prior research into the relationship of immigration status and academic performance. The goals of the current study are then outlined, followed by a description of the study methods, the results of the study, and a discussion of these results are presented.

IMMIGRATION STATUS AND ACADEMIC PERFORMANCE

There have been a number of studies examining the academic performance of children from immigrant families, particularly in comparison with their native born peers. This work has shown that generally speaking, immigrant students had lower mean reading, math, and science achievement test scores across 34 Organization for Economic Co-operation and Development (OECD) nations, including Australia, Belgium, Canada, France, Germany, Japan, Mexico, Norway, Portugal, Spain, Sweden, the United Kingdom, and the United States, among others (e.g., Schleicher, 2006; Levels and Dronkers, 2008; Marx and Stanat, 2011; Shapira, 2012; Andon, et al., 2014; Pivovarova and Powers, 2019; Borgonovi and Ferrara, 2020; He and Fischer, 2020). This achievement gap has been identified consistently across a number of nations that have quite different immigrant populations. In addition, thsee gaps were found to hold across academic disciplines, and appear to have been most marked for first generation (born outside the host country) students, when compared to those who were second generation (born in the host country, but whose parents were immigrants).

Researchers have identified a number of potential factors that were associated with the achievement test score gap for immigrant students. For example, in research involving more than 20 nations, including Argentina, Switzerland, Costa Rica, Hong Kong, Slovenia, and Turkey, a number of individual/family level variables were associated with lower achievement test performance for immigrant students, including lower income (Andon, et al., 2014; Giannelli and Rapallini, 2016; Radišić et al., 2021). Likewise, in Germany, Austria, the United Kingdom, Sweden, Canada, and the Czech Republic, a lack of facility in the host nation language (Marks, 2005; Schnepf, 2006; Pivovarova and Powers, 2019). Cultural barriers in Canada, Korea, Finland, Greece, Chile, Estonia, the Netherlands, Spain, and Ireland (Rindermann and Thompson, 2014; Bilican Demir and Yildirim, 2020) along with lower parental education attainment (Schnepf, 2006; Andon, et al., 2014) were associated with lower relative educational test performance on the part of immigrants vis-à-vis native born students.

In addition to the impact of these individual level factors, researchers have also found that school effects are also associated with the academic achievement of immigrant students. For example, Martin, et al. (2012) found that in multiple OECD nations (e.g., Germany, France, Korea, Japan, Russia, the United Kingdom, the United States) the availability of resources for teachers working with immigrant children was associated with the academic performance of their students. Other researchers have found that having a positive learning environment for immigrant students, and positive student attitudes towards their schools were also associated with higher achievement (Schleicher, 2006). Because the Schleicher study did not use an experimental design, it is not possible to determine whether more positive learning environments were causal with respect to higher academic achievement, but the link between the two was found to be relatively large. Rodríguez et al. (2020) examined the extent to which a sense of belonging school was associated with academic achievement for immigrant students in 17 OECD nations, including Hong Kong, Macao, Switzerland, New Zealand, Luxembourg, and the Netherlands. They found that students who felt a sense of welcome and belongingness at their school exhibited higher levels of academic achievement than did those who felt less welcome. When immigrant students were segregated/concentrated in a few schools, the achievement gap vis-à-vis native born students was greater in Portugal, France, Germany, and Canada (Melkonian, et al., 2019; Pivovarova and Powers, 2019).

Researchers have also demonstrated that broader societal attitudes towards immigration, as well as specific government policies were associated with the academic achievement gap between immigrant and native born students. More specifically, immigrant students living in nations where the citizens had more positive attitudes towards immigration had higher levels of academic performance than did immigrants living in nations with less positive attitudes (Rindermann and Thompson, 2014). Likewise, research in a variety of nations including Qatar, Italy, the United Kingdom, the United States, Hungary, Russia, and Serbia, found that students living in nations with more restrictive immigration policies generally performed worse than those living in nations with less restrictive policies (Radišić et al., 2021). Radisic, et al. also demonstrated that nations emphasizing more traditional pedagogical practices exhibited a greater immigrant to native student achievement test performance gap than did those which used a wider array of teaching methods.

STUDY GOALS

There were two primary goals for this study. First, it was of interest to ascertain whether there were different trajectories between 2007 and 2018 in the refugees and internally displaced persons component score across nations for which the fragile states index (FSI) and Programme for International Student Assessment (PISA) were collected. Specifically, the first research question being addressed focused on how these refugee impact scores changed over time, and whether there were subgroups with distinct change trajectories over time. The second goal of this study was to investigate the extent to which membership in subgroups based on change trajectories of the refugee/displaced person index were associated with

student academic performance. In addition, we investigated whether subgroup membership moderated the relationships between selected student and school level variables and academic performance.

Multiple research hypotheses were assessed in this study. First, it was hypothesized that there were different change trajectories in the refugee impact scores over time, with some nations exhibiting increasingly negative attitudes during the time period under study, others exhibiting increasingly positive attitudes, and some nations exhibiting no change over time. In addition, it was hypothesized that the academic performance of immigrant students residing in nations with increasingly negative refugee impacts would be lower than the academic performance of native born students living in these nations. Third, it was hypothesized that the growth trajectory subgroup variable would moderate the relationship between teachers' attitudes toward different cultures and academic achievement. Specifically, we hypothesized that the relationship between teacher attitudes toward different cultures and academic achievement would be positive, but that this relationship would be weaker in nations with negative trajectories in the attitudes towards immigrants. Finally, it was hypothesized that growth trajectory membership would moderate the relationship between the attitudes toward immigration in 2018 and academic achievement in 2018.

METHODS

This project examined data collected through the PISA (OECD, 2018), which is the largest international assessments of students' achievement in mathematics, reading, and science. The data collected from PISA includes demographic information as well as teacher, administrator, and self-reported data. In addition, national level data from the FSI was also used in the study, and is described below. Following is a description of the study participants, the variables used in the analyses, and the analyses used to address the study goals outlined above.

Participants

The sample included a total of 612,004 15-year old students (49.6% female) from the 80 nations participating in PISA (OECD, 2018). More than 20,000 different schools were included in the sample. School samples were selected to be representative within each country. Within each participant nation, all schools registered with the government (both private and public) were included in the sampling frame with the probability of being randomly selected weighted by the number of 15 year old students enrolled, and schools stratified by region of the country. Within each strata, schools were then randomly selected to participate in the PISA project. For each of the selected schools, students were randomly selected from a list of all enrolled 15-year olds. Student samples within each school were chosen using *KeyQuest* software. PISA student data is weighted to be representative of their school and their country, and these weights were used in the current study.

Independent Variables

The independent variables used to address the study goals outlined above included measures of family socioeconomic status (SES), native born status (native), and teacher attitudes towards multiculturalism and equity (SCMCEG) from the PISA 2018 dataset. In addition, the growth trajectory subgroup with respect to treatment of refugees, and the score from this index in 2018 were also included as independent variables in the statistical analyses. These variables are described in more detail below.

PISA Variables

For Family SES, the PISA SES index was derived by OECD (2018) from a factor analysis of variables that include parent education and occupation, home background, as well as possessions in the home (mean = 0, standard deviation = 1). The variable is expressed on a standard normal scale (mean = 0, standard deviation = 1), with higher scores indicating a higher family SES. Native born status was measured by an item on the PISA survey that asked whether students were first generation immigrants, second generation, or more than second generation. The native variable used in this analysis was initially recoded as either more than second generation (1) or first/second generation (0). This coding decision was made in order to differentiate families where both the parents and the children were born in the country of residence (native) from those where at least one member of the family (parents and/or children) were immigrants to the country of residence. As described in the results section, follow up analyses disaggregated this native variable into the original three groups collected by PISA. The SCMCEG variable was based on a scale developed by Hachfeld, et al. (2011) and assessed school leaders' opinions regarding teachers' attitudes towards multiculturalism and equity. The four likert-type items (4 response options) that comprise this scale appear in the **Appendix Table A1**. Higher scores on the SCMCEG score indicated more positive attitudes towards multiculturalism and equity.

Fragile States Index

The attitudes towards refugees and internally displaced persons (RIDP) score from the fragile states index (The Fund for Peace, 2020) was used to assess the pressure on nations that resulted from the inflow and outflow of refugees, as well as movement within nations of internally displaced persons. This variable was designed to measure violence against refugees, sufficiency of resources available to refugees, safety of refugees, and impact of refugee immigration on resource availability for those already residing in the country. The score of each nation included in the study for each year between 2006 and 2018 was used as the variable of interest. Higher RIDP scores (i.e., larger numbers) indicate more negative outcomes (e.g., greater violence, fewer resources, less safety) for refugees. The fragile states index can be accessed at https://fragilestatesindex.org/.

Academics Achievement

Achievement scores for reading and math was used from the individual level PISA database. PISA is not tied to any particular

TABLE 1 | Model fit statistics for GMM solutions.

Model	AIC	BIC	aBIC	BLRT *p*-value
1 class	7341.84	7386.38	7342.04	0 < 0.001
2 classes	7139.38	7209.38	7139.71	0.003
3 classes	7112.09	7194.81	7112.47	0.03
4 classes	7040.33	7135.79	7040.78	0.43
5 classes	7060.55	7161.41	7061.41	0.70

curriculum, but is designed to examine students' higher level thinking skills, such as analysis, synthesis, and evaluation, through the application of knowledge and skills to real-life situations. The 2-h test contains a mixture of multiple-choice and open-ended items. PISA is methodologically complex and state-of-the-art in development (OECD, 2018). PISA proficiency scales were created using Item Response Theory. Research work (Kankaraš and Moors, 2014) has demonstrated that the quality of measurement of the achievement variables was essentially equivalent across nations, which is crucial for the current study, given its emphasis on estimating relationships among variables. For this reason, it is possible to have confidence in the cross-national meaning of the achievement test scores, given that we are not comparing their means, but rather examining relationships among them.

Data Analysis

In order to determine whether there were distinct trajectories in the RIPD scores, a series of growth mixture models (GMM) were fit to the data using a maximum likelihood estimator. Models with from 1 to 5 classes were considered, and for each the AIC, BIC, and sample size adjusted BIC (aBIC) information indices were calculated. In addition, the bootstrap likelihood ratio test (BLRT) test was used to compare the fit of adjacent models (e.g., 1 vs. 2 class, 2 vs. 3 classes, etc.). The null hypothesis being tested by the BLRT was that the statistical fit of the models yielded the same fit to the data. Thus, a rejection of the null would indicate that the fit of the two models to the data differed, and the information indices were used to determine which number of classes yielded the best fit to the data. With regard to the information indices, the model with the lowest value was deemed to provide the best fit to the data after applying a penalty for model complexity. Once the optimal model was identified, the resulting classes were retained for use in the multilevel model described below. The GMM was fit using Mplus, version 8 (Muthén and Muthén, 2020).

Because students were nested within multiple systems, including schools and countries, a 3-level multilevel regression model with level-1 being students, level-2 being schools, and level-3 being countries was fit to the data in order to ascertain the nature of relationships among the independent variables and reading achievement score. It should be noted that the same model was fit to the data treating the mathematics achievement score as the dependent variable, and the results were nearly identical to those for the reading test. Therefore, only the results for the reading test are reported here. The independent variables described above served as the fixed effects variables in the analysis, including national growth trajectory group, national 2018 RIPD score, school SCMCEG score, student native born status, and family SES. A random intercepts model was fit to the data with the random effects being the variances associated with school and nation. Prior to fitting the full model including all of the independent variables, a null model including only the variances associated with nation, school, and random error was first fit to the data. The null model provided an estimate of the intraclass correlation (ICC) for reading test scores. The proportion of variance explained in the outcome variables (R^2), and the proportion of reduction in variance compared to the null model (ICC_Δ; Roberts et al., 2011) were calculated for the full model. These data analyses were conducted using SPSS, version 27 (IBM, 2020), and sampling weights were used.

TABLE 2 | Model parameter estimates of the GMM by latent class.

Term	Coefficient	Standard error
Class 1		
Intercept	7.12[a]	0.39
Linear slope	0.01	0.04
Class 2		
Intercept	5.58[a]	0.70
Linear slope	–0.18[a]	0.05
Class 3		
Intercept	2.88[a]	0.91
Linear slope	0.71[a]	0.11
Class 4		
Intercept	2.74[a]	0.81
Linear slope	–0.05	0.09

[a]*Statistically significant,* $\alpha = 0.05$.

The assumption of normality of the errors was assessed using a QQ-plot and was found to hold, as was homogeneity of variance which was checked using a plot of the residual and predicted values obtained from the model. Given that these assumptions were met, the use of maximum likelihood estimation was deemed to be appropriate. Collinearity among the predictor variables was assessed using the variance inflation factor (VIF), with values in excess of 10 being indicative of collinearity (Fox, 2016). None of the VIF values for this sample exceeded 2, meaning that collinearity was not present. Across the sample, 1.6% of the data were missing. In keeping with recommendations in the literature (e.g., Snijders and Bosker, 2012), missing data was dealt with using full information maximum likelihood (FIML) to obtain model parameter estimates. Grand mean centering was used, per standard practice (Heck and Thomas, 2015).

RESULTS

Hypothesis 1: Differing Change Trajectories in Refugee Impact Scores

The results of the GMM revealed that the 4 class solution yielded the best fit to the data, based on the information indices and the results of the LRT (**Table 1**). More specifically, the 4 class solution had the lowest AIC, BIC, and aBIC values. In addition, this was the first model in the sequence from 1-5 that did not have a statistically significant BLRT result. Recall that the null hypothesis for this test is that there is not a difference in the model fit to the data for the current model (e.g., 4 classes) and the

TABLE 3 | Random effects estimates for null and full models.

Random effect	Estimate	Standard error	95% confidence interval
Null model (school ICC = 0.35, country ICC = 0.17)			
Error	8329.75	18.13	8294.29, 8365.36
School	6062.15	74.14	5918.57, 6209.22
Nation	2839.46	455.54	2073.36, 3888.63
Full Model (School ICC = 0.32, Country ICC = 0.10, Fixed effects R^2 = 0.29)			
Error	7132.23	18.53	7095.99, 7168.65
School	3871.02	58.04	3758.93, 3986.54
Nation	1208.73	235.16	825.52, 1769.81

TABLE 4 | Hypothesis test results for fixed effects in the full model.

Term	DF	F	*p*
Intercept	1	413.73	<0.001
Non-native	1	1737.68	<0.001
SES	1	16959.34	<0.001
SCMCEG (Teacher multiculturalism/equity)	1	189.08	<0.001
FSI immigration 2018	1	0.425	0.52
GMM class	3	3.96	0.01
Non-native × FSI immigration 2018	1	0.08	0.78
Non-native × Teacher equity	1	0.01	0.94
GMM class × Teacher equity	3	33.81	<0.001
GMM class × FSI immigration 2018	3	1.03	0.27
Non-native × GMM class	3	101.91	<0.001
Non-native × GMM class × Teacher equity	6	1.60	0.14

model with one additional class (e.g., 5 classes). The statistically significant result for the 3 class solution indicates that 4 classes fit the data better than 3 classes, whereas the non-significant 4 class solution means that adding a fifth class does not improve the fit. Taking the information indices and hypothesis test results together, further discussion will be focused on this solution.

Table 2 includes the model intercept and coefficient for linear growth for each latent class from the GMM. The intercept is an estimate of the mean RIDP in the first year of analysis (2006), with the coefficient reflecting the mean change from 1 year to the next. Class 1 was characterized by having the largest mean starting value in the immigrants score (least favorable treatment of refugees in 2006) with no change over time. The nations in class 2 were characterized by having the second largest mean RIDP in 2006, and their score declined (improved) over time. On average the scores for the nations in this class declined by 0.18 each year. The nations in class 3 had the second lowest mean starting score (exhibited the second best treatment of refugees on average in 2006), but had an increase in scores of 0.71 each year on average (i.e., treatment of refugees became less positive over time). Finally, class 4 had the lowest mean RIDP in 2006 and exhibited no change in the scores over time. Latent class membership by nation appears in **Appendix Table A2**. Taken together, these results support the first hypothesis that there would be different change trajectories in the refugee impact scores over time.

Multilevel Modeling of Achievement Data

As described in the Methods section, a multilevel model was used in order to investigate the relationship between the GMM latent classes and reading achievement, while controlling for individual and school level variables. **Table 3** includes the random effects estimates for the null model for which there were no predictor variables, and the full model for which all of the predictors described in the Methods section were included. The variances associated with school and country were statistically significant for both models (the confidence intervals did not include 0), indicating that there were differences in mean reading scores across schools within nations, and across the nations themselves. The ICC values in **Table 3** reflect the proportion of variance in the reading test scores that were associated with both school and nation. When no fixed effects were included in the model, school accounted for 35% of the reading score variance and nation accounted for 17%. When the full fixed effects portion of the model was specified, school accounted for 32% of the reading score variance and nation accounted for 10%. The set of fixed effects variables accounted for approximately 29% of the variance in the reading test scores.

TABLE 5 | Mean reading score and standard error by GMM class and non-native status.

Non-native status	GMM class	Mean	Standard error
Non-native	1	396.77	21.12
	2	399.23	10.16
	3	444.84	15.12
	4	449.48	14.37
Native	1	436.59	21.09
	2	425.51	10.09
	3	483.60	15.06
	4	453.05	14.35

Cohen's d *values for mean comparisons between native and non-native examinees. GMM class 1, Native > Non-native, Cohen's* d = *0.41. GMM class 2, Native > Non-native, Cohen's* d = *0.26. GMM class 3, Native > Non-native, Cohen's* d = *0.39. GMM class 4, Native > Non-native, Cohen's* d = *0.03.*

Hypothesis 2: Academic Achievement of Immigrant Students Residing in Nations With Negative Attitudes Towards Immigrants Will Perform Worse Than Native Born Students in Those Nations

Table 4 includes the hypothesis test results for the fixed effects included in the full multilevel model with PISA reading score serving as the dependent variable. The main effects of non-native status, family SES, school mean teacher equity score (SCMCEG), GMM class, as well as the interactions between GMM class and SCMCEG, and GMM class and non-native status were statistically significantly associated with the reading test score. The only statistically significant main effect that was also not involved in a statistically significant interaction was family SES, which had a coefficient of 30.45 (**Table 5**), indicating that examinees from families with a higher SES also had higher scores on the reading test.

The mean reading scores by non-native status and country GMM class appear in **Table 5** and **Figure 1**. The means between the native and non-native examinees within each latent class were compared using simple contrasts with a Bonferroni correction

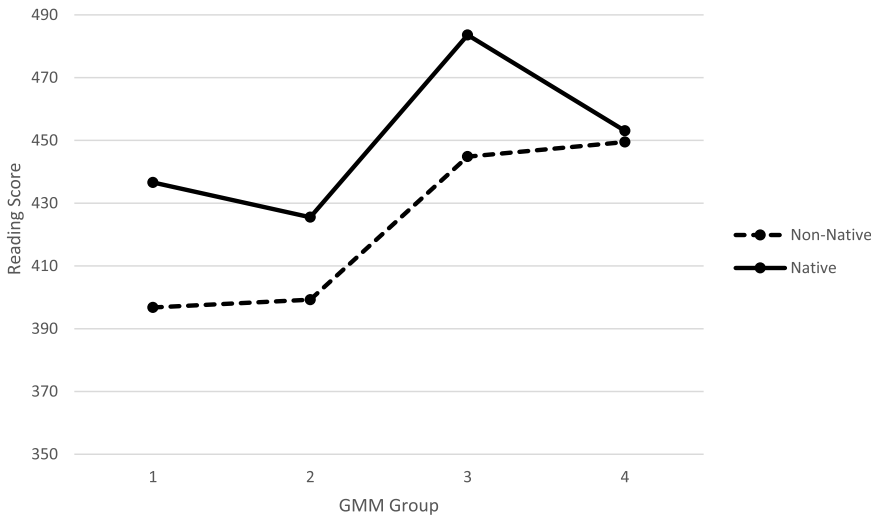


FIGURE 1 | Mean reading score by GMM group and non-native status.

used to control the Type I error rate. The results of these tests revealed that the mean reading scores for native born examinees were significantly higher than those of immigrant students within GMM classes 1, 2, and 3, but not for class 4. The values of Cohen's *d* for the comparison of group means within latent class appear as footnotes in **Table 5**. Based on Cohen's (1988) guidelines for interpreting these values, the difference in mean scores for classes 1–3 fell within the small range, whereas the mean difference for class 4 was negligible in size.

In order to more fully explore the relationships between GMM latent class, native/non-native status, and reading test scores, the multilevel model was refit to the data treating non-native status in its original form as collected by PISA with 3 categories, including first generation, second generation, and native. The results for this model were very similar to those for the model including the dichotomous native/non-native variable, and will therefore not be discussed in more detail here. A set of contrasts comparing the means of the three non-native status categories within GMM latent classes revealed that within each class, the means for the native category were significantly higher than those of second generation students. In turn, the mean reading scores of the second generation students were significantly higher than those for first generation students within each class. The means for the paired comparisons within GMM latent class appear in **Table 6** and **Figure 2**, and the Cohen's *d* values for these comparisons appear in **Table 7**. In addition, the means for the native status groups by GMM latent class appear in **Figure 2**. For latent classes 1 and 3, the Cohen's *d* values for the native vs first generation comparisons fell within Cohen's (1988) large guidelines. Cohen's *d* values for the other mean comparisons in latent classes 1–3 were in the small range, including native vs second generation and second generation vs first generation. In contrast, the Cohen's *d* values were in the negligible range for GMM class 4. These results support the second hypothesis that immigrant students living in nations with more negative attitudes towards immigration will perform worse than native born students living in those nations.

TABLE 6 | Mean reading score and standard error by GMM class and residency status.

Non-native status	GMM class	Mean	Standard error
1st Generation	1	378.03	20.98
	2	378.67	10.59
	3	419.66	15.33
	4	445.96	14.80
2nd Generation	1	398.14	21.77
	2	401.83	10.07
	3	448.28	15.31
	4	453.05	14.08
Native	1	436.59	21.09
	2	425.51	10.09
	3	483.60	15.06
	4	455.25	14.35

Cohen's d *values for mean comparisons between native and non-native examinees.*

Hypothesis 3: The Relationship Between Teacher Attitudes Toward Different Cultures and Academic Achievement Will Be Moderated by National Attitudes Toward Immigration

The coefficients for the teacher equity score by GMM latent class appear in **Table 8**. Across latent classes, students attending schools where teachers had a more positive attitude towards cultural diversity and equity (higher SCMCEG scores) also had higher reading test scores. These coefficients were compared between latent classes using *z*-tests with the type I error rate controlled using the Bonferroni procedure. These results revealed that the relationship between teacher equity and the reading test score for class 3 (5.74) was significantly smaller than for classes 1 (6.61), 2 (6.84), and 4 (7.96). In addition, the coefficient for class 4

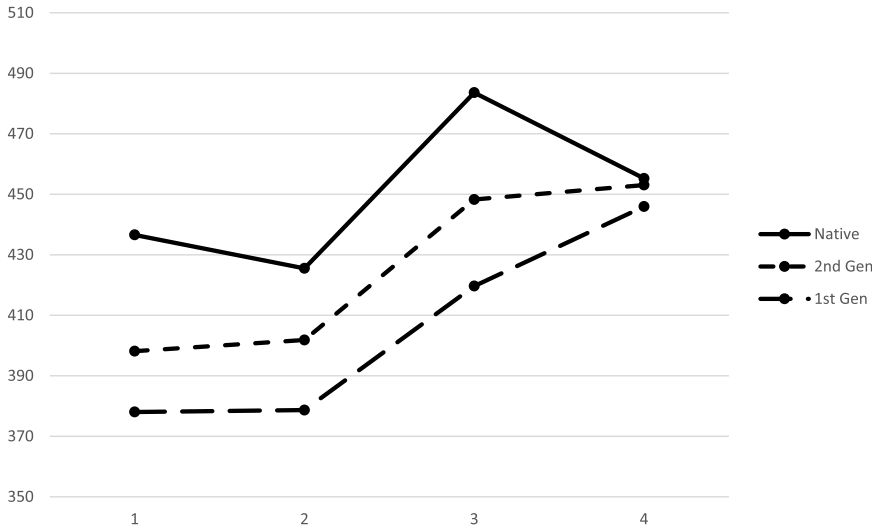


FIGURE 2 | Mean reading score by GMM group and residency status.

TABLE 7 | Mean contrast results and Cohen's *d* effect sizes for non-native status groups.

GMM class	Comparison 1	*d*	Comparison 2	*d*	Comparison 3	*d*
Class 1	Native v 2nd gen	0.36	Native v 1st gen	0.55	2nd gen v 1st gen	0.19
Class 2	Native v 2nd gen	0.25	Native v 1st gen	0.45	2nd gen v 1st gen	0.22
Class 3	Native v 2nd gen	0.37	Native v 1st gen	0.64	2nd gen v 1st gen	0.29
Class 4	Native v 2nd gen	0.02	Native v 1st gen	0.09	2nd gen v 1st gen	0.06

TABLE 8 | Covariate coefficients, standard errors, and 95% confidence intervals for highest level statistically significant effects.

Term	Estimate[a]	Standard error	95% confidence interval
Intercept	478.63[a]	8.92	460.75, 496.51
SES	30.45[a]	0.16	30.13, 30.77
SCMCEG: GMM class 1	6.61[a]	0.33	5.96, 7.26
SCMCEG: GMM class 2	6.84[a]	0.32	6.22, 7.45
SCMCEG: GMM class 3	5.74[a]	0.53	4.70, 6.79
SCMCEG: GMM class 4	7.96[a]	0.28	7.40, 8.52

[a]*Statistically significant,* $\alpha = 0.05$.

was significantly larger than those for classes 1 and 2, which did not differ from one another. Thus, the relationship between teacher attitudes towards multiculturalism and equity were more strongly positively related to reading test score for those in class 4 than in the other classes, whereas the relationship was weakest for those in class 3. This result supports hypothesis 3.

Hypothesis 4: The Growth Trajectory Membership Will Moderate the Relationship Between Attitudes Towards Immigration and Academic Achievement

The fourth hypothesis was that the change in attitudes toward immigration over time would moderate the relationship between attitudes in 2018 and academic achievement in that year. The statistically non-significant interaction between GMM class and FSI 2018 immigration score (**Table 4**) indicates that such moderation was not found to be present in the current study. Therefore, hypothesis 4 was not supported in this study.

DISCUSSION

The results of the analyses presented above revealed the existence of four distinct trajectories in attitudes and treatment of refugees across nations participating in the PISA testing program. These groups differed both in terms of the mean starting RIDP score, as well as the change in these scores over time. Two of the groups were characterized by relatively higher RIDP values (worse treatment of refugees), with nations in one experiencing

improvements over time (Latent class 2) and the other having no change (Latent class 1). The other two latent classes had lower mean starting scores (relatively better treatment of refugees), with one having no mean change in the scores (Latent class 4) and nations in the other having their RIDP increase in value (increasingly worse treatment of refugees between 2006 and 2018; Latent class 3). These results support the first hypothesis that was assessed by this study, namely that there are different subgroups in the population of nations with respect to how attitudes and treatment of refugees evolved over time.

The second research hypothesis investigated by this work focused on whether for nations where attitudes towards refugees became worse over time, non-native born students would exhibit worse academic performance than native students. This hypothesis was partially supported by the results presented above. It was true that the reading means for non-native examinees were significantly lower than for native born examinees in latent class 3, which experienced a significant decline in the treatment of refugees. However, such differences were also found for latent classes 1 and 2, both of which had high scores but either no change, or improvement in RIDP scores. Indeed, only for latent class 4 was there not a statistically significant difference in reading score means between the native and non-native groups. Taken together, these results suggest that for nations with relatively positive treatment/attitudes regarding refugees and for which this did not change between 2006 and 2018 there was no difference in reading achievement test performance for native and non-native examinees, but for all other nations native born individuals tended to perform better on the reading assessment. This finding supports results from earlier studies that demonstrated a relationship between societal and school treatment of immigrants and their performance on academic tasks (Marks, 2005; Rindermann and Thompson, 2014; Radišić et al., 2021).

The third hypothesis, that the latent class would moderate the relationship between teacher attitudes towards multiculturalism and equity (SCMCEG) and academic achievement was supported. The relationship between SCMCEG and reading achievement was most strongly positive for examinees living in latent class 4, and weakest for those in latent class 3. Thus, student attendance at schools where teachers valued multiculturalism more was associated with better reading achievement, and this relationship was strongest in nations with more positive and stable treatment/attitudes towards refugees. Interestingly, the weakest relationship between reading scores and SCMCEG was found for the nations that exhibited the best treatment of refugees in 2006, but which saw the greatest diminution of this support over time. Those nations that exhibited the most negative treatment of refugees in 2006, but that either improved or remained the same in this regard also had a stronger positive relationship between teacher attitudes towards multiculturalism and equity than in nations with more positive refugee treatment, but where that treatment degraded over time. Prior research (Schleicher, 2006; Martin, et al., 2012; Melkonian, et al., 2019; Pivovarova and Powers, 2019; Rodríguez et al., 2020) has shown that school level factors such as a positive learning environment and resource availability are associated with the academic performance of immigrant students. The current study furthers this work by showing that these school level effects on immigrant's academic performance is itself influenced by the broader cultural attitude toward and treatment of immigrants.

Finally, the fourth hypothesis, that latent class would moderate the relationship between national attitudes toward refugees in 2018 (the year of the PISA test) and reading achievement was not supported. Indeed, there was not a relationship between the single year RIDP score and reading achievement across the latent classes, in addition to there not being a statistically significant interaction between latent class and the 2018 RIDP score.

The results of this study both support and amplify earlier findings that have been reported in the literature. Prior work has demonstrated that there is an achievement gap between immigrant and native born students (e.g., Andon, et al., 2014; Borgonovi and Ferrara, 2020). This study also found evidence of such an achievement gap for many nations, though not all. Specifically, these results suggest that in nations that exhibit generally positive treatment of refugees consistently over multiple years, the gap does not exist. In contrast, in nations where this treatment has traditionally been less positive, or where it began positive and then deteriorated over time, the native versus immigrant achievement gap was present. As noted above, these results supported earlier work showing that a positive school climate (in the form of teacher attitude toward multiculturalism in the current study) was associated with greater academic achievement (Schleicher, 2006; Rodríguez et al., 2020). However, this effect was lower for schools in countries with more negative and/or declining attitudes towards refugees. Finally, it should be noted that these results regarding the relationship of national level measures of refugee treatment and academic achievement were found to be present above and beyond family SES, which has consistently been related to academic achievement, as was the case here.

Limitations and Directions for Future Research

The current study was designed to ascertain whether there existed distinct subgroups within the population of nations with respect to the course of change in treatment of refugees and displaced persons over time, and whether these groupings were associated with academic achievement. The results presented above serve, we believe, to extend on earlier work investigating the achievement gap between immigrant and native born students. As with all research efforts, the current study has limitations that future research should be conducted to address. First, although this study was designed to ascertain the role of national level treatment of refugees on academic achievement, it would be of interest to learn the extent to which these policies impact individual schools. Future research should be conducted to investigate how national policies impact school practice, and in turn how these practices specifically impact student achievement. Future research should also examine whether the results presented here also carry over to other national level testing programs such as TIMMS and PIRLS. In order to

investigate the issue from a more nuanced perspective, results for specific nations could also be studied. This study treated all immigrants within a nation as being monolithic. However, it is clear that this is not truly the case. Therefore, future work should more closely consider immigration and home language status for specific nations to determine whether students from different cultural and language backgrounds exhibit divergent academic achievement for nations in the various latent classes. Finally, it is very possible that there exist relationships between the structures and practice of national educational systems, changes in national attitudes towards immigration, and academic performance of immigrants. Therefore, future research should more carefully examine the relationships among these factors in an attempt to identify what aspects of educational systems are associated with the attitudes towards immigration change trajectories identified by the mixture model.

CONCLUSION

Results of this study highlight the relationship between broad national trends in the treatment of and attitudes toward refugees, and academic achievement for non-native students. Specifically, in nations where the treatment of refugees degraded over time, there was a wider gulf between the reading performance of native and non-native students than was the case for nations where this treatment did not change and was relatively positive. In addition, non-native students also exhibited lower average reading scores than their native peers in nations where the treatment of refugees was relatively negative and stable, or even slightly improved over time. In short, how refugees are treated by the nations in which they reside, particularly if that treatment becomes worse over time, appears to be associated with greater gaps in the academic achievement of native and non-native born students. This trend was even true for those who were born in the country of residence but whose parents were not. Therefore, education policy makers should carefully consider how their nation's policies towards refugees may deleteriously impact first and second generation students, and what mechanisms could be put into place to mitigate these negative effects. In contrast, non-native students living in nations for which treatment of refugees was stable and positive appear to perform comparably to their native born classmates.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.oecd.org/pisa/data/2018database/.

AUTHOR CONTRIBUTIONS

HF ran statistical analysis and did the majority of writing. MH developed core research ideas, identified data sources, and wrote portions of the manuscript. BA did literature search and formatted references.

REFERENCES

Andon, A., Thompson, C. G., and Becker, B. J. (2014). A Quantitative Synthesis of the Immigrant Achievement Gap across OECD Countries. *Large-Scale Assessments Educ.* 2 (1). doi:10.1186/s40536-014-0007-2

Bilican Demir, S., and Yildirim, O. (2020). Indirect Effect of Economic, Social, and Cultural Status on Immigrant Students' Science Performance through Science Dispositions: A Multilevel Analysis. *Educ. Urban Soc.* 53 (3), 336–356. doi:10.1177/0013124520928602

Borgonovi, F., and Ferrara, A. (2020). Academic Achievement and Sense of Belonging Among Non-native-speaking Immigrant Students: The Role of Linguistic Distance. *Learn. Individ. Differ.* 81, 101911. doi:10.1016/j.lindif.2020.101911

Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models*. Thousand Oaks, CA: Sage.

Giannelli, G. C., and Rapallini, C. (2016). Immigrant Student Performance in Math: Does it Matter where You Come from? *Econ. Educ. Rev.* 52, 291–304. doi:10.1016/j.econedurev.2016.03.006

Hachfeld, A., Hahn, A., Schroeder, S., Anders, Y., Stanat, P., and Kunter, M. (2011). Assessing Teachers' Multicultural and Egalitarian Beliefs: The Teacher Cultural Beliefs Scale. *Teach. Teach. Educ.* 27 (6), 986–996. doi:10.1016/j.tate.2011.04.006

He, J., and Fischer, J. (2020). Differential Associations of School Practices with Achievement and Sense of Belonging of Immigrant and Non-immigrant Students. *J. Appl. Dev. Psychol.* 66, 101089. doi:10.1016/j.appdev.2019.101089

Heck, R. H., and Thomas, S. L. (2015). *An Introduction to Multilevel Modeling Techniques: MLM and SEM Approaches Using Mplus*. 3rd ed. Boca Raton, FL: Routledge/Taylor & Francis Group.

IBM (2020). *IBM SPSS Statistics for Windows*. Version 27.0. Armonk, NY: IBM Corp.

Kankaraš, M., and Moors, G. (2014). Analysis of Cross-Cultural Comparability of PISA 2009 Scores. *J. Cross-Cultural Psychol.* 45 (3), 381–399. doi:10.1177/0022022113511297

Levels, M., and Dronkers, J. (2008). Educational Performance of Native and Immigrant Children from Various Countries of Origin. *Ethnic Racial Stud.* 31 (8), 1404–1425. doi:10.1080/01419870701682238

Marks, G. N. (2005). Accounting for Immigrant Non-immigrant Differences in reading and Mathematics in Twenty Countries. *Ethnic Racial Stud.* 28 (5), 925–946. doi:10.1080/01419870500158943

Martin, A. J., Liem, G. A. D., Mok, M. M. C., and Xu, J. (2012). Problem Solving and Immigrant Student Mathematics and Science Achievement: Multination Findings from the Programme for International Student Assessment (PISA). *J. Educ. Psychol.* 104 (4), 1054–1073. doi:10.1037/a0029152

Marx, A. E., and Stanat, P. (2011). Reading Comprehension of Immigrant Students in Germany: Research Evidence on Determinants and Target Points for Intervention. *Read. Writ.* 25 (8), 1929–1945. doi:10.1007/s11145-011-9307-x

Melkonian, M., Areepattamannil, S., Menano, L., and Fildago, P. (2019). Examining Acculturation Orientations and Perceived Cultural Distance Among Immigrant Adolescents in Portugal: Links to Performance in reading, Mathematics, and Science. *Soc. Psychol. Educ.* 22 (4), 969–989. doi:10.1007/s11218-019-09506-5

Muthén, L. K., and Muthén, B. O. (2020). *Mplus User's Guide*. Eigth Edition. Los Angeles, CA: Muthén & Muthén.

OECD (2018). *PISA 2018 Results: What Students Know and Can Do*. Available: https://www.oecd.org/pisa/publications/pisa-2018-results-volume-i-5f07c754-en.htm (Accessed September 17, 2021).

Pivovarova, M., and Powers, J. M. (2019). Generational Status, Immigrant Concentration and Academic Achievement: Comparing First and Second-Generation Immigrants with Third- Plus Generation Students. *Large-Scale Assessments Educ.* 7 (1). doi:10.1186/s40536-019-0075-4

Radišić, J., Selleri, P., Carugati, F., and Baucal, A. (2021). Are Students in Italy Really Disinterested in Science? A Person-Centered Approach Using the PISA 2015 Data. *Sci. Educ.* 105 (2), 438–468. doi:10.1002/sce.21611

Rindermann, H., and Thompson, J. (2016). The Cognitive Competences of Immigrant and Native Students across the World: An Analysis of Gaps, Possible Causes and Impact. *J. Biosoc. Sci.* 48 (1), 66–93. doi:10.1017/S0021932014000480

Roberts, J. K., Monaco, J. P., Stovall, H., and Foster, V. (2011). "Explained Variance in Multilevel Models," in *Handbook for Advanced Multilevel Analysis*. Editors J. J. Hox and J. K. Roberts, 219–230.

Rodríguez, S., Valle, A., Gironelli, L. M., Guerrero, E., Regueiro, B., and Estévez, I. (2020). Performance and Well-Being of Native and Immigrant Students. Comparative Analysis Based on Pisa 2018. *J. Adolesc.* 85, 96–105. doi:10.1016/j.adolescence.2020.10.001

Schleicher, A. (2006). Where Immigrant Students Succeed: A Comparative Review of Performance and Engagement in Pisa 20031. *Intercultural Educ.* 17 (5), 507–516. doi:10.1080/14675980601063900

Schnepf, S. V. (2006). Immigrants' Educational Disadvantage: an Examination across Ten Countries and Three Surveys. *J. Popul. Econ.* 20 (3), 527–545. doi:10.1007/s00148-006-0102-y

Shapira, M. (2012). An Exploration of Differences in Mathematics Attainment Among Immigrant Pupils in 18 OECD Countries. *Eur. Educ. Res. J.* 11 (1), 68–95. doi:10.2304/eerj.2012.11.1.68

Snijders, T. A. B., and Bosker, R. J. (2012). *Multilevel Analysis*. Thousand Oaks, CA: Sage Publishers.

The Economist (2019). *Why the Arguments against Are So Popular*. London, UK: Economist.

The Fund for Peace (2020). *Fragile States Index*. Available at: https://fragilestatesindex.org/ (Accessed September 17, 2021).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

TABLE A1 | SCMCEG items.

1. It is important for students to learn that people from other cultures can have different values.
2. Respecting other cultures is something that students should learn as early as possible.
3. In the classroom, it is important that students of different origins recognise the similarities that exist between them.
4. When there are conflicts between students of different origins, they should be encouraged to resolve the argument by finding common ground.

TABLE A2 | Nation by latent class.

Class	Nations
1	Azerbaijan, Bosnia, Colombia, Croatia, Georgia, Israel/West Bank, Jordan, Lebanon, Morocco, Philippines, Serbia, Thailand, Turkey
2	Bulgaria, China, Costa Rica, Dominican Republic, Indonesia, Malaysia, Malta, Mexico, Moldova, Montenegro, Peru, Russia, Saudi Arabia, Vietnam
3	Albania, Austria, Belarus, Brazil, Germany, Greece, Hungary, Italy, Japan, Macedonia, Poland, Romania, Slovenia, Sweden, Ukraine
4	Argentina, Australia, Belgium, Brunei, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Iceland, Ireland, Kazakhstan, Latvia, Lithuania, Luxembourg, Netherlands, New Zealand, Norway, Panama, Portugal, Qatar, Singapore, Slovak Republic, South Korea, Spain, Switzerland, UAE, United Kingdom, United States

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 18 May 2022
doi: 10.3389/fpsyg.2022.876485

COVID-19 Pandemic and Student Reading Achievement: Findings From a School Panel Study

Ulrich Ludewig[1]†, Ruben Kleinkorres[1†], Rahim Schaufelberger[1†], Theresa Schlitter[1†], Ramona Lorenz[1†], Christoph König[2†], Andreas Frey[2†] and Nele McElvany[1†]*

[1]Center for Research on Education and School Development (IFS), TU Dortmund University, Dortmund, Germany, [2]Department of Educational Psychology, Goethe University Frankfurt, Frankfurt, Germany

OPEN ACCESS

Edited by:
Sebastian Urquijo,
CONICET Institute of Basic, Applied and Technology Psychology (IPSIBAT), Argentina

Reviewed by:
Sebastian Paul Suggate,
University of Regensburg, Germany
Elena Florit,
University of Verona, Italy

***Correspondence:**
Ulrich Ludewig
ulrich.ludewig@tu-dortmund.de

†ORCID:
Ulrich Ludewig
orcid.org/0000-0001-9614-847X
Ruben Kleinkorres
orcid.org/0000-0002-9689-8796
Rahim Schaufelberger
orcid.org/0000-0003-4489-6761
Theresa Schlitter
orcid.org/0000-0001-5269-0451
Ramona Lorenz
orcid.org/0000-0002-5733-5421
Christoph König
orcid.org/0000-0003-3172-7029
Andreas Frey
orcid.org/0000-0001-5334-9538
Nele McElvany
orcid.org/0000-0001-8649-5523

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: 15 February 2022
Accepted: 01 April 2022
Published: 18 May 2022

Citation:
Ludewig U, Kleinkorres R, Schaufelberger R, Schlitter T, Lorenz R, König C, Frey A and McElvany N (2022) COVID-19 Pandemic and Student Reading Achievement: Findings From a School Panel Study. Front. Psychol. 13:876485. doi: 10.3389/fpsyg.2022.876485

Since 2020, the COVID-19 pandemic had an impact on education worldwide. There is increased discussion of possible negative effects on students' learning outcomes and the need for targeted support. We examined fourth graders' reading achievement based on a school panel study, representative on the student level, with $N = 111$ elementary schools in Germany (total: $N = 4{,}290$ students, age: 9–10 years). The students were tested with the *Progress in International Reading Literacy Study* instruments in 2016 and 2021. The analysis focused on (1) total average differences in reading achievement between 2016 and 2021, (2) average differences controlling for student composition, and (3) changes in achievement gaps between student subgroups (i.e., immigration background, socio-cultural capital, and gender). The methodological approach met international standards for the analysis of large-scale assessments (i.e., multiple multi-level imputation, plausible values, and clustered mixed-effect regression). The results showed a substantial decline in mean reading achievement. The decline corresponds to one-third of a year of learning, even after controlling for changes in student composition. We found no statistically significant changes of achievement gaps between student subgroups, despite numerical tendencies toward a widening of achievement gaps between students with and without immigration background. It is likely that this sharp achievement decline was related to the COVID-19 pandemic. The findings are discussed in terms of further research needs, practical implications for educating current student cohorts, and educational policy decisions regarding actions in crises such as the COVID-19 pandemic.

Keywords: reading comprehension, reading achievement, COVID-19, elementary school, achievement gaps, large-scale assessment

INTRODUCTION

Since the beginning of 2020, the COVID-19 pandemic has led to a substantially new situation for education systems. To contain the spread of the virus that causes COVID-19, schools in many countries around the world have partially or completely closed, learning groups have been rearranged, and students or teachers had to be absent from school for various amounts of time (cf., Woessmann et al., 2020; Meinck et al., 2022). Teachers had to carry out learning

activities without the usual face-to-face lessons, learners had to self-regulate at home, and parents had to support their children's learning more than before. How these learning conditions affected students' achievement is of considerable interest for educational policy, administration, and practice. This is especially true for reading literacy, a key competence that influences students' achievement in other subjects and enables them to participate in society throughout their entire life course. Additionally, there is reason to assume that the COVID-19 pandemic had a differential effect on students. Even within a given education system, certain groups of students might have been affected more severely than others.

In Germany, the sudden shift from face-to-face instruction to more technologically mediated interaction and emergency remote education (ERE) was especially hard. ERE required German schools and teachers to catch up in terms of the digitalization process in education, which had been shown to lag behind other countries in the years prior to the pandemic (cf., Voogt and Roblin, 2012; Eickelmann et al., 2019; Lorenz et al., 2021). Studies have repeatedly shown that teachers lacked pedagogical skills related to technology and that students had problems accessing and using technological devices during the COVID-19 pandemic (e.g., Huber and Helm, 2020; Reimers and Schleicher, 2020; Rožman et al., 2022). Therefore, Germany might have had particular problems in adapting to the pandemic schooling situation.

A variety of recent publications have shown that schools, instruction, and stakeholders—school administrators, teachers, students, and parents—were only partially prepared for a crisis with substantial restrictions on school life such as the COVID-19 pandemic (e.g., Huber et al., 2020). Accordingly, teachers as well as parents subjectively perceived a decline in student learning (Dong et al., 2020a; Rožman et al., 2022). In contrast, some studies based on student reports found (tendentially) positive learning experiences compared to usual instruction, but students pointed out that they felt more uncertain about estimating their learning status (e.g., Huber and Helm, 2020; Rožman et al., 2022). However, there is a lack of country-specific results related to effects of the COVID-19 pandemic on key achievement measures *via* standardized tests. Highly aggregated results show that school closures due to COVID-19 had an effect of about $d = -0.08$ (Hammerstein et al., 2021) and $d = -0.17$ (König and Frey, 2022) on average student achievement across subject areas, grades, and countries. Data for Germany regarding achievement in one domain that is generalizable to a well-defined student population is missing so far.

Elementary school, and fourth grade in particular, is a pivotal moment in students' educational biographies. At this point, reading literacy should be developed to the point where students can acquire further knowledge through reading in all subjects and continue their educational biography through independent learning. Additionally, in most federal states in Germany, after 4 years of compulsory elementary education (Grades 1–4 in age-homogenous classes of 21 students on average; Destatis, 2018), typically starting at age 6, students finish elementary school and go on to secondary schools of different tracks (Lohmar and Eckhardt, 2015). At the end of elementary school, studies before the COVID-19 pandemic repeatedly indicated that disadvantaged student groups exhibit lower reading literacy (e.g., Mullis et al., 2017). The COVID-19 pandemic might pose further risks for successful education, especially for disadvantaged student subgroups.

Taken together, students' achievement level in important areas (e.g., reading) is of special interest after a long period of restrictions related to the COVID-19 pandemic. Additionally, whether achievement differences between student subgroups are currently greater than before is an important research question. To provide reliable comparative information on key competences before and during the COVID-19 pandemic, the present study examined reading achievement among fourth graders in German elementary schools. In this study, samples representative for the student population of all fourth graders in Germany were examined in the same 111 elementary schools in 2016 and 2021. Both samples were tested with the reading achievement tests from the international school achievement comparison study *Progress in International Reading Literacy Study* (PIRLS). We accounted for changes in student composition and investigated achievement means and how achievement gaps have evolved.

READING ACHIEVEMENT

The acquisition of reading literacy is key for further learning in other school subjects and students' subsequent educational and life paths (Savolainen et al., 2008). Reading achievement is a core component of reading literacy, along reading motivation and behavior. In international achievement studies such as PIRLS, reading achievement represents students' ability to extract relevant information from narrative and informational texts and to understand, use, and reflect on written texts in areas of life that are relevant to the individual and required by society (Mullis et al., 2015). Reading achievement involves multiple levels of text comprehension: surface structure, text base, situation model, rhetorical structure, and pragmatic communication (Kintsch, 1988; Graesser and McNamara, 2011). Mastering text comprehension requires sufficient word recognition (e.g., decoding skills; Wang et al., 2019), language comprehension (e.g., verbal reasoning), and bridging processes (e.g., vocabulary knowledge; see Kim, 2020), as well as active self-regulation, motivation, and engagement (Duke and Cartwright, 2021).

In the first years of schooling, students learn to read at the letter, word, and sentence level in the sense of automating reading and propositional comprehension processes. By the end of fourth grade, which is the end of elementary school in most German federal states, students are expected to comprehend increasingly longer and more complex texts (e.g., Fitzgerald et al., 2015) and to build situation models for age-appropriate texts.

There are important differences concerning comprehension of narrative and informational texts when it comes to different subprocesses (e.g., Ozuru et al., 2009). However, for pragmatic

reasons, many comparative studies report on global reading achievement (e.g., Mo, 2019) that reflects comprehension of narrative and informational texts as well as other genres.

READING AND THE IMPACT OF THE COVID-19 PANDEMIC

Various factors must be considered in ascertaining whether and to what extent reading achievement has been affected by the restrictions related to the pandemic. Students learn to read *via* formal school-based instruction, including homework, and in their leisure time through informal reading activities. The transition from face-to-face instruction in school to ERE because of the COVID-19 restrictions led to less time for formal school-based instruction (Reimers and Schleicher, 2020). In addition, there was less instructional time available in ERE, so that overall students spent less time on learning than they would have in school (Woessmann et al., 2020). In Germany, compared to before the time spent on learning activities dropped by 62% and 42% during the first and second lockdown phases (spring 2020 and autumn/winter 2020/2021), respectively (Woessmann et al., 2020; Werner and Woessmann, 2021). At the same time, students' leisure time behavior partly changed during ERE (Grewenig et al., 2020; Woessmann et al., 2020): the time spent on reading activities, creative work, and exercise stayed on a comparable level during the school closures in Germany (spring 2020: +11%; autumn/winter 2020/2021: −14%). But the time spent on screen-based activities such as watching TV, gaming, social media, and online media increased by a notable 21% (spring 2020) to 34% (autumn/winter 2020/2021). Children from non-college-educated households spent 1 h more on such screen-based activities than children from college-educated households (Woessmann et al., 2020). The reduction in total time spent on formal and informal reading activities and the shift toward more screen-based activities may have affected students' achievement in reading.

Besides these substantial reductions in learning time, reading development could be negatively affected by the reduced effectiveness of instruction during the pandemic. Reading instruction could have been hampered by limited experience with technical equipment necessary for digital instruction and learning during ERE (e.g., Reimers and Schleicher, 2020; Rožman et al., 2022). This problem had been recognized in Germany even before the COVID-19 pandemic (e.g., Lorenz et al., 2021). Compared to other subjects, there are less rigorous curricular frameworks and less readily available exercises, instruction, and materials for reading teachers when reading is done (in part) at a distance (Maldonado and De Witte, 2020). Additionally, fourth graders are confronted with informational texts that involve new challenges, for instance, an increasing amount of instructional pictures (e.g., graphs, maps, and diagrams). This new challenge of cognitively demanding integrated text-picture comprehension might be difficult for teachers to support in distance learning situations (McElvany et al., 2012; Hochpöchler et al., 2013).

Currently, there is no differentiated picture of student achievement, and particularly of elementary school children's reading achievement, during or after the restrictions related to the COVID-19 pandemic compared to before the pandemic. Several publications have already dealt with the effects of the COVID-19 pandemic on students in terms of wellbeing, school achievement, and their interactions (e.g., Hammerstein et al., 2021; Rose et al., 2021; Sánchez Amate et al., 2021). Different approaches were pursued, including a focus on theoretical considerations (e.g., Schneider et al., 2021), teacher surveys (e.g., Reimers and Schleicher, 2020; for Germany: McElvany et al., 2021), and parent surveys (e.g., Reimers and Schleicher, 2020; Steinmayr et al., 2021).

In a first systematic review on student achievement across multiple countries and grades, Hammerstein et al. (2021) focused on the effects of school closures related to COVID-19 on the subjects of math and reading. They reported heterogeneous effect sizes ($d = -0.37$ to $d = 0.25$) across studies, with a small negative effect (median $d = -0.08$) on average. These results for the first lockdown phase were corroborated by two meta-analyses. König and Frey (2022) reported an average impact of $d = -0.12$ of later school closures (after summer 2021) on average student achievement. Storey and Zhang (2021) found an effect of $d = -0.15$ across domains. Furthermore, Zierer (2021) found an average effect of $d = -0.17$ for elementary school students. Among studies examining reading achievement in elementary school children, two studies (Depping et al., 2021; Gore et al., 2021) reported very small positive effect sizes ($d = 0.00$ to $d = 0.04$). In contrast, the four studies finding negative effects on reading achievement reported larger but still small effect sizes (Engzell et al., 2021: $d = -0.09$; Maldonado and De Witte, 2020: $d = -0.29$; Schult et al., 2021: $d = -0.07$; Tomasik et al., 2020: $d = -0.37$). However, it is not yet known how the situation during the COVID-19 pandemic affected reading achievement in elementary school in Germany as a whole.

READING ACHIEVEMENT GAPS

International large-scale assessments of student achievement have repeatedly shown that Germany has some of the most pronounced social disparities (Hußmann et al., 2017; Reiss et al., 2019). There are several theories offering explanations for gaps in achievement related to family background and student variables such as gender (e.g., primary and secondary effects: Boudon, 1974; Grätz and Wiborg, 2020; expectancy-value approaches: Wigfield and Eccles, 2000; Guo et al., 2015; cultural theory: Bourdieu, 1983; and motivation as mediator: Wang and Finch, 2018; Steinmayr et al., 2021). When examining the relationship between family background and reading achievement, studies often refer to socio-cultural capital and the immigration background. Additionally, reading achievement and reading motivation are known to be systematically related to gender (Wigfield et al., 2016).

Family Background

Children with different family backgrounds experience different levels of support from home and their reading socialization varies accordingly. Following the home literacy model (Sénéchal and LeFevre, 2002), such support may involve different literacy experiences, for instance shared reading between parents and children, teaching the alphabet, or reading words. These literacy experiences explain children's growth in reading and vocabulary knowledge (e.g., Becker et al., 2010). Among other factors, these home literacy experiences could explain that the reading achievement of children and adolescents in Germany and many other countries is systematically associated with family background characteristics, such as socio-cultural capital or immigration background (Mullis et al., 2017; for Germany: Wendt and Schwippert, 2017).

Socio-Cultural Capital of the Family

Socio-cultural capital describes the social assets of a person (e.g., intellect and education). More highly educated parents are often able to support their children better and promote their children's reading socialization more comprehensively, due to their own educational experiences and by being educational role models (Dong et al., 2020b). Therefore, higher socio-cultural capital is positively associated to reading achievement.

The number of books at home has become a frequently used indicator to approximate socio-cultural capital in large-scale assessments (e.g., Schwippert, 2019). There are large differences in reading achievement between children from families with different amounts of books at home in many countries (international: Mullis et al., 2017). In Germany, children from families with more than 100 books at home have substantially higher reading achievement on average, than children from families with a maximum of 100 books at home (Hußmann et al., 2017). There are different mechanisms that could explain these differences. (1) More books at home represent an opportunity for children to engage in reading. (2) Parents with more books are more likely to read by themselves, making them positive role models. (3) Furthermore, they are probably able to support their children to a higher degree. (4) The presence of books indicates parents' appreciation for reading and intellectual stimulating activities and (5) is associated with a relatively stable, wealthy and spacious living situation. In sum, the amount of books at home represents a broad indicator for a family background with favorable conditions for becoming a good reader.

Immigration Background

On the one hand, families from immigrant backgrounds often place high value on and strongly promote their children's education, as suggested by the immigration optimism hypothesis (Kao and Tienda, 1995). On the other hand, an immigrant background can also represent a challenge, as it is often confounded with a lower socioeconomic status, a lack of experience with the education system in the host country, and a different family language than the language of instruction, which is associated with children's lower language skills on average (Kristen and Dollmann, 2012; Mullis et al., 2017). Immigrant parents often do not speak the language of instruction as well as native speakers, so their children may not learn the language implicitly to the same extent as their classmates, which could also affect their reading skills. This is supported by the results of PIRLS 2016, where children who always or almost always spoke German at home scored substantially higher on average than children who never or almost never spoke German at home (Wendt and Schwippert, 2017; for an in-depth longitudinal analysis, see Kigel et al., 2015).

Prior to 2021, Germany underwent a number of societal developments that have affected education. One such development is an increase in the number of immigrants coming to Germany. In 2020, about 24 percent of people living in Germany had an immigrant background. Among 5–10 year-old, 38.8 percent of children have a primary or secondary immigration background. This proportion increased by 2.7 percentage points compared to 2019 (Destatis, 2021).

Gender

Several theoretical approaches have attempted to explain gender differences in reading achievement (for an overview of gender differences in reading and language, see Eagly and Wood, 1999; Hyde, 2014). For example, socio-cultural theory explains differences based on societal stereotypes regarding reading and learning activities (Schunk and Zimmerman, 2006). According to social-cognitive learning theory, the gender gap in reading can be explained by girls' better self-regulatory abilities and their higher self-efficacy (cf., Hyde, 2014; McElvany et al., 2017). Additionally, reading achievement is substantially related to reading motivation (Toste et al., 2020). On average, girls have higher reading motivation and read more often in their leisure time (Ainley et al., 2002; Wigfield et al., 2016; Lepper et al., 2021), which promotes their reading achievement. Thus, a wealth of studies indicate that girls have a higher level of reading achievement than boys on average (Logan and Johnston, 2010; Mullis et al., 2017). The PIRLS 2016 results for Germany showed that fourth grade girls scored systematically higher than boys; the achievement gap favoring girls in Germany was about the same as the average achievement gap in the EU and OECD countries overall (McElvany et al., 2017).

READING ACHIEVEMENT GAPS AND THE IMPACT OF THE COVID-19 PANDEMIC

To date, there is no clear evidence on how the restrictions related to COVID-19 influenced reading achievement gaps among elementary school students. It is possible that the COVID-19-related restrictions had differential effects for different subgroups of students and therefore exacerbated educational inequality. Generally, the aforementioned achievement differences related to students' socio-cultural capital, immigrant backgrounds, and gender can be expected to hold for the COVID-19 pandemic period as well. In fact, they may be even more pronounced

because school-based support was difficult during full or partial school closures and children's learning was left in the hands of families to a greater extent than before the pandemic (e.g., Huber and Helm, 2020). For students with lower socio-cultural capital and/or from immigrant backgrounds, the need for greater parental involvement in the learning process might have led to widening achievement gaps. As described above, parents with more socio-cultural capital are more engaged and provide more support for their children's learning (Dong et al., 2020b). Therefore, it seems plausible that children from these families might benefit from spending more time learning with their parents. With respect to immigrant families, if learners speak a language other than the language of instruction at home, they may receive inadequate support in the language of instruction, which is particularly important for reading achievement and might have therefore affected educational outcomes in this domain during or after the COVID-19 pandemic (see Maldonado and De Witte, 2020). ERE was associated with additional costs if families had to purchase technological devices for their children to participate in the digital lessons. This may have further disadvantaged students from low-income families (Eickelmann et al., 2019; Wrase, 2020). Regarding gender, a widening achievement gap might be expected, as female students tend to have higher reading motivation and more frequently read for pleasure than male students (e.g., McElvany et al., 2017; Mullis et al., 2017). A decline in extrinsic school-based reading motivation during the COVID-19 pandemic may have led to these gender differences playing a greater role in reading improvement, which could exacerbate gender achievement gaps in the current cohort of students. Empirical evidence has shown that students' leisure time behavior changed during the COVID-19 pandemic (e.g., Woessmann et al., 2020; Werner and Woessmann, 2021), which could affect the trends in achievement gaps. Students with more highly educated parents spent less time on leisure activities detrimental to learning than their peers and more time on conducive activities (Grewenig et al., 2020; Woessmann et al., 2020). First evidence by Engzell et al. (2021) shows a 40% larger learning loss among students from poorly educated families compared to children from highly educated families in the Netherlands.

CURRENT STUDY AND RESEARCH AIM

The COVID-19 pandemic affected many areas of education, resulting in a need for empirical research how students' learning was affected during this time. First studies indicate negative effects on students' learning outcomes and learning behavior due to the COVID-19 restrictions. More differentiated results on reading achievement among German elementary school students are lacking so far.

The aim of this study is to provide more differentiated results on trends in elementary school students' reading achievement by applying rigorous methodological standards and using data from a school panel study. Differences in reading achievement across different cross-sectional cohorts may be explained by changes in student composition, even when the same schools participate. Thus, the present study also controlled for changes in the student composition within each school. Furthermore, the development of reading achievement gaps during the pandemic was investigated. The students examined in this study are representative for fourth graders in Germany. We used the reading achievement tests from PIRLS 2016.

The research questions and hypotheses investigated are as follows:

1. How does the average reading achievement of fourth grade elementary school students in Germany differ in 2021 compared to before the COVID-19 pandemic in 2016?

   H1: Due to theoretical considerations on the impact of COVID-19-related restrictions on schooling, we expect a decline in average reading achievement from 2016 to 2021.

2. How does the average reading achievement of fourth grade elementary school students in Germany differ in 2021 compared to before the COVID-19 pandemic in 2016 after controlling for student composition?

   H2: We expect a decline in average reading achievement from 2016 to 2021 even when adjusting for student composition.

3. Considering achievement gaps between subgroups of students, (3a) to what extent do differences in reading achievement exist across student subgroups (socio-cultural capital, immigration background, and gender) in 2021 and (3b) how do these gaps differ in 2021 compared to 2016?

   I. There is a gap in average reading achievement to the disadvantage of students with lower socio-cultural capital (H3.1.1) and this gap is larger in 2021 than in 2016 (H3.1.2).
   II. There is a gap in average reading achievement to the disadvantage of students from immigrant backgrounds (H3.2.1) and this gap is larger in 2021 than in 2016 (H3.2.2).
   III. There is a gap in average reading achievement to the disadvantage of boys (H3.3.1) and this gap is larger in 2021 than in 2016 (H3.3.2).

MATERIALS AND METHODS

Participants

The target population for the school panel analyses was the cohort of fourth graders attending a general education German elementary school (i.e., one that does not cater exclusively to special education students) that existed in both 2016 and 2021 (i.e., excluding closed and newly founded schools). The analysis was based on the responses of $N = 2{,}208$ fourth grade students in 2016 and $N = 2{,}082$ fourth grade students in 2021 from a panel of $N = 111$ general education schools (with one class per school participating). All schools participated in PIRLS 2016 and were examined again 5 years later for the school panel

study. Participation in the reading achievement test was mandatory in both years. Students required parental consent to fill out the student background questionnaire. Students with intellectual or physical disabilities (e.g., blindness or deafness) and recently immigrated children with less than 1 year of German instruction were free to participate but were excluded from the data set.

Data collection in 2021 was slightly affected by the COVID-19 pandemic and took place four to 6 weeks later in the school year than in 2016 (May 2 to June 3, 2016, vs. June to July 3, 2021). The absence rate on the test day was slightly higher in 2021 compared to 2016 (6.03% in 2016 vs. 9.01% in 2021). In 2021, at the time of the study, students were required to stay home at the first sign of illness. We will discuss possible consequences for the interpretation of the results later.

Sampling Procedure

PIRLS 2016 followed a two-stage (i.e., sampling first schools and then classes within schools) stratified cluster design (Martin et al., 2017). In 2016, a total of 208 schools were randomly sampled from a complete list of elementary schools in Germany, considering strata regarding school type (e.g., general education vs. special education schools) and the proportion of students from immigrant backgrounds as well as the additional condition that at least one school from each German Federal State had to participate. In 2021, 116 schools were sampled for the panel study as a random sample of the original $N = 208$ schools in PIRLS 2016, considering the strata school type and proportion of children from immigrant backgrounds. For the analysis, we excluded special education schools ($n = 5$) because they are structurally very different from general education schools (i.e., much smaller classes, less bound to state-mandated curricula, and students do not transition to secondary schools after fourth grade). This resulted in a sample of $N = 111$ general education elementary schools.

Weights

The overall weights were calculated to adjust for clustered sampling (i.e., at the school level), the combination of school, class and student weights, as well as non-response adjustment at each level (Martin et al., 2017). On average, each student in our sample from 2016 represented 294 students in the target population for 2016, and each student in our sample from 2021 represented 325 students in the target population for 2021. The 2016 sample represented 648,297 and the 2021 sample 677,762 students.

Instruments

Reading Achievement Test

The reading achievement test used in PIRLS consisted of six narratives and six informational texts and different comprehension tasks developed for them (Mullis et al., 2015; Martin et al., 2017). In 2016, 181 items were administered across 15 different test versions, with each student answering items about two texts. The reading achievement test in 2021 was a subset of 120 items of the test in 2016, spread over eight different booklets. Each student answered 28.31 items on average ($SD = 4.70$) in 2016 and 27.24 items on average ($SD = 4.50$) in 2021. The items were a mixture of multiple-choice (MC) and constructed response (CR) items. The MC items were scored as either correct or incorrect. CR items were rated by trained personnel from the study administration based on scoring rubrics, as either incorrect, partially correct, or completely correct. Omitted items were scored as if they were incorrect responses and not reached items were treated as if they were not administered. The overall scoring procedure was the same in 2016 and 2021. More details on test construction can be found in Martin et al. (2017).

Student Composition Variables

All of the following variables are based on questions asked in both cycles (i.e., 2016 and 2021) with the same phrasing, at a similar location in the questionnaire, to the same group of respondents (i.e., students, teachers, parents, and school administrators). For binary variables, we chose a coding that sets the majority group (>50%) to 0 and the minority group (<50%) to 1, unless indicated otherwise.

Gender

The gender variable was based on administrative data indicating students' gender as reported in official documents. We used contrast coding for gender, because there is no majority group (1 = Male; −1 = Female). A third category (i.e., "Other") was only collected in 2021 and not in 2016 and could therefore not be considered in the analysis.

Age, Enrolment, and Grade Retention

We aimed at comparing same-aged students in 2016 and 2021. Generally, students' age within and across cohorts of fourth graders in Germany is biased by school enrolment deadlines in Germany's federal school system (i.e., the deadlines by which students have to turn 6 years old in order to enroll in first grade in a given year vary from August 5 to September 30 across different federal states). Additionally, the average age of participating students is higher in 2021 due to the fact that the survey period shifted slightly toward later in the school year. Furthermore, individual students' age in fourth grade depends on whether they enrolled in school late or early relative to their birth date, and whether they were held back a grade during elementary school. Generally, being older relative to the rest of a cohort could be a developmental advantage, whereas late enrolment and grade retention are negatively associated with achievement (e.g., Bell et al., 2009). Based on these considerations, we used three variables to control for age-related aspects:

1. Relative cohort age: Students' age within a cohort in a federal state, excluding individual deviations from regular enrolment (i.e., enrolment at age 6) and excluding grade retention. This variable represents a child's age if all federal states had the same enrolment deadline and excludes age shifts

of entire years caused by irregular enrolment and grade retention. This age variabe had a range of 1 year.
2. Enrolment: Individual deviations from regular school enrolment in years (regular enrolment is at age 6; deviations would include enrolment at age 5 or 7).
3. Grade retention: Individual deviations in the number of years of schooling in years (regular is four).

Immigration Background

We chose to define immigration background in three different ways based on the students' responses.

1. The student was not born in Germany (=1) vs. the student was born in Germany (=0).
2. One or both of the students' parents were not born in Germany (three-level factor with both parents born in Germany as the reference group: both parents born in Germany, one parent not born in Germany, and both parents not born in Germany). Place of birth for both the mother and father had to have been filled in; otherwise, the variable was set to missing.
3. The student's family almost never or never speaks German at home (=1) vs. the family almost always or always speaks German at home (=0).

Socio-Cultural Capital

We used students' responses regarding the number of books at home to approximate their cultural capital. The first group included students who reported that their families owned 100 books or less (=1) vs. students who reported that their families possessed more than 100 books (=0).

Special Educational Needs

In Germany, students with special educational needs have been diagnosed by an official institution as having a disability that necessitates special learning support. Specific disorders regarding scholastic skills such as dyslexia do not qualify a student for special educational support. We distinguish students with no special educational needs (=0) from students with diagnosed special educational needs (=1).

Procedure

PIRLS 2016 and the 2021 panel study were administered by the International Association for the Evaluation of Educational Achievement (IEA) in Hamburg. Both studies were conducted entirely on paper and took place during the first half of the school day. The study was administered by trained test administrators in each class, assisted by a teacher known to the class. The test administrators were university students from related disciplines (teacher training, educational science, and psychology) who attended a mandatory workshop on international testing guidelines and the standardized testing manuals.

The testing procedure was structured the same in both cycles. First, students worked on the PIRLS achievement test in two 40-min blocks with a 10-min break in between. During these blocks, students were allowed to ask questions to clarify the instructions but not regarding how to solve the tasks. Second, after another break, students completed several further standardized tests (for cognitive ability, decoding, vocabulary, and sentence comprehension). The cognitive ability test was administered with different variations in the two cycles (e.g., different time constraints), and different instruments were used to assess the reading subprocesses, so we did not use them for the analyses presented here. Lastly, to obtain background information, students completed a questionnaire that took 45 min for PIRLS 2016 and 60 min for the panel study 2021. However, the fact that the questionnaire was longer in 2021 was not relevant to our analysis because all the questions we were interested in (immigration background and socio-cultural capital) were at the beginning of the questionnaire. In total, the study took 138 min in 2016 and 160 min in 2021, mainly because of the longer questionnaire at the end of the study.

Data Analysis

Data preparation and analyses were performed using *R* Studio Version 4.0.3 (R Core Team, 2020). First, we used multi-level imputation to treat missing data in the background variables. Second, we scaled the test data using a multi-group IRT model. Third, plausible values were drawn based on the imputed background variables for conditioning. Fourth, we used linear mixed-effects models to examine our research questions.

Missing Values and Multiple Imputation

We used multiple imputation to address missing values occurring in our data. All student composition variables are based on either administrative data (e.g., age and gender) or students' responses (e.g., books at home and immigration background). For administrative variables, the missing rate was very low, <1%. In 2016, about 10% and in 2021, about 12% of student responses on the background questionnaire were completely missing (i.e., mostly due to missing parental consent). Missing student responses were not systematically clustered within classes.

The multiple imputation was carried out separately for 2016 and 2021 with the same variables and specifications. In addition to student composition, we included parents' reported number of books at home from the parent questionnaire and city size as auxiliary variables. For the imputation, we used a two-level imputation with predictive mean matching at level one for continuous variables (e.g., age). Furthermore, we used predictive mean matching for level two variables (i.e., city size) and logistic regression for binary variables (i.e., immigration background) within the *R* packages *miceadds* (Robitzsch et al., 2017) with 20 iterations and 10 imputed datasets.

Scaling and Plausible Values

Scaling for the reading achievement test was performed using a multi-group generalized partial credit model (Van der Linden, 2016). The model was estimated using the marginal maximum likelihood method (MML) with the *R* package *TAM* (Robitzsch et al., 2019). The model estimates a difficulty and a discrimination parameter for each item or response category. Prior to model

estimation, we excluded two items because fewer than 5% or more than 95% of responses were correct (i.e., leaving 179 items for 2016 and 118 for 2021). The slopes within each CR item with multiple response categories were set to be equal to each other. We used a multi-group approach instead of separate scaling with linking because the achievement tests and test procedures in 2021 and 2016 were very similar. All items had a root mean squared deviation (RMSD) <0.08, so that none of the items indicated large misfit (Köhler et al., 2020). Because the item fit was acceptable for all included items, we considered the multi-group approach to be appropriate. The EAP reliability was good at $REL_{EAP} = 0.87$. For all analyses, we used 10 plausible values to provide a measurement error-adjusted and unbiased estimation of effects. Plausible values were drawn using item parameters anchored at their estimated values from the calibration and random draws from the marginal posterior of the latent distribution for each student (Monseur and Adams, 2009). We used all student composition and auxiliary variables as well as their interaction with the cycle (2016 vs. 2021) for conditioning. We performed five draws with each of the 10 sets of imputed conditioning variables, resulting in 50 data sets. Finally, we used a scale that sets the mean and SD in 2016 to 1,000 and 100, respectively, to make the results of the reading achievement test easier to interpret.

Analysis

Proportions, means, and SDs were calculated with multiple imputed variables, overall student weighting and school clustering using the *R* package *survey* (Lumley, 2020).

Students' reading achievement was statistically modeled using a linear mixed-effects model framework in the *R* package *lme4* (Bates et al., 2014) with the weights for 2016 and 2021. We estimated three models: (1) a gross differences model (i.e., without student composition) to compare the overall difference between the study cycles (2016 vs. 2021) and a (2) net differences model that considered changes in student composition. Additionally, we estimated (3) an achievement gap model that considers possible changes in the achievement gaps.

Models

First, we modeled the reading achievement (θ_{pc}) of a student $p = 1, \ldots, N$ in school $c = 1, \ldots, C$ using a linear mixed-effect model (Bates et al., 2014). In the gross model (GM), reading achievement was modeled as a function of an intercept β_0 (i.e., the average reading achievement in 2016), the fixed effect of the year β_{cycle} (0 = 2016, 1 = 2021), and the random intercept of the school ζ_c [the variance of ζ_c was normally distributed with $\zeta_c \sim N(0, \sigma^2\zeta)$]. Thus, in our GM, β_0 represented the average reading achievement in 2016 and β_{cycle} the difference between 2021 and 2016.

$$GM: \theta_{pc} = \beta_0 + \beta_{cycle} + \zeta_c$$

Second, the net model (NM) included all student composition variables ($\mathbf{X}_{pk}$), $k = 1, \ldots, K$ as fixed effects β_k. In the NM, β_0 represented the expected average reading achievement of the reference group across cycles. The reference group represented the majority groups (born in Germany, both parents born in Germany, speaking German at home, more than 100 books at home, and no special educational needs) with average age and regular enrolment and without grade retention. The regression coefficient β_{cycle} represented the reading achievement difference between the cycles if the students' composition and the fixed effect β_k of the student composition variables were the same in both cycles.

$$\mathrm{NM}: \theta_{pc} = \beta_0 + \beta_{cycle} + \sum_{k=1}^{K} \beta_k X_{pk} + \zeta_c$$

Third, the achievement gap model (AM) included an additional interaction between student composition and cycle. As in the other models, in the AM, β_0 represented the reading achievement of the reference group in 2016. β_{cycle} represents the difference between the reference group in 2016 and 2021. The interaction effect represents the difference in the deviation between the reference group and the student subgroup in 2016 vs. 2021.

$$\mathrm{AM}: \theta_{pc} = \beta_0 + \beta_{cycle} + \sum_{k=1}^{K} \beta_k X_{pk} + \sum_{k=1}^{K} \theta_k X_{pk} * \mathrm{cycle} + \zeta_c$$

RESULTS

Descriptive Statistics

Descriptive statistics for reading achievement are reported in Implications: Research, Support, Educational Policy, **Appendix A**. The student composition changed statistically significantly between 2016 and 2021, with (a) a slightly higher relative cohort age in 2021 due to later test administration dates in 2021 ($t = 14.13$, $p < 0.001$), (b) a higher percentage of children enrolled in school after turning age 6 ($t = 2.59$, $p = 0.009$), (c) a higher percentage of students from immigrant backgrounds in terms of children who were themselves born abroad ($t = 9.28$, $p < 0.001$), both of whose parents were not born in Germany ($t = 3.59$, $p < 0.001$) and who did not speak German at home ($t = 3.59$, $p = 0.006$), and (d) the percentage of students with special educational needs in general education schools ($t = 2.01$, $p = 0.044$). There were no statistically significant differences in grade retention, gender distribution, one parent being born abroad, or number of books at home across the two study cycles in 2016 and 2021 (see details in **Appendix A**).

Does Student Reading Achievement in 2021 Differ From Pre-COVID-19 Times in 2016?

The average reading achievement in 2021 was 980 points. In 2016, fourth graders from the same schools had a mean reading achievement of 1,000 points. The gross model (Model 1) describes the difference in reading achievement between the study cycles without taking into account changes in student composition (see **Table 1**), but including school random

intercepts. The fixed effect for the difference between the study cycles was 19 points ($\beta_{cycle} = -18.93$, $SE = 3.04$, $p < 0.001$) for an average student in an average school. This difference of 19 points was statistically significant and corresponded to a standardized effect size of $d = 0.19$ (note that the SD is 100). The slight deviation from the average score difference (20 points) results from controlling for the random intercept. In conclusion, on average, students' reading achievement was lower in 2021 than in 2016. This result supported our Hypothesis 1.

Does Student Reading Achievement in 2021 Differ From Pre-COVID-19 Times in 2016 When Adjusting for Student Composition?

The net model (Model 2) displays the difference in reading achievement between 2016 and 2021 adjusted for student composition. The net model displayed a significant effect of study cycle $\beta_{cycle} = -13.80$, $SE = 3.03$, $p < 0.001$, indicating that the difference between 2016 and 2021 cannot fully be explained by the student composition variables. The corresponding effect size was $d = 0.14$. This supports H2 that average reading achievement declined from 2016 to 2021 even when adjusting for student composition. The mean expected reading achievement for 2016 given the student composition in 2016 is 1,000 (i.e., mean for 2016), while the mean expected reading achievement for 2021 given the student composition in 2021 is 980 (i.e., mean for 2021). However, we can estimate the expected mean reading achievement for 2021 based on the student composition for 2016. The expected mean reading achievement for 2021 given the student composition for 2016 is 986, and thus, 14 points (i.e., $d = 0.14$) lower than 2016.

In sum, these results indicate that the average reading achievement is lower in 2021 independently of student composition. This supports Hypothesis 2 that average reading achievement declined even when adjusting core characteristics of student composition.

Are There Achievement Gaps Between Subgroups of Students and Did They Change Over Time?

Table 2 shows the estimated subgroup differences in reading achievement, achievement gaps, and changes in achievement gaps. Overall, the results suggest that the achievement gap between students born in Germany and students born in other countries widened from 2016 to 2021. The gap between students with both parents born in Germany and students with both parents born abroad tend to be larger in 2021 than it was in

TABLE 1 | Linear mixed-effect model explaining reading achievement.

	Model 1 gross study cycle difference (GM)		Model 2 net study cycle difference (NM)		Model 3 achievement gap differences (AM)	
	Estimate	***SE***	**Estimate**	***SE***	**Estimate**	***SE***
Intercept	**1001.32**	**4.39**	**1039.49**	**4.25**	**1044.04**	**5.13**
Study cycle (2016 = 0, 2021 = 1)	**−18.93**	**3.04**	**−13.80**	**3.03**	**−22.16**	**6.04**
Gender[a] (male = 1, female = −1)			**−6.03**	**1.46**	**−6.54**	**2.05**
Relative cohort age (years)*			−2.95	5.73	−3.79	7.52
Enrolled (years)*			**−13.19**	**4.27**	**−13.78**	**5.49**
Retention (years)*			**−53.97**	**4.32**	**−50.20**	**6.67**
Child not born in Germany[+]			**−21.09**	**6.60**	−13.44	9.72
One parent not born in Germany[+]			**−17.78**	**4.62**	**−25.16**	**7.54**
Both parents not born in Germany[+]			**−31.51**	**4.78**	**−27.57**	**6.73**
German spoken at home[+,a]			**−10.20**	**5.86**	−13.18	7.06
Number of books at home[+,b]			**−36.16**	**3.73**	**−42.20**	**5.00**
Need for special education[+,c]			**−78.92**	**8.19**	**−88.00**	**16.47**
Year 21 x Gender					0.80	2.92
Year 21 x Age					1.19	11.28
Year 21 x Enrolled					0.68	8.45
Year 21 x Retention					−6.83	9.57
Year 21 x Child not born in Ger.					−10.17	13.76
Year 21 x One parent not born in Ger.					13.57	11.06
Year 21 x Both parents not born in Ger.					−7.70	9.93
Year 21 x German spoken at home					5.87	11.03
Year 21 x Number of books at home					11.25	6.68
Year 21 x Need for special education					14.11	21.43
Explained variance between schools	0.006		0.588		0.598	
Explained variance overall	0.010		0.167		0.168	

*Study 2016 N = 2,208 and 2021 N = 2,082 with each N = 111 schools. *Continuous variable centered. SE, Standard error and Bold estimates: p < 0.05.*

[+]Dichotomous variables with dummy coding (0 vs. 1).

[a]Percentage of children who answered "I always speak German at home" or "almost always speak German at home."

[b]Percentage of children who answered "Enough to fill two bookshelves (101–200)" or more.

[c]Children with an official diagnosis that justifies special educational needs (i.e., emotional disability).

TABLE 2 | Reading achievement gaps in different student subgroups.

Student subgroup	Reading achievement (*SE*)		Achievement Gap	Δ Gap
Gender	Girls	Boys		
2016	1,008 (4.4)	994 (4.9)	−14 (2.2)	2 (3.1)
2021	988 (5.3)	976 (6.5)	−12 (3.8)	
Country of birth (child)	Germany	Other		
2016	1,004 (4.2)	958 (10.1)	−46 (9.2)	−17 (11.4)
2021	991 (5.3)	928 (15.6)	−63 (14.7)	
Country of birth (one parent)	Germany	Other		
2016	1,004 (4.6)	987 (9.0)	−17 (7.8)	13 (11.4)
2021	983 (5.7)	979 (15.0)	−4 (13.8)	
Country of birth (both parents)	Germany	Other		
2016	1,010 (4.2)	971 (7.1)	−39 (5.8)	−16 (8.2)
2021	997 (5.5)	942 (11.4)	−55 (10.0)	
Language at home	German	Not German		
2016	1,008 (4.3)	975 (7.4)	−33 (6.0)	−8 (8.6)
2021	991 (5.6)	951 (12.0)	−41 (10.6)	
Books at home	More then 100	100 or less		
2016	1,034 (5.3)	985 (7.3)	−50 (5.0)	5 (6.8)
2021	1,012 (7.6)	967 (11.4)	−45 (10.6)	

Study 2016 N = 2,208 and 2021 N = 2,082 with each N = 111 schools.

2016. Similarly, the gaps between students who primarily spoke German at home and students who did not primarily speak German at home tended to widen. There was no increase in the gender gap between 2016 and 2021. Lastly, the gap between children with one parent born in another country and children with both parents born in Germany and children with more and less than 100 books seemed to close. However, none of these differences was statistically significant.

The achievement gap model (Model 3) considers differential effects of the student composition variables. The model displays no significant interaction between the year and any of the student composition variables. This suggests that the achievement gaps in the student composition variables did not change significantly between 2016 and 2021. With respect to our hypotheses, we did find a gap between students with different socio-cultural capital, which is in accordance with H3.1.1. However, we did not find a widening gap between 2016 and 2021 (i.e., H3.1.2 was rejected). Furthermore, we found a gap between students from immigrant and non-immigrant backgrounds, which is in accordance with H3.2.1. However, we did not find a widening gap between 2016 and 2021 (i.e., H3.2.2 was rejected). Finally, we found a gender gap in reading achievement, which is in accordance with H3.3.1, but did not find a widening gap from 2016 to 2021 (i.e., H3.3.2 was rejected). In sum, none of the achievement gaps statistically significantly changed between 2016 and 2021.

DISCUSSION

The present work provided first empirical evidence on the status of reading achievement among German fourth graders after the COVID-19-related changes to schooling. Our study makes a cohort comparison of reading achievement among students from 111 elementary schools in Germany before the COVID-19 pandemic in 2016 and more than 1 year after the outbreak of the pandemic in 2021. We adjusted the results for student composition in both study cycles. In sum, there is clear evidence that reading achievement, a core learning outcome, is lower on average among current fourth graders compared to the pre-COVID-19 situation in 2016. The difference between 2016 and 2021 can only partially be explained by student composition. A difference of 19 points is way beyond changes in average reading achievement found in large-scale assessment over the past decades. Thus, it is likely that this decline in average reading achievement is at least partly due to COVID-19-related measures. The observed effects are in the range of the average impact of COVID-19-related school closures as reported in the meta-analysis by König and Frey (2022) ($d = -0.18$).

The observed decline in average reading achievement is remarkable. Baird and Pane (2019) discussed translating standardized effect sizes into years of learning to make them more interpretable. The average annual reading achievement gains in fourth grade are often considered $d = 0.40$ with a margin of error of ±0.06 (Hill et al., 2008). Thus, the decline of $d = -0.19$ means that fourth graders in 2021 are around half a year of learning behind fourth graders in 2016. The decrease of $d = -0.14$ when controlling for student composition would represent slightly more than 4 months of learning. Note that the effect size of annual literacy gains was not measured directly, and average annual literacy gains vary across studies (e.g., $d = 0.29$: Ditton and Krüsken, 2009; $d = 0.48$: Krüsken, 2007), so the half-year or 4-month learning time are not necessarily very precise estimates. Nonetheless, fourth graders in 2021 are substantially behind fourth graders in 2016, even with more conservative estimates. Hence, even though elementary schools implemented a variety of support measures during the COVID-19 pandemic (Huber et al., 2020; Lorenz et al., 2020; Meinck et al., 2022), the results presented here support the concern that younger students were particularly affected by the pandemic schooling situation (see also Tomasik et al., 2020).

Contrary to expectations, we did not find statistically significant effects indicating widening achievement gaps between subgroups of students—here: socio-cultural capital, immigration background, and gender. However, the statistical power for such interaction effects is limited in our study. Our study considered different sources of statistical uncertainty, plausible value variance, sampling variance, and imputation variance, as well as weighting, which imposed a high standard on finding significant changes in achievement gaps. There are recent findings from the German federal state Baden-Württemberg based on an annual population survey suggesting that schools with a large proportion of students with migration background and with lower average socio-cultural capital, respectively, had larger average losses in achievement than other schools (Schult et al., 2022). Therefore, it is likely that studies using larger samples or longitudinal designs can identify significant differences in achievement gaps. Thus, in light of the existing gaps and the

low achievement levels of a substantial share of the student population, targeted support measures are clearly necessary. This finding is in line with previous studies (for Germany: Stanat et al., 2019, internationally: Mullis et al., 2017).

Strengths and Limitations

There is a need for empirical evidence on the academic achievement of current student cohorts in order to understand how these students perform compared to their expected achievement in the absence of the COVID-19 pandemic. Our study is one of the first studies worldwide—and the first of its kind in Germany—to apply a rigorous methodology in order to estimate the actual status of students' reading achievement in elementary schools. The presented analyses are based on a representative sample taking the standardized, well-established PIRLS reading achievement test. In contrast to other comparative studies, we present a school panel analysis. This has the main benefit of holding a number of key variables related to the educational environment, such as general school conditions (e.g., reading curricula) and school location, constant, allowing for a very high degree of comparability. Thus, the instrument and study design enable us to obtain reliable information on developments in achievement over time controlling for student composition as well as evidence on achievement gaps.

However, as a main limitation, it must be stated that no causal inferences on the effect of the containment measures during the COVID-19 pandemic on reading achievement since 2016 can be drawn. The prerequisites for causal inferences are not given. A control group is not available, since the COVID-19-related measures were applied to all schools, and our study is not longitudinal at the student level and therefore cannot control for pre-pandemic individual student characteristics. At least one of these two conditions (as well as a few others) would be necessary to estimate the causal effect of specific pandemic measures such as school closures of different lengths. In addition, there may be a slight underestimation of the full effect, as the measurement date in 2021 was on average 1 month later than in 2016.

Furthermore, we only investigated reading achievement as a comprehensive construct. However, reading is a multi-faceted construct (Graesser and McNamara, 2011) with many contributing subprocesses such as word recognition (e.g., decoding skills), language comprehension (e.g., verbal reasoning), and bridging processes (e.g., vocabulary knowledge) and additionally, active self-regulation, motivation, and engagement (Duke and Cartwright, 2021). All of these subprocesses could be influenced by the COVID-19 pandemic conditions. Further insights into which particular reading subprocesses were especially impaired could help to further improve post-COVID-19 reading interventions. We will have to leave this to further research, as the panel study was not originally designed to allow for these in-depth analyses.

Implications: Research, Support, and Educational Policy

However, the presented findings lead to important conclusions regarding further research, educational practice, and educational policy. Further analyses may provide more in-depth insights. These include differentially considering reading achievement for literary texts compared with informational texts, which may lead to more gender-specific findings, as girls' performance advantages at the end of fourth grade are especially prominent for literary texts (Mullis et al., 2017), and this may have been further reinforced by increased reading for pleasure during the COVID-19 pandemic-related restrictions. In addition, it should be examined whether the results also apply to other domains such as mathematics or to older groups of students. Finally, international comparisons are urgently needed to clarify whether the pattern found for Germany holds for other countries as well. This will be possible in the future using data from internationally comparative school achievement surveys such as PIRLS 2021 (elementary school, to be published in December 2022) and PISA 2022 (secondary schools, assessed in 2022). Similarly, national large-scale assessments of student achievement can also be insightful (Stanat et al., 2019) and could help to refine our findings in the future.

Regarding educational practice, it should be noted that compensatory measures have not been sufficiently effective for elementary school students in Germany more than a year after the onset of the COVID-19 pandemic-related restrictions on school operations but since then comprehensive measures have started to take place in Germany. Indeed, the findings highlight the need for comprehensive support—for all learners, as shown by the overall effect, but also targeted support for specific groups of students, as illustrated by the significant achievement gaps at the end of fourth grade, even if these were not further amplified compared to 2016. Here, coordinated targeted support approaches must be used that focus on systematically support reading skills in the classroom, extracurricular support during students' leisure time, and during school vacations, as well as support from the family. Lastly, we assessed reading achievement shortly before most students in Germany transition to secondary schools. Therefore, the study provides information that could help secondary school teachers better understand the needs of rising fifth graders in post-COVID-19 times.

The findings are also informative for the design of educational policy. It should be concluded that the framework and conditions for learning in crisis situations need to be strengthened. This includes but is not limited to expanding the framework conditions and use of digital media, but also promoting resilience at all levels (i.e., among learners and their families, teachers, schools, and the educational system). Furthermore, self-regulated learning should be fostered among students of all ages, and last but not least, reading skills should be effectively supported at an early stage, as a key competency for all learners that enables them to acquire learning content relatively independently even in extraordinary learning situations such as distance learning.

The aim of the present study was to gain profound insights into the status of students' achievement in the key competence of reading after a long period of COVID-19-related restrictions on learning at school and to identify any necessary support needs. In conclusion, society, as well as educational practice and educational policy more specifically, are now tasked with implementing effective supports for the children and adolescents

affected by the COVID-19 pandemic in order to effectively secure their educational and life chances.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because publication restrictions apply until the end of 2022. When available, the datasets will be available here: https://www.fdz-bildung.de/home.

ETHICS STATEMENT

Ethical review and approval was not required for the current study in accordance with the local legislation and institutional requirements. Written informed consent was not required in accordance with the national legislation and the institutional requirements. The original studies that led to the creation of the dataset were reviewed and approved by the Ministers of Education ("Kultusminister der Länder") of all 16 federal states in Germany, and written informed consent to participate in these studies was provided by the participants' legal guardian/next of kin.

AUTHOR'S NOTE

Content, ethical aspects, and data protection have been thoroughly examined by the responsible (data protection) officers of each of the 16 German federal states.

AUTHOR CONTRIBUTIONS

NM and RL contributed to the conception and design of the study. UL prepared the database, performed the statistical analysis, and wrote the first draft of the method and result section. NM wrote the first draft of introduction and discussion. TS, RS, and RK wrote paragraphs of the manuscript. UL, NM, RL, TS, RS, RK, CK, and AF contributed to manuscript revision, read, and approved the submitted version. All authors contributed to the article and approved the submitted version.

FUNDING

The study was funded by the Federal Ministry of Education and Research (BMBF) and the Standing Conference of the Ministers of Education and Cultural Affairs of the German federal states (KMK).

ACKNOWLEDGMENTS

We would like to thank our intern Carina Jüschke for her support in the preparation of this manuscript and Sebastian Weirich, Matthias Trendtel, and Jakob Schwerter for their advice on methodological implementation. In addition, we would like to thank Keri Hartmann for proofreading of an earlier version of this manuscript.

REFERENCES

Ainley, M., Hillman, K., and Hidi, S. (2002). Gender and interest processes in response to literary texts: situational and individual interest. *Learn. Instr.* 12, 411–428. doi: 10.1016/S0959-4752(01)00008-1

Baird, M. D., and Pane, J. F. (2019). Translating standardized effects of education programs into more interpretable metrics. *Educ. Res.* 48, 217–228. doi: 10.3102/0013189X19848729

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv [Preprint].

Becker, M., McElvany, N., and Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy: a longitudinal study. *J. Educ. Psychol.* 102, 773–785. doi: 10.1037/a0020084

Bell, J. F., Sykes, E. D., and Vidal, C. (2009). *Birthdate Effects: A Review of the Literature From 1990-On*. UK: University of Cambridge.

Boudon, R. (1974). *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. Göttingen, Germany: Wiley.

Bourdieu, P. (1983). "Ökonomisches Kapital, kulturelles Kapital, soziales Kapital [Economic capital, cultural capital, social capital]," in *Soziale Ungleichheiten [Social Unequalities]*. ed. R. Kreckel (North-Holland: Schwartz), 183–198.

Depping, D., Lücken, M., Musekamp, F., and Thonke, F. (2021). "Kompetenzstände hamburger Schüler* innen vor und während der Corona-Pandemie [Competence levels of Hamburg students before and during the Corona pandemic]," in *Schule während der Corona-Pandemie. Neue Ergebnisse Und Überblick über Ein Dynamisches Forschungsfeld [School during the Corona Pandemic. New Results and Overview of a Dynamic Field of Research]*. eds. D. Fickermann and B. Edelstein (Münster, Germany: Waxmann), 51–79.

Destatis (2018). Allgemeinbildende Schulen - Fachserie 11 Reihe 1 – Schuljahr 2016 [General education schools - Section 11 Subsection 1 - Year 2016]. *Statistisches Bundesamt*. Available at: https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00071539

Destatis (2021). Allgemeinbildende Schulen - Fachserie 11 Reihe 1 – Schuljahr 2021 [General education schools - Section 11 Subsection 1 - Year 2021]. *Statistisches Bundesamt*. Available at: https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00071539

Ditton, H., and Krüsken, J. (2009). Denn wer hat, dem wird gegeben werden? Eine Längsschnittstudie zur Entwicklung schulischer Leistungen und den Effekten der sozialen Herkunft in der Grundschulzeit. *J. Educ. Res. Online* 1, 33–61. doi: 10.25656/01:4555

Dong, C., Cao, S., and Li, H. (2020a). Young children's online learning during COVID-19 pandemic: Chinese parents' beliefs and attitudes. *Child Youth Serv. Rev.* 118:105440. doi: 10.1016/j.childyouth.2020.105440

Dong, Y., Wu, S., Dong, W.-Y., and Tang, Y. (2020b). The effects of home literacy environment on children's reading comprehension development: a meta-analysis. *Educ. Sci.: Theory Pract.* 20, 63–82. doi: 10.12738/jestp.2020.2.005

Duke, N. K., and Cartwright, K. B. (2021). The science of reading progresses: communicating advances beyond the simple view of reading. *Read. Res. Q.* 56, 25–44. doi: 10.1002/rrq.411

Eagly, A. H., and Wood, W. (1999). The origins of sex differences in human behavior: evolved dispositions versus social roles. *Am. Psychol.* 54, 408–423. doi: 10.1037/0003-066X.54.6.408

Eickelmann, B., Bos, W., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., et al. (eds.) (2019). ICILS 2018# Deutschland: Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking [ICILS 2018# Germany: Computer and information-related competencies of students in the second international comparison and Computational Thinking competencies]. (Münster: Waxmann Verlag).

Engzell, P., Frey, A., and Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. SocArXiv [Preprint]. doi: 10.31235/osf.io/ve4z7

Fitzgerald, J., Elmore, J., Koons, H., Hiebert, E. H., Bowen, K., Sanford-Moore, E. E., et al. (2015). Important text characteristics for early-grades text complexity. *J. Educ. Psychol.* 107, 4–29. doi: 10.1037/a0037289

Gore, J., Fray, L., Miller, A., Harris, J., and Taggart, W. (2021). The impact of COVID-19 on student learning in New South Wales primary schools: an empirical study. *Aust. Educ. Res.* 48, 605–637. doi: 10.1007/s13384-021-00436-w

Graesser, A. C., and McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Top. Cogn. Sci.* 3, 371–398. doi: 10.1111/j.1756-8765.2010.01081.x

Grätz, M., and Wiborg, Ø. N. (2020). Reinforcing at the top or compensating at the bottom? Family background and academic performance in Germany, Norway, and the United States. *Eur. Sociol. Rev.* 36, 381–394. doi: 10.1093/esr/jcz069

Grewenig, E., Lergetporer, P., Werner, K., Woessmann, L., and Zierow, L. (2020). COVID-19 and Educational Inequality: How School Closures Affect Low- and High-Achieving Students. CESifo Working Paper No. 8648. Munich: CESifo. Available at: https://www.cesifo.org/de/publikationen/2020/working-paper/COVID-19-and-educational-inequality-how-school-closures-affect-low

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S., and Yeung, A. S. (2015). Expectancy-value in mathematics, gender, and socioeconomic background as predictors of achievement and aspirations: a multi-cohort study. *Learn. Individ. Differ.* 37, 161–168. doi: 10.1016/j.lindif.2015.01.008

Hammerstein, S., König, C., Dreisörner, T., and Frey, A. (2021). Effects of COVID-19-related school closures on student achievement: a systematic review. *Front. Psychol.* 12:746289. doi: 10.3389/fpsyg.2021.746289

Hill, C. J., Bloom, H. S., Black, A. R., and Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Dev. Perspect.* 2, 172–177. doi: 10.1111/j.1750-8606.2008.00061.x

Hochpöchler, U., Schnotz, W., Rasch, T., Ullrich, M., Horz, H., McElvany, N., et al. (2013). Dynamics of mental model construction from text and graphics. *Eur. J. Psychol. Educ.* 28, 1105–1126. doi: 10.1007/s10212-012-0156-z

Huber, S. G., Günther, P. S., Schneider, N., Helm, C., Schwander, M., Schneider, J., et al. (2020). *COVID-19 und aktuelle Herausforderungen in Schule und Bildung. Erste Befunde des Schul-barometers in Deutschland, Österreich und der Schweiz [COVID-19 and Current Challenges in Schools and Education. First Findings of the School Barometer in Germany, Austria and Switzerland]* Münster, Germany: Waxmann.

Huber, S. G., and Helm, C. (2020). Lernen in Zeiten der Corona-Pandemie. Die Rolle familiärer Merkmale für das Lernen von Schüler*innen: Befunde vom Schul-Barometer in Deutschland, Österreich und der Schweiz [Learning in Times of the Corona Pandemic. The Role of Family Characteristics in Student Learning: Findings From the School Barometer in Germany, Austria, and Switzerland]. DDS – Die Deutsche Schule, Beiheft 16, 37–60.

Hußmann, A., Stubbe, T. C., and Kasper, D. (2017). "Soziale Herkunft und Lesekompetenzen von Schülerinnen und Schülern [Social background and reading skills of students]," in *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich [IGLU 2016: Reading Competencies of Primary School Children in Germany in International Comparison]*. eds. A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper and E.-M. Lankes et al. (Münster, Germany: Waxmann), 195–217.

Hyde, J. (2014). Gender similarities and differences. *Annu. Rev. Psychol.* 65, 373–398. doi: 10.1146/annurev-psych-010213-115057

Kao, G., and Tienda, M. (1995). Optimism and achievement: the educational performance of immigrant youth. *Soc. Sci. Q.* 76, 1–19.

Kigel, R. M., McElvany, N., and Becker, M. (2015). Effects of immigrant background on text comprehension, vocabulary, and reading motivation: a longitudinal study. *Learn. Instr.* 35, 73–84. doi: 10.1016/j.learninstruc.2014.10.001

Kim, Y. S. G. (2020). Hierarchical and dynamic relations of language and cognitive skills to reading comprehension: testing the direct and indirect effects model of reading (DIER). *J. Educ. Psychol.* 112, 667–684. doi: 10.1037/edu0000407

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychol. Rev.* 95, 163–182. doi: 10.1037/0033-295X.95.2.163

Köhler, C., Robitzsch, A., and Hartig, J. (2020). A bias-corrected RMSD item fit statistic: an evaluation and comparison to alternatives. *J. Educ. Behav. Stat.* 45, 251–273. doi: 10.3102/1076998619890566

König, C., and Frey, A. (2022). The impact of COVID-19-related school closures on student achievement: a meta-analysis. *Educ. Meas. Issues Pract.* 41, 16–22. doi: 10.1111/emip.12495

Kristen, C., and Dollmann, J. (2012). "Migration und Schulerfolg: Zur Erklärung ungleicher Bildungsmuster [Immigration and school success: towards explaining unequal educational patterns]," in *Handbuch Migration Und Bildung [Handbook Immigration and Education]*. ed. M. Matzner (Weinheim, Germany: Beltz), 102–117.

Krüsken, J. (2007). "Entwicklung von Schülerleistungen und Zensuren in der Grundschule," in *Kompetenzaufbau und Laufbahnen im Schulsystem*. ed. H. Ditton (Frankfurt, Germany: Waxmann), 41–62.

Lepper, C., Stang, J., and McElvany, N. (2021). Gender differences in text-based interest: text characteristics as underlying variables. Reading Research Quarterly [Preprint]. doi: 10.1002/rrq.420

Logan, S., and Johnston, R. (2010). Investigating gender differences in reading. *Educ. Rev.* 62, 175–187. doi: 10.1080/00131911003637006

Lohmar, B., and Eckhardt, T. (2015). The Education System in the Federal Republic of Germany 2013/2014: A Description of the Responsibilities, Structures, and Developments in Education Policy for the Exchange of Information in Europe. Available at: https://www.kmk.org/fileadmin/Dateien/pdf/Eurydice/Bildungswesen-engl-pdfs/dossier_en_ebook.pdf

Lorenz, R., Brüggemann, T., and McElvany, N. (2020). Unterricht während der Corona-Pandemie: Lehrkräftebefragung Ergebnisse Teil I "Der Unterricht" [Teaching During the Corona Pandemic: Teacher Survey Results Part I "The Lessons"]. Institut für Schulentwicklungsforschung der Technischen Universität Dortmund. Available at: https://www.tu-dortmund.de/storages/zentraler_bilderpool/user_upload/UCP_Kurzbericht_final.pdf

Lorenz, R., Yotyodying, S., and Eickelmann, B., and Endberg, M. (2021). Schule digital – der Länderindikator 2021. Erste Ergebnisse und Analysen im Bundesländervergleich [Digital School: The State Indicator 2021. First Results and Analyses Comparing the Federal States]. Available at: https://eldorado.tu-dortmund.de/handle/2003/40566

Lumley, T. (2020). Survey: Analysis of Complex Survey Samples. R Package Version 4.0. 1–19. Available at: http://ftp.acc.umu.se/mirror/CRAN/web/packages/survey/survey.pdf

Maldonado, J., and De Witte, K. (2020). The effect of school closures on standardized student test outcomes. *Br. Educ. Res. J.* 48, 49–94. doi: 10.1002/berj.3754

Martin, M., Mullis, I., and Hooper, M. (2017). Methoden und Verfahren in PIRLS 2016. *International Association for the Evaluation of Educational Achievement*. Chestnut Hill, USA: TIMSS & PIRLS International Study Center. Available at: https://files.eric.ed.gov/fulltext/ED580352.pdf

McElvany, N., Kessels, U., Schwabe, F., and Kasper, D. (2017). "Geschlecht und Lesekompetenz [Gender and reading competence]," in *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich IGLU [2016: Reading Competencies of Primary School Children in Germany in International Comparison]*. eds. A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper and E.-M. Lankes et al. (Münster, Germany: Waxmann), 177–194.

McElvany, N., Lepper, C., Lorenz, R., and Brüggemann, T. (2021). "Unterricht während der Corona-Pandemie [Lessons during the corona pandemie]," in *Generation Corona? Wie Jugendliche durch die Pandemie benachteiligt werden [How Are Young People Disadvantaged by Pandemic? How Young People Are Affected by Pandemic]*. eds. D. Dohmen and K. Hurrelmann (Weinheim: Beltz Juventa), 64–79.

McElvany, N., Schroeder, S., Baumert, J., Schnotz, W., Horz, H., and Ullrich, M. (2012). Cognitively demanding learning materials with texts and instructional pictures: teachers' diagnostic skills, pedagogical beliefs, and motivation. *Eur. J. Psychol. Educ.* 27, 403–420. doi: 10.1007/s10212-011-0078-1

Meinck, S., Fraillon, J., and Strietholt, R. (2022). The Impact of the COVID-19 Pandemic on Education. Interational Evidence From the Responses to Educational Disruption Survey (REDS). IEA. Available at: https://www.iea.nl/publications/international-evidence-responses-to-educational-disruption-survey

Mo, J. (2019). "How Does PISA Define and Measure Reading literacy?" PISA in Focus, No. 101, OECD Publishing, Paris.

Monseur, C., and Adams, R. (2009). Plausible values: how to deal with their limitations. *J. Appl. Meas.* 10, 1–15.

Mullis, I., Martin, M., Foy, P., and Hooper, M. (2017). PIRLS 2016. International Results in Reading. International Association for the Evaluation of Educational Achievement (IEA). Available at: https://files.eric.ed.gov/fulltext/ED580353.pdf

Mullis, I., Martin, M., and Sainsbury, M. (2015). "PIRLS 2016 reading framework," in *PIRLS 2016 Assessment Framework*. eds. I. Mullis and M. O. Martin *2nd*

Edn. (Boston College, Chestnut Hill, MA: TIMSS & PIRLS International Study Center), 11–29.

Ozuru, Y., Dempsey, K., and McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learn. Instr.* 19, 228–242. doi: 10.1016/j.learninstruc.2008.04.003

R Core Team (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing Available at: https://www.R-project.org

Reimers, F., and Schleicher, A. (2020). Schooling Disrupted, Schooling Rethought. How the COVID-19 Pandemic is Changing Education. OECD. https://globaled.gse.harvard.edu/files/geii/files/education_continuity_v3.pdf

Reiss, K., Weis, M., Klieme, E., and Köller, O. (eds.) (2019). *PISA 2018: Grundbildung Im Internationalen Vergleich [PISA 2018: Basic Education in International Comparison]*. Münster, Germany: Waxmann Verlag.

Robitzsch, A., Grund, S., Henke, T., and Robitzsch, M. A. (2017). Package 'Miceadds'. R Package: Madison, WI, USA. Available at: https://cran.r-project.org/web/packages/miceadds/miceadds.pdf

Robitzsch, A., Kiefer, T., and Wu, M. (2019). TAM: Test Analysis Modules. R Package Version 3.3–10. https://CRAN.R-project.org/package=TAM

Rose, S., Twist, L., Lord, P., Rutt, S., Badr, K., Hope, C., et al. (2021). Impact of School Closures and Subsequent Support Strategies on Attainment and Socio-Emotional Wellbeing in Key Stage 1: Interim Paper 1. Education Endowment Foundation (EEF). https://educationendowmentfoundation.org.uk/public/files/Publications/COVID-19_Resources/Impact_of_school_closures_KS1_interim_findings_paper_-_Jan_2021.pdf

Rožman, M., Meinck, S., and Chen, M. (2022). "Impact of the pandemic on classroom teaching and learning," in *The Impact of the COVID-19 Pandemic on Education. International evidence from the Response to Education Disruption Survey (REDS)*. eds. S. Meinck, J. Fraillon and R. Strietholt (Paris, France: United Nations Educational, Scientific and Cultural Organization (UNESCO)), 54–83.

Sánchez Amate, J. J., Luque de la Rosa, A., Gutiérrez Cáceres, R., and Vargas Serrano, A. (2021). The effects of COVID-19 in the learning process of primary school students: a systematic review. *Educ. Sci.* 11:654. doi: 10.3390/educsci11100654

Savolainen, H., Ahonen, T., Aro, M., Tolvanen, A., and Holopainen, L. (2008). Reading comprehension, word reading and spelling as predictors of school achievement and choice of secondary education. *Learn. Instr.* 18, 201–210. doi: 10.1016/j.learninstruc.2007.09.017

Schneider, S., Beege, M., Nebel, S., Schnaubert, L., and Rey, G. D. (2021). The cognitive-affective-social theory of learning in digital environments (CASTLE). *Educ. Psychol. Rev.* 34, 1–38. doi: 10.1007/s10648-021-09626-5

Schult, J., Mahler, N., Fauth, B., and Lindner, M. A. (2021). Did students learn less during the COVID-19 pandemic? Reading and mathematics competencies before and after the first pandemic wave. [Preprint]. doi: 10.31234/osf.io/pqtgf

Schult, J., Mahler, N., Fauth, B., and Lindner, M. A. (2022). Long-term consequences of repeated school closures during the COVID-19 pandemic for reading and mathematics competencies. Psyarxiv [Preprint] doi: 10.31234/osf.io/dfqeg

Schunk, D. H., and Zimmerman, B. J. (2006). Influencing Children's self-efficacy and self-regulation of reading and writing through modeling. *Read. Writ. Q.* 23, 7–25. doi: 10.1080/10573560600837578

Schwippert, K. (2019). Was wird aus den Büchern? Sozialer Hintergrund von Lernenden und Bildungsungleichheit aus Sicht der international vergleichenden Erziehungswissenschaft. *J. Educ. Res. Online* 11, 92–117. doi: 10.25656/01:16789

Sénéchal, M., and LeFevre, J. A. (2002). Parental involvement in the development of children's reading skill: a five-year longitudinal study. *Child Dev.* 73, 445–460. doi: 10.1111/1467-8624.00417

Stanat, P., Schipolowski, S., Mahler, N., Weirich, S., and Henschel, S. (2019). *IQB-Bildungstrend 2018: Mathematische Und Naturwissenschaftliche Kompetenzen Am Ende der Sekundarstufe I Im Zweiten Ländervergleich [IQB Education Trend 2018: Mathematics and Science Competencies at the End of Secondary School Level I in the Second Country Comparison]*. Münster, Germany: Waxmann.

Steinmayr, R., Lazarides, R., Weidinger, A., and Christiansen, H. (2021). Teaching and learning during the first COVID-19 school lockdown: realization and associations with parent-perceived students' academic outcomes. *Zeitschrift für Padagogische Psychol.* 35, 85–106. doi: 10.1024/1010-0652/a000306

Storey, N., and Zhang, Q. (2021). A meta-analysis of COVID learning loss. EdArXiv [Preprint] doi: 10.35542/osf.io/qekw2

Tomasik, M. J., Helbling, L. A., and Moser, U. (2020). Educational gains of in-person vs. distance learning in primaryelementary and secondary schools: a natural experiment during the COVID-19 pandemic school closures in Switzerland. *Int. J. Psychol.* 56, 566–576. doi: 10.1002/ijop.12728

Toste, J. R., Didion, L., Peng, P., Filderman, M. J., and McClelland, A. M. (2020). A meta-analytic review of the relations between motivation and reading achievement for K–12 students. *Rev. Educ. Res.* 90, 420–456. doi: 10.3102/0034654320919352

Van der Linden, W. J. (2016). "Unidimensional logistic item response models," in *Handbook of Item Response Theory*. (New York, USA: CRC Press), 13–30.

Voogt, J., and Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st century competences: implications for national curriculum policies. *J. Curric. Stud.* 44, 299–321. doi: 10.1080/00220272.2012.668938

Wang, L., and Finch, H. (2018). Motivation variables mediate the relationship between socioeconomic status and academic achievement. *Psychol. Educ.* 55, 123–136. doi: 10.1002/per.842

Wang, Z., Sabatini, J., O'Reilly, T., and Weeks, J. (2019). Decoding and reading comprehension: a test of the decoding threshold hypothesis. *J. Educ. Psychol.* 111, 387–401. doi: 10.1037/edu0000302

Wendt, H., and Schwippert, K. (2017). "Lesekompetenzen von Schülerinnen und Schülern mit und ohne Migrationshintergrund [Reading competencies of students with and without a immigration background]," in *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. eds. A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper and E.-M. Lankes et al. (Münster, Germany: Waxmann), 219–234.

Werner, K., and Woessmann, L. (2021). The legacy of covid-19 in education. CESifo Working Paper, No. 9358, Center for Economic Studies and ifo Institute (CESifo). Available at: http://hdl.handle.net/10419/248903

Wigfield, A., and Eccles, J. (2000). Expectancy–value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps.1999.1015

Wigfield, A., Gladstone, J., and Turci, L. (2016). Beyond cognition: reading motivation and reading comprehension. *Child Dev. Perspect.* 10, 190–195. doi: 10.1111/cdep.12184

Woessmann, L., Freundl, V., Grewenig, E., Lergetporer, P., Werner, K., and Zierow, L. (2020). Bildung in der Coronakrise: Wie haben die Schulkinder die Zeit der Schulschließungen verbracht, und welche Bildungsmaßnahmen befürworten die Deutschen? ["Education in the Corona Crisis: How Have Schoolchildren Spent the Period of School Closures, and What Educational Measures do German's Advocate?"], ifo Schnelldienst, 73. Available at: https://www.ifo.de/DocDL/sd-2020-09-woessmann-etal-bildungsbarometer-corona.pdf

Wrase, M. (2020). "Schulrechtliche Herausforderungen in Zeiten der Corona-Pandemie [School law challenges in times of the Corona pandemic]," in *"Langsam vermisse ich die Schule…" – Schule während und nach der Corona-Pandemie*. eds. D. Fickermann and B. Edelstein (Münster, Germany: Waxmann), 105–116.

Zierer, K. (2021). Effects of pandemic-related school closures on pupils' performance and learning in selected countries: a rapid review. *Educ. Sci.* 11:252. doi: 10.3390/educsci11060252

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX A

Descriptive results comparing the student composition in 2016 and 2021.

	Year	*M*	*SD*	Δ2021–2016	*t*	Mis (%)
Reading achievement	2016	1000.00 (5.29)	100			–
	2021	980.46 (5.32)	102	**−19.55 (3.08)**	−6.34	–
Relative cohort age (years)	2016	10.19 (0.01)	0.30			0.50
	2021	10.32 (0.01)	0.29	**0.13 (0.01)**	14.13	0.96
Late enrollment (years)	2016	0.03 (0.02)	0.43			0.50
	2021	0.06 (0.02)	0.43	**0.03 (0.01)**	2.59	0.96
Grade retention (years)	2016	0.13 (0.02)	0.35			0.50
	2021	0.13 (0.02)	0.37	0.01 (0.01)	0.77	0.98
Females (%)	2016	50.12 (0.95)				0
	2021	50.02 (1.05)		−0.10 (1.53)	−0.07	0
Not born in Germany						
Child (%)	2016	5.49 (0.65)				12.63
	2021	13.83 (1.53)		**8.34 (0.9)**	9.28	17.45
One parent (%)	2016	13.32 (1.10)				18.40
	2021	12.78 (1.09)		−0.53 (1.03)	−0.52	22.61
Both parents (%)	2016	22.01 (2.11)				18.41
	2021	26.72 (2.38)		**4.71 (1.31)**	3.59	22.61
German not spoken at home[a] (%)	2016	18.86 (1.58)				10.58
	2021	22.26 (1.78)		**3.40 (1.23)**	2.76	13.04
Number of books at home (>100)[b]	2016	67.69 (1.69)				12.36
(%)	2021	67.25 (1.62)		−0.44 (1.43)	−0.31	15.38
Special educational needs[c] (%)	2016	2.98 (0.80)				1.96
	2021	4.12 (0.66)		**1.14 (0.57)**	2.01	0.00

Study 2016 N = 2,208 and 2021 N = 2,082, with N = 111 schools each. Bold estimates: $p < 0.05$

[a]*Percentage of children selecting "I always speak German at home" or "almost always speak German at home."*

[b]*Percentage of children selecting "Enough to fill two bookshelves (101–200)" or more.*

[c]*Children with an official diagnosis of special educational needs (i.e., emotional disability).*

frontiers | Frontiers in Education

ORIGINAL RESEARCH
published: 24 May 2022
doi: 10.3389/feduc.2022.768959

Academic Self-Efficacy, Procrastination, and Attrition Intentions

Efim Nemtcan, Rannveig Grøm Sæle, Thor Gamst-Klaussen and Frode Svartdal*

Department of Psychology, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway

Why do students leave universities? The current study addresses the problem of academic attrition from the perspective of students' intentions. Specifically, we focus on the roles of academic self-efficacy and procrastination in exploring their relationships with attrition intentions. Based on existing research, we expected a negative relationship between academic self-efficacy and attrition intentions, with procrastination as a possible mediator. Furthermore, it was expected that this relationship would differ depending on the type of attrition (i.e., drop-out, transfer university, transfer study field). These hypotheses were investigated among Norwegian students in a questionnaire study (N = 693). Results showed that procrastination partially mediated the relationship between academic self-efficacy and three attrition intentions categories. Although procrastination was a significant mediator of self-efficacy for all types of intentions, the sizes of the direct and indirect effects were different. We conclude that academic procrastination is important in understanding the relationship between students' self-efficacy beliefs and attrition intentions.

Keywords: academic attrition, attrition intentions, drop-out, transfer-out, academic self-efficacy, procrastination, mediation

OPEN ACCESS

Edited by:
Marcos Cupani, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

Reviewed by:
Aude Villatte, University of Quebec in Outaouais, Canada
Valeria Moran, Siglo 21 Business University, Argentina
Javier Sanchez Rosas, National University of Córdoba, Argentina

****Correspondence:***
Efim Nemtcan
efim.nemtcan@uit.no

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

Received: *29 September 2021*
Accepted: *28 April 2022*
Published: *24 May 2022*

Citation:
Nemtcan E, Sæle RG, Gamst-Klaussen T and Svartdal F (2022) Academic Self-Efficacy, Procrastination, and Attrition Intentions. Front. Educ. 7:768959. doi: 10.3389/feduc.2022.768959

INTRODUCTION

The rates of students' departure before degree completion (i.e., academic attrition) remain relatively high across Europe, with 24% of students leaving higher education before obtaining formal degree qualifications (OECD, 2019). Internationally, academic attrition remains on the agenda of higher education stakeholders. The increased importance of formal education, detrimental societal and personal consequences of academic attrition are among the main reasons for increased attention to the issue. For example, personal consequences might include short- and long-term economic consequences (i.e., needing to pay back study loans while earning lower wages due to the lack of formal qualifications) as well as reduced physical health and general well-being (Mayhew et al., 2016; Zajacova and Lawrence, 2018; Kirp, 2019). The leading social consequence is an inefficient use of government funding which might have more detrimental consequences in countries with state-funded higher education systems (OECD, 2021a,c). Therefore, research providing good explanations seems required to facilitate more effective solutions.

Academic attrition has usually been addressed from the perspective of students' actual behavior, despite research evidence on the role of intentions in explaining human behavior (e.g., Sheeran, 2002; Morwitz and Munz, 2020). Although some theoretical models address the role of attrition intentions (e.g., Bean, 1982; Tinto, 1993; Bean and Eaton, 2000), they do not differentiate between types of students' attrition (e.g., leaving permanently, changing university). However, evidence

shows that the predictive ability of intentions might be dependent on the behavior in question (e.g., Sheeran, 2002). Although intentions might be a good predictor of students' permanent departure from the university, the same might not be the case for changing university.

In turn, focusing on different types of students' attrition intentions might enable institutions to better address and assist students during the process of disengagement. For example, students intending to change the place of education or study field but continue their education may receive better support before (e.g., considering alternative solutions, providing information about the process) and during (e.g., grades transfer assistance) the actual transfer which might be beneficial for time spent on obtaining a degree (Li, 2010; Spencer, 2021). Students intending to leave altogether might need different types of support. Hence, counselors and university staff might adjust assistance or intervention strategies accordingly by knowing students' intentions. Interventions based on a vague definition of the target population and their intentions, on the other hand, may be limited in their effectiveness (e.g., Hovdhaugen, 2011).

The present study aims to address some of the central student-related factors and mechanisms involved in the process of attrition intentions formation. Understanding the mechanisms involved may assist researchers and practitioners in developing, assessing, and refining the assistance programs. In particular, the present study aims to assess the relationship of two psychologically grounded factors (i.e., academic self-efficacy and procrastination) with different types of attrition intentions. As will be discussed, academic self-efficacy is related to procrastination and students' attrition. However, few studies have investigated the relationship between procrastination and academic attrition. Further, these factors have not been examined accounting for different attrition intentions (e.g., leaving entirely, changing an academic institution). Hence, we first present a brief overview of academic attrition and the role of behavioral intentions. Then we proceed with an overview of the factors of interest in the present paper, self-efficacy and procrastination.

ACADEMIC ATTRITION AND ITS VARIABILITY

Researchers have used different terminology to describe that some students leave their studies before getting an official degree qualification. The operationalization of the phenomenon varied from "wastage" (e.g., Cross and Hall, 1954) to more recent "attrition." However, a common feature shared by both operationalizations is their negative connotation[1]. Although neither wastage nor attrition are appropriate to fully describe student departure, we will use the term *academic attrition*, an umbrella term for all types of academic discontinuations.

However, it is important to acknowledge that there are different forms of academic attrition. The notion that all students leaving higher education are not the same can be traced back to Tinto (1993). In his seminal work, Tinto (1993, Chapter 2) provides a synthesis of research and, importantly, distinguishes between two main categories of students' departure, *institutional* and *system* departure. The first type of departure describes a pattern of attrition when students switch academic institutions (i.e., transferring out), while the second distinguishes students who leave the wider education system altogether (i.e., dropping out). The categorization was primarily based on the registry data and pattern of students' behavior after leaving university. This distinction was seen as crucial since different factors were assumed to be involved. If an academic institution aims to handle departure, it is essential to know which type of departure a university is dealing with, institutional or system.

The institutional-system distinction is supported by research evidence indicating non-uniformity of the student population (Hoyt and Winn, 2004; Hovdhaugen, 2009; Jones-White et al., 2010; Kehm et al., 2019). For example, previous and current academic performance, or "problems related to meeting academic standards," are reported more frequently as reasons for leaving by drop-out than by transfer-out students (Hoyt and Winn, 2004; Hovdhaugen, 2009, 2011; Hovdhaugen and Aamodt, 2009). Indeed, transfer-out students have comparable performance with direct-entry students (Aulck and West, 2017; Quinn-Nilas et al., 2019). Also, Hovdhaugen (2009) found that background characteristics such as age, gender, and school grades are significantly related to dropping out, but not so for transfer-out behaviors. Transfer-out was more strongly related to students' motivation, educational goals, and field of study.

Behavioral Intentions to Leave Education

Behavioral intention is one of the most studied factors in basic and applied research on human behavior (Morwitz and Munz, 2020). Based on a meta-analysis by Sheeran (2002), intentions explain 28% of the variance in behaviors including alcohol consumption, weight loss, seatbelt use, training, smoking, and cancer screening, to name but a few. These findings align with the assumptions of the Theory of Reasoned Action (TRA)/Theory of Planned Behavior (TPB) stating that intentions are the closest antecedents of actual behaviors (Fishbein and Ajzen, 1975; Ajzen, 1991). According to TRA and TPB, intentions capture the motivational factors influencing actual behaviors. Intentions are indicators of how hard people are willing to try and how much effort they are planning to exert to perform behaviors. It is assumed that the stronger the intention to perform a behavior is, the more likely a person is to perform the behavior.

However, behavioral intentions or intentions to leave education have been rarely included in a theoretical discussion on academic attrition. This can be partially explained by the predominance of the sociological perspective on the issue (for review, see Melguizo, 2011; Aljohani, 2016; Behr et al., 2020). Nevertheless, some classical theories of academic attrition and

[1]Still, it is worth mentioning that not all types of students' departures are necessarily negative or, at least, not for everyone and not in every case (e.g., Faas et al., 2018). For example, changing university might be perceived as something positive from a student's perspective since he/she is presumably aiming for a degree qualification, only in a more suitable institution/place. Also, some students might take only specific courses to increase their qualifications while being employed. In addition, students may take a break from their studies for one or another reason and subsequently re-enroll to receive their academic degrees. Yet another group might find that higher education is not for them but may go on to something else without any negative consequences.

their reevaluations acknowledged the importance of students' intentions. For example, the ideas from TRA/TPB (i.e., intentions as antecedents of behaviors) were implemented in the student attrition models by Bean (1982) and Cabrera et al. (1993). The authors found that intentions to leave were the best predictor of students' actual attrition. Also, the importance of student's attrition intention as an antecedent of actual behavior is asserted in the models by Tinto (1993) and Bean and Eaton (2000). Although the models agree on the role of intentions, they do not address the variability of academic attrition. As discussed, different factors are related to the different types of attrition, and thus it might be the case for students' intentions. Moreover, based on the analysis of items used to measure students' intentions, the classical studies might have assessed students' persistence intentions (e.g., "Do you expect to return to this university next fall"; Bean, 1982). Still, it is evident that reasons for staying can differ from reasons for leaving.

To summarize, emerging evidence shows that transfer-out and drop-out students leave universities for different reasons. Thus, operationalizing and measuring students' departure in general terms such as wastage or attrition may lead to imprecise results and conclusions. For example, the overrepresentation of drop-out students in a study sample might lead to findings that are hardly applicable to transfer-out students, and the other way around. Further, few studies investigated differences in factors related to students' intentions. The central assumption of the majority of proposed theoretical models and frameworks is that students' attrition results from their interaction with the academic environment. Still, what is lacking in the interactionalist perspective and research on academic attrition is factors that are relevant for students and their learning. Further, relatively few studies have focused on factors that are malleable and for which evidence on possible interventions is available. In the present study, we aim to address these issues by assessing the relationship of academic self-efficacy and procrastination with students' drop-out, transfer university, and transfer study field intentions. As will be discussed, both factors may have theoretical and practical utility.

FACTORS AND MECHANISMS INVOLVED IN ACADEMIC ATTRITION

Academic Self-Efficacy

From a student's perspective, attrition can be seen as a manifestation of a flaw in motivation. According to results of multiple meta-analyses and reviews (e.g., Robbins et al., 2004; Richardson et al., 2012; Schneider and Preckel, 2017), academic self-efficacy shows the strongest relationship with both academic performance[2] and persistence. Also, indirect evidence shows that self-efficacy might be related to both dropping and transferring out behaviors. According to the Social Cognitive Theory (Bandura, 1997), individuals' confidence in their ability to perform a required course of action to solve a problem or achieve a desired goal (i.e., *self-efficacy*) is important for understanding human motivation and behavior. The basic principle behind self-efficacy is that individuals are more likely to engage, exert more effort, and persist in activities for which they have high self-efficacy. By and large, the evidence supports the theoretical predictions on the relationship of self-efficacy beliefs with the amount of effort devoted to and persistence on a certain task (Van Dinther et al., 2011; Jackson et al., 2012; Komarraju and Nadler, 2013; Puente−Diaz and Cavazos−Arroyo, 2018). In turn, students' efforts are related to both drop-out and transfer-out behaviors (Hovdhaugen, 2009).

Further, self-efficacy beliefs play a major role in Bean and Eaton's (2000) model of academic attrition. Similar to Tinto's (1975, 1993) and related theoretical models, student-university interaction is an important part of the model by Bean and Eaton (2000). Nevertheless, it adds an individual perspective or students' self-assesments of their interaction with university into the explanation of the attrition process. In particular, Bean and Eaton (2000) assumed that as the result of interaction with the university's environment, students' academic and social self-efficacy increases or decreases facilitating persistence or attrition intentions and actual behavior. Hence, the relationship between self-efficacy and students' attrition intentions can be assumed. Also, according to the Theory of Planned Behavior, self-efficacy as a dimension of behavioral control is a crucial aspect in the formation of behavioral intentions and has a direct relationship with actual behavior (Ajzen, 1991, 2002, 2020). According to TPB (Ajzen, 1991, 2002), behavior is primarily determined by attitudes toward behavior, subjective social norms or pressure from significant others, and perceived behavioral control (PBC). Individual's attitudes, subjective norms, and PBC influence behavior by facilitating intention to act. The theory assumes that behavioral intentions, which summarize the motivational forces (i.e., attitudes, subjective norms, and PBC), are the most approximate predictors of behaviors. In addition, the theory also suggests that PBC can have a direct impact on behavior.

Hence, academic self-efficacy is related to students' attrition intentions and actual attrition behaviors. Although the results of Robbins et al. (2004) meta-analysis support the importance of self-efficacy for students' retention, the size of the relationship was only moderate. Nevertheless, we argue that this relationship is crucial and has a great theoretical and practical utility. First, from a practical perspective, self-efficacy is a cognitive belief that is malleable to change (Bandura, 1997; Van Dinther et al., 2011; Bartimote-Aufflick et al., 2016). Second, from a theoretical perspective, the evidence on the relationship of self-efficacy with different categories of students' attrition (i.e., drop-out, transfer-out) is scarce. Third, according to Weissberg and Owen (2005), the findings of Robbins et al. (2004) might not be equally applicable to commuter students, which is the case for many European universities and our study sample. Thus, research on the importance of students' self-efficacy for different attrition intentions is of particular interest.

[2]Academic performance is the most stable predictor of drop behaviors (Tinto, 1975, 1993; Bean, 1982; Bean and Metzner, 1985; Bean and Eaton, 2000; Robbins et al., 2004). In addition, as discussed in the section on the variability of academic attrition, performance may be important in the distinction of drop-out and transfer-out students.

Malleability of Self-Efficacy

As noted, self-efficacy is assumed to effect engagement, effort, and persistence in tasks and behaviors (Bandura, 1997; Van Dinther et al., 2011). According to the Social Cognitive Theory (Bandura, 1997), there are four primary sources of information that influence or create self-efficacy: mastery experience (previous success experience), vicarious (observational) experience, social persuasion, and physical/affective states. The common characteristic describing these four sources of self-efficacy is that they are based on personal experience meaning that self-efficacy may be improved. Indeed, the evidence supports the theory's assertion. For example, Bartimote-Aufflick et al. (2016) reviewed 64 articles indicating 17 intervention studies investigating if certain teaching strategies or approaches can improve students' self-efficacy. Among these studies, ten interventions demonstrated improvement in participants' self-efficacy beliefs. In particular, facilitating opportunities to work with peers, helping students identify their misconceptions, including multimedia into the learning process, providing additional resources and activities for challenging concepts, and encouraging students to share their personal experiences were effective. Also, Van Dinther et al. (2011) note that interventions based on the Social Cognitive Theory are more effective with mastery experiences having the most powerful influence on self-efficacy beliefs. Here, providing practical experience such as performing a task while applying knowledge and skills in a demanding situation is argued to facilitate mastery experience. In addition, goal setting combined with self-reflection (i.e., self-regulation components) may influence students' perception of progress leading to mastery experience.

Procrastination and Academic Attrition

Procrastination has been defined as a voluntary delay of an intended course of action despite expecting to be worse off for doing so (Steel, 2007; Klingsieck, 2013). Procrastination can occur in all possible areas but is especially prevalent in the academic context (i.e., academic procrastination; Steel, 2007). Poor academic achievement, perceived stress, depression, and anxiety are among the potential outcomes of students' tendency to procrastinate (Steel, 2007; Klassen et al., 2008; Kim and Seo, 2015; Rozental et al., 2015; Sirois, 2016). To the best of our knowledge, only few studies have investigated the role of procrastination in academic attrition. For example, Grau and Minguillon (2013) demonstrated that students taking online programs who procrastinated in returning to university after taking a break from studies were more likely to leave permanently (i.e., drop out). Further, Bäulke et al. (2018) found that procrastination is related to drop-out intentions and mediated the relationship between motivational regulation and students' intentions. Also, results of a qualitative study by Visser et al. (2018) indicated that students scoring high on academic procrastination reported that they considered quitting their studies. Finally, Herrmann and Brandstätter (2015) found that an *action crisis* was predictive of disengagement from academic goals (i.e., dropout). An action crisis is a decisional conflict between continuing and disengaging from the pursuit of a personal goal. As defined by Herrmann and Brandstätter (2015), this conflict is characterized by six dimensions, including procrastination. However, to the best of our knowledge, there is no evidence on whether procrastination is related to other types of academic attrition (i.e., transfer-out intentions and behaviors).

Malleability of Procrastination

Similar to self-efficacy beliefs, research evidence shows that academic procrastination can be ameliorated (see meta-analysis by Van Eerde and Klingsieck, 2018, Malouff and Schutte, 2019). According to Van Eerde and Klingsieck (2018), cognitive-behavior therapy is the most effective approach. Still, such interventions are usually either *ad hoc*, time-consuming or require the involvement of professionals. Thus, interventions that would enable educators to support students effectively within their natural academic environment with little additional effort are of particular interest. According to Wäschle et al. (2014), one of such approaches may be strengthening students' self-efficacy beliefs. These authors argued that high self-efficacy facilitates students' achievement by increasing their motivation and application of effective learning strategies. Achievement, in turn, contributes to and raises self-efficacy which should facilitate students' motivation and achievement during the next learning cycle (i.e., virtuous cycle of self-efficacy). The results of the study supported these assumptions and indicated that self-efficacy beliefs have an important role in counteracting procrastination.

Procrastination as a Mediator

The research shows a close relationship between self-efficacy and procrastination. According to the Social Cognitive Theory (Bandura, 1997), high self-efficacy should increase students' effort and persistence devoted to a task. Hence, a negative relationship between self-efficacy and procrastination characterized by reduced effort and persistence is not unexpected (Van Eerde, 2003; Klassen et al., 2008; Wu and Fan, 2017). In addition, experimental evidence shows that altering students' negative and irrational thoughts (e.g., low self-efficacy) may be effective in reducing procrastination (Visser et al., 2017). The findings can be explained by the Temporal Motivational Theory (TMT; Steel and König, 2006). According to TMT, self-efficacy (an indicator of the expectancy construct) is crucial in explaining procrastination. In particular, motivation to perform a behavior (i.e., utility) is increased when people are confident of acquiring the desired reward (i.e., expectancy) or outcome (i.e., value). In turn, increased motivation should increase task performance or reduce task delay (i.e., procrastination).

In addition, although direct evidence on the environmentally driven nature of procrastination is scarce, different lines of research suggest that procrastination may be ingrained into the academic environment (Klingsieck, 2013; Svartdal et al., 2020). Hence, procrastination might represent an unintended environmental characteristic (i.e., academic system; Tinto, 1993) facilitating students' attrition intentions and actual attrition behaviors (Bean and Eaton, 2000). Likewise, evidence on the negative relationship of procrastination with academic performance is well-established (Steel, 2007; Kim and Seo, 2015). In turn, students' performance is a central aspect of the

student-university interaction perspective where performance is commonly defined as a mediating factor in the process of academic attrition (Aljohani, 2016). Finally, seen from a different perspective, academic attrition can be seen as a result of a goal-disengagement process (Brandstätter and Bernecker, 2021). In turn, action crisis characterized by delaying a goal pursuit (i.e., procrastination) has been commonly found to precede actual goal-disengagement (Herrmann and Brandstätter, 2015). Action crisis typically arises when individuals suffer from repeated setbacks. In the case of students, the setbacks may be determined by their self-efficacy beliefs (for review, see Honicke and Broadbent, 2016).

In sum, different lines of research suggest that having low self-efficacy beliefs may be detrimental to students' academic success and persistence. In this study, we will investigate whether this relationship can be explained (i.e., mediated) by students' tendency to procrastinate. As discussed, although the assumption is reasonable, there is no evidence on whether procrastination is related to other types of academic attrition beyond dropout (i.e., transfer-out intentions and behaviors). Hence, we aim to elucidate this aspect which may have practical utility for universities since both academic self-efficacy and procrastination are malleable to change (e.g., Van Dinther et al., 2011; Wäschle et al., 2014; Bartimote-Aufflick et al., 2016; Van Eerde and Klingsieck, 2018).

BACKGROUND FACTORS

Also, we considered several potentially relevant covariates including gender, age, high-school GPA, study field, university affiliation, years studied, parents' education, and history of changing study field or university. Previous empirical research suggests a relationship between students' background factors and actual attrition. For example, Hovdhaugen (2009) found that females, younger students, students whose parents have higher education, and students having better high-school GPAs are less likely to drop out. In contrast, transferring to another university is less likely when students are older and study natural sciences. Also, some evidence shows that females are more likely to switch majors (i.e., transfer study field) than males (Astorne-Figari and Speer, 2018; Meyer et al., 2021). Still, based on the findings of Ishitani and Flood (2018a), females may be less prone to change university (i.e., transfer university). Further, researchers note that attrition, including transferring out, varies across study fields and programs (DesJardins et al., 2003; Danaher et al., 2008; Ishitani and Flood, 2018b; Korhonen and Rautopuro, 2019). According to Wolter et al. (2014), students who have previously changed their study field or major are more likely to drop out. Similarly, changing university (i.e., history of changing university) was found to be negatively related to students' degree attainment and persistence (Ishitani, 2008; Li, 2010). Finally, Willcoxson (2010), Willcoxson et al. (2011), and Ishitani and Flood (2018b) found that different factors may drive students to drop and transfer out depending on their study year and university affiliation.

The findings that background factors (i.e., age, gender, high-school GPA) are important in the process of students attrition are in line with available theoretical models and frameworks (e.g., Tinto, 1975, 1993; Pascarella et al., 1983; Bean and Metzner, 1985). Still, the described associations are primarily found for students' actual behavior while evidence on students' intentions is scarce. Based on TPB, stating that intentions are the closest antecedents of actual behaviors, we assumed that the described factors are important for students' attrition intentions and, therefore, appropriate to control for in the analyses.

THE CURRENT STUDY

Academic attrition and persistence have been commonly viewed as the result of interaction between students and their academic environment leading to either persistence or attrition. Still, the mechanisms involved in the process of student-university interaction have rarely been addressed explicitly. In the present paper, we focus on the relationship of students' self-efficacy with different categories of attrition intentions (i.e., drop-out, transfer university, and transfer study field). As discussed, the relationship between self-efficacy beliefs and students' persistence/attrition is well-documented in the research literature (Robbins et al., 2004). Still, the evidence on the relationship of self-efficacy with other types of departure (i.e., transfer university or study field) is less clear. Also, there is little evidence on the mechanisms that explain this relationship. In the present study, we investigated if procrastination is one of such mechanisms. Self-efficacy beliefs are relatively strongly related to procrastination (Van Eerde, 2003; Klassen et al., 2008) which, in turn, is related to students' drop-out intentions (Bäulke et al., 2018). As discussed, low self-efficacy may incline students to delay and devote less effort to academic tasks facilitating students' attrition intentions (Van Eerde, 2003; Klassen et al., 2008; Hovdhaugen, 2009; Wu and Fan, 2017). In addition, seen as an environmental characteristic, procrastination may be important in the student-environment interaction process traditionally used to explain academic attrition (Tinto, 1993; Bean and Eaton, 2000; Svartdal et al., 2020). Finally, although it remains unknown whether interventions aimed at self-efficacy and procrastination substantially reduce academic attrition, the literature suggests that both factors are amenable to change. In this study, we assume that self-efficacy is negatively related to procrastination and attrition intentions (*Hypothesis 1*). Further, the relationship between self-efficacy and attrition intentions is mediated by students' procrastination tendency (*Hypothesis 2*). Finally (*Hypothesis 3*), we aim to explore if the observed relationships (i.e., direct and indirect) would differ for three types of attrition intentions (i.e., drop-out, transfer university, and transfer study field). It is expected that the observed relationships would differ for three types of intentions.

MATERIALS AND METHODS

Sample and Setting

Participants were 693 students (65% females) in different stages of their education: first-year (26%), second-year (25%), third-year (19%), fourth-year (13%), fifth-year (10%), and sixth-year or

more (7%). Age ranged from 19 to 54 with a mean of 23.9 years (*SD* = 4.79). The data was collected at the beginning of the spring semester (January- March) 2020 before the COVID restriction. The response and completion rates were satisfactory (41.2 and 88.5%, respectively).

Assessment and Measurement

Procedure and Ethics

Students were contacted via the university's e-mail and received an invitation to the study containing a brief study summary. Following the link, respondents were presented with a consent form, informed that they were anonymous and could refrain from answering or withdraw from the study at any time. Participants agreed to participate in the web-based survey by pressing a start survey button after reading information about the study. Uncompleted and suspicious responses (e.g., fast completion time) were excluded from analyses. The study was approved by the Norwegian Center for Research Data (NSD) in accordance with the requirements of data protection legislation (reference code 651244). The data for the present study is available on Open Science Framework (OSF)[3]. Participants could also participate in a random tracking of a gift card with a value of 1000 NOK. These participants provided their phone numbers which were recorded and stored separately from the rest of the data. Phone numbers were deleted when a winner had been chosen.

Covariates of Attrition Intentions

Students were asked to report their gender, age, high-school GPA, study field, university affiliation, years studied, parents' education, and history of changing study field or university (see **Supplementary Table 8** for descriptive data). *Age* was an open-ended question. *High-school GPA* was a categorical variable consisting of six categories (1 = Lowest grade; 6 = Highest grade). *Study field* was recorded into five categories: psychology; humanities and social science; science, technology, engineering and math (STEM); medicine and health science; biology and fishery. *Parents' education* included four categories: lower-secondary education, upper-secondary education, higher education, and other. Responses of students who chose "other" were recorded as missing. Parents' education was not distinguished into the mother's and father's levels of education based on data privacy considerations. *University affiliation* consisted of two categories: University of Tromsø (UiT) and Norwegian University of Science and Technology (NTNU). Only 18 participants were from other universities and, thus, were recorded as missing. *Number of years studied* at university was a six-categories variable (1 = 1 year; 6 = 6 years or more). Participants who have studied for 4 years and above were merged into one category due to the small sample size in the last two categories (i.e., 5, 6 years, and more). We also included two questions about students' *previous history of changing study fields* and *history of changing academic institutions* (0 = No; 1 = Yes). *Parents' education* (with university's education as the reference group), *university affiliation* (with students from NTNU as the reference group), *number of years studied* (with 1 year as the reference category), and *study field* (with medicine as the reference group) were dummy coded for subsequent mediation analyses. The reference category was chosen based on the easiness of interpretation (e.g., years studied). The medicine field was chosen as the reference group based on present results showing the most differences with other study fields. High-school GPA was subsequently excluded from the mediation analysis. According to the Social Cognitive Theory (Bandura, 1997), high school GPA is the antecedent of self-efficacy (i.e., previous or mastery experiences). In the study, it was related to students' academic self-efficacy beliefs (i.e., independent variable) and was insignificantly related to attrition intentions. Exclusion of high-school GPA did not lead to substantial changes in the estimated relationships.

Academic Self-Efficacy

The measurement index was borrowed from a Danish study by Herrmann et al. (2017). The scale is based on MSLQ (Motivated Strategies for Learning Questionnaire) by Pintrich (1991). Three items were chosen based on the reported highest factor loadings (Herrmann et al., 2017). An example item is: "I am confident that I can acquire the skills necessary to excel within my field of study" with higher scores indicating stronger self-efficacy beliefs (1 = Totally disagree; 5 = Totally agree). The items were translated to Norwegian with forward-back translation. Internal reliability (Cronbach's alpha) was 0.80. The measure was significantly related to students' self-reported academic performance (r = 0.39) and three study strategies subscales (relating ideas, r = 0.32; time-management, r = 0.38; unrelated memorizing, r = –0.39) consistent with the research literature (Robbins et al., 2004; Diseth, 2011; Richardson et al., 2012). This particular scale was chosen since the pure self-efficacy scale (i.e., task- or subject-specific) was deemed inappropriate in the context of the present study (i.e., students from different study fields). Still, it is worth mentioning that such decision could raise some questions about the construct validity (i.e., self-concept/self-efficacy distinction; Marsh et al., 2019).

Procrastination

A subset of four items from the Academic Procrastination Scale (APS; Mccloskey and Scielzo, 2015; Yockey, 2016) measured academic procrastination (e.g., "I know that I should work on a school work, but I just don't' do it"; "Cramming and last-minute studying is the best way that I study for a big test"). Based on the exploratory factor analysis performed before the main analysis, one item was excluded due to factor loading below 0.40 and low communality. The items were translated to Norwegian with forward-back translation. All items are rated on a 5-point scale with higher scores indicating more procrastination (1 = Totally disagree; 5 = Totally agree). Cronbach's alpha for 25 items was 0.94 (Mccloskey and Scielzo, 2015). The three items used in this study had Cronbach's alpha of 0.85. The measure was significantly related to students' self-reported academic performance (r = –0.20) and three study strategies subscales (relating ideas, r = –0.08; time-management, r = –0.71; unrelated

[3]https://osf.io/k8ax4/?view_only=f8cf1a2b15ab4da7b552e4a20a79e125

memorizing, $r = 0.23$) consistent with the research literature (Richardson et al., 2012; Saele et al., 2017).

Attrition Intentions

In the present study, we used four-item measure of students' intentions to drop out, transfer to another university, and transfer to another study field. Although the research on behavioral intentions is extensive (Sheeran, 2002), there is scarce evidence on validated and psychometrically sound measures of intentions (Fishman et al., 2020). Based on findings that intentions/thoughts of performing an action can vary in the degree of their specificity (Mashburn, 2000; Gollwitzer, 2012; Bäulke et al., 2021), we borrowed the first two items from the study by Hardre and Reeve (2003). Based on the face validity, they represented the first two (i.e., deliberation; intention or Rubicon) mindset phases of goal pursuit (Gollwitzer, 2012). The items were: "I sometimes consider dropping out of university before graduation," "I intend to drop out of school before graduation." Further, we designed two additional items for the study: "I sometimes think that other job opportunities suit me better than those I can get with my current education"; "I know what I am going to do if I withdraw from my studies." The items were intended to measure the deliberation and planning phases. Similar items were designed for transfer university intentions: "I sometimes think about how my life would be if I change my study place"; "I have a plan for when and how I will change my study place." The second pair of items measuring transfer study field intentions were the following: "I sometimes think about advantages and disadvantages of changing study field"; "I am waiting for the possibility to change my study field." Participants were also presented with a descriptive text for transfer study field intentions specifying the high-cost transfer (e.g., history → science; Meyer et al., 2021). Exploratory factor analysis was performed to test the dimensionality of the items. Based on the results, only two items for each type of intention were retained. All items are rated on a 5-point scale with higher scores indicating higher intentions. Spearman-Brown coefficient for drop-out, transfer university and transfer study field intentions were 0.73, 0.76, and 0.82 (Eisinga et al., 2013).

Analysis

Model Specification and Estimation

A structural equation model (SEM) using weighted least squares parameter (WLSMV) estimation was employed. The WLSMV estimation is appropriate when manifest variables are categorical or ordinal, and the sample size is relatively large (Muthén and Muthén, 2017). Model fit data were examined using the chi-square test ($\chi 2$), Comparative Fit Index (CFI), Tucker-Lewis Fit Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). For a more detailed description and discussion of the fit indices, the reader is referred to Hu and Bentler (1999) and Brown (2015). Standard fit cut-off values were applied: CFI, TLI > 0.95, SRMR < 0.08, and RMSEA < 0.06 (Hu and Bentler, 1999). Values equal to or lesser/higher than cut-off values indicate good or close fit. Although the traditional approach to mediation using ordinary least squares or Baron and Kenny's (1986) stepwise approach is widely used, we chose the SEM alternative. Based on the recent evaluations of the approaches to mediation analysis (e.g., Iacobucci et al., 2007; Kline, 2015), SEM seems to be superior to Baron and Kenny's (1986) regression approach. For example, SEM provides more accurate or less biased estimations due to adjustment for measurement error which is not possible with traditional mediation approaches. Confirmatory factor analysis (CFA) was performed to assess the validity of the measurement model (see **Supplementary Materials**). The results of CFA indicated an excellent fit: $\chi 2 = 94.737$, $df = 44$, $p < 0.001$; CFI = 0.993; TLI = 0.989; RMSEA = 0.041 (90% CI 0.029 – 0.052); SRMR = 0.028.

The results of observed indirect effects were interpreted in concordance with Zhao et al. (2010) approach to mediation analysis. The main characteristic and the difference of this approach from the traditionally applied Baron and Kenny's (1986) mediation analysis is the dependent-independent variables relationship. In particular, Zhao et al. (2010) argue that a zero-order relationship between dependent and independent variables should not necessarily be significant for proceeding with the mediation analysis. Under certain conditions (e.g., presence of mediator variables with opposite effects, presence of suppressing variables, temporal distance), a mediator variable may be exercising its effect even when no significant dependent-independent variables relationship is found. The main requirement for mediation is the significant interaction effect (i.e., indirect effect). Further, consistent with the proposed mediation approach, the authors provided an alternative to the «full, partial, and no mediation» categorization of mediation patterns. *Complementary mediation* is present when mediated and direct effects are significant and point in the same direction. In contrast, *competitive mediation* assumes that the same effects are present but point in the opposite direction. *Indirect-only mediation* describes a pattern when the mediated effect is significant while the direct effect is not. *Direct-only non-mediation* and *no-effect non-mediation* are patterns when either only direct effect is significant or all the relationships between variables are insignificant.

RESULTS

Academic Self-Efficacy and Drop-Out Intentions via Procrastination

The chi-square test was significant ($\chi 2 = 99.820$, $df = 44$, $p < 0.01$) for the model *without covariates*. However, the chi-square test statistics is sensitive to sample size and is usually significant in large samples (Hooper et al., 2008). Other fit indices indicated a very good model fit, CFI = 0.991; TLI = 0.986; RMSEA = 0.043 (90% CI 0.032–0.054); SRMR = 0.028. As seen in **Figure 1**, academic self-efficacy was negatively related to procrastination ($\beta = -0.265$, boot $SE = 0.047$, $p < 0.001$), which in turn was positively related to drop-out intentions ($\beta = 0.277$, boot $SE = 0.054$, $p < 0.001$). The direct effect from academic self-efficacy to drop-out intentions was significant and in expected direction ($\beta = -0.395$, boot $SE = 0.052$, $p < 0.001$). The indirect effect via procrastination was also significant and in the same direction as the direct effect ($\beta = -0.074$, boot $SE = 0.019$, $p < 0.001$), indicating *complementary mediation*

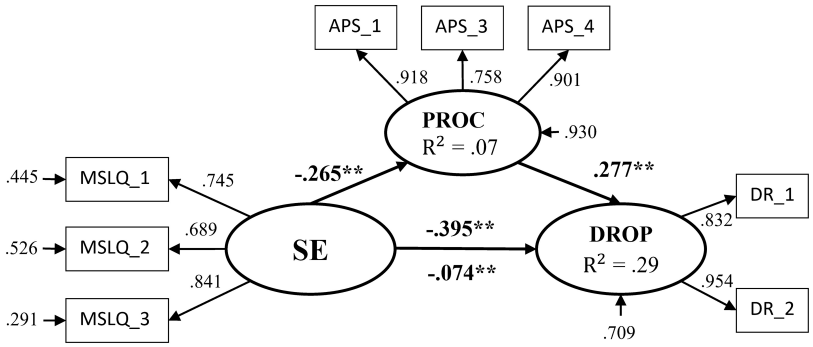


FIGURE 1 | Mediation model for drop-out intentions (n = 693). Indirect effect is provided below the path line. SE, academic self-efficacy; PROC, procrastination; DROP, drop-out intentions. $^{**}p \leq 0.001$.

(Zhao et al., 2010). The total effect was significant (β = –0.469, boot *SE* = 0.048, $p < 0.001$). Hence, procrastination "partially" mediated the relationship between academic self-efficacy and drop-out intentions. All additional estimates are provided in the **Supplementary Table 1**. Including covariates into the model did not substantially alter either model fit or mediation model relationships (see **Supplementary Table 2**).

Academic Self-Efficacy and Transfer University Intentions via Procrastination

The overall model fit for transfer university intentions *without covariates* was very good. The chi-square test was significant ($\chi 2 = 99.820$, $df = 44$, $p < 0.01$); CFI = 0.991; TLI = 0.986; RMSEA = 0.043 (90% CI 0.032–0.054); SRMR = 0.028. As seen in **Figure 2**, academic self-efficacy was negatively related to procrastination (β = –0.265, boot *SE* = 0.047, $p < 0.001$), which in turn was positively related to transfer university intentions (β = 0.168, boot *SE* = 0.066, $p < 0.01$). The direct effect from academic self-efficacy to transfer university intentions was insignificant and in expected direction (β = –0.102, boot *SE* = 0.063, $p = 0.11$). Still, the indirect effect via procrastination was significant and in the same direction as the direct effect (β = –0.045, boot *SE* = 0.020, $p < 0.01$), indicating *indirect-only mediation*. The total effect was significant (β = –0.212, boot *SE* = 0.049, $p < 0.001$). Hence, procrastination "fully" mediated the relationship between academic self-efficacy and transfer university intentions. All additional estimates are provided in the **Supplementary Table 3**. Including covariates into the model did not substantially change the overall model fit: $\chi 2 = 274.910$, $df = 188$, $p < 0.001$; CFI = 0.982; TLI = 0.975; RMSEA = 0.029 (90% CI 0.021 – 0.036); SRMR = 0.054. However, type of mediation changed from the indirect-only to complementary (see **Supplementary Table 4**). In particular, the direct relationship between academic self-efficacy and transfer university intentions became significant (β = –0.204, boot *SE* = 0.069, $p < 0.01$).

Academic Self-Efficacy and Transfer Study Field Intentions via Procrastination

The overall model fit for transfer study field intentions *without covariates* was very good. The chi-square test was significant ($\chi 2 = 99.820$, $df = 44$, $p < 0.01$); CFI = 0.991; TLI = 0.986; RMSEA = 0.043 (90% CI 0.032–0.054); SRMR = 0.028. As seen in **Figure 3**, academic self-efficacy was negatively related to procrastination (β = –0.265, boot *SE* = 0.047, $p < 0.001$), which in turn was positively related to transfer study field intentions (β = 0.181, boot *SE* = 0.057, $p < 0.001$). The direct effect from academic self-efficacy to transfer study field intentions was significant and in expected direction (β = –0.229, boot *SE* = 0.053, $p < 0.001$). The indirect effect via procrastination was also significant and in the same direction as the direct effect (β = –0.048, boot *SE* = 0.018, $p < 0.001$), indicating *complimentary mediation*. The total effect was significant (β = –0.276, boot *SE* = 0.049, $p < 0.001$). Hence, procrastination "partially" mediated the relationship between academic self-efficacy and transfer study field intentions. All additional estimates are provided in the **Supplementary Table 5**. Including covariates into the model did not substantially alter either model fit or mediation model relationships (see **Supplementary Table 6**).

RESULTS SUMMARY

The results of the three mediatory analyses supported *Hypothesis 1* that academic self-efficacy is negatively related to procrastination and attrition intentions. Also, *Hypothesis 2* was supported by results showing that the relationship of self-efficacy with drop-out and transfer study field intentions was complementary (partially) mediated by academic procrastination. These findings may indicate that the investigated models have an omitted mediator. In turn, the relationship between self-efficacy and transfer study field intentions was complementary mediated only when covariates were included in the model. Without covariates, procrastination indirect-only or fully mediated the investigated relationship. Hence, the inclusion of covariate variables into the model was reasonable. Finally, *Hypothesis 3* was supported by results indicating stronger relationships (i.e., direct and indirect) between self-efficacy and drop-out intentions than it was the case for two types of transfer-out intentions. Also, self-efficacy and procrastination accounted for a larger amount of variance in drop-out intentions ($R^2 = 29\%$)

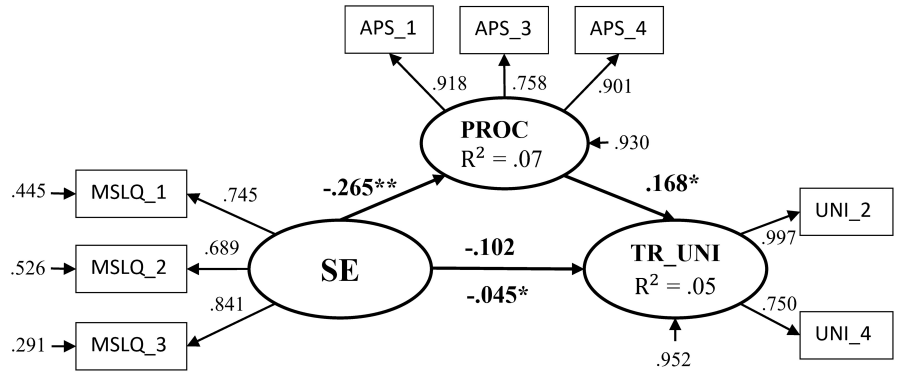


FIGURE 2 | Mediation model for transfer university intentions (n = 693). Indirect effect is provided below the path line. SE, Academic self-efficacy; PROC, procrastination; TR_UNI, transfer university intentions. *$p \leq 0.01$, **$p \leq 0.001$.

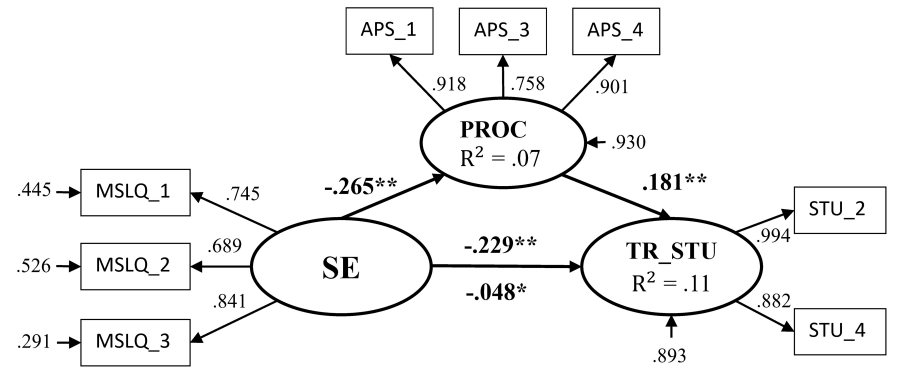


FIGURE 3 | Mediation model for transfer study field intentions (n = 693). Indirect effect is provided below the path line. SE, academic self-efficacy; PROC, procrastination; TR_STU, transfer study field intentions. *$p \leq 0.01$, **$p \leq 0.001$.

than in transfer university (R^2 = 5%) and transfer study field intentions (R^2 = 11%). The inclusion of control variables did not substantially change the observed relationships for drop-out and transfer study field intentions. In contrast, the relationship between academic self-efficacy and transfer university intentions has become complementary after the inclusion of covariates. In addition, different covariates turned out significant depending on the type of attrition intention. For instance, students' intentions to drop out and transfer study field differed between medicine and STEM fields with medicine students having fewer intentions. In contrast, no significant difference was found across the study fields for transfer university intentions (see **Supplementary Tables 2**, **4**, **6**). In sum, the results indicated the importance of the distinction between different categories of attrition intentions.

DISCUSSION

The present paper aimed to investigate the significance of the distinction between different categories of students' attrition intentions. Although students' motivation in general (Demetriou and Schmitz-Sciborski, 2011) and self-efficacy in particular (Bean and Eaton, 2000; Robbins et al., 2004; Willcoxson, 2010; Willcoxson et al., 2011; Tinto, 2017) are important for academic success and persistence, there is scarce evidence on the role of procrastination in academic attrition. In turn, understanding the involved mechanisms might assist researchers and practitioners in developing, assessing, and refining the assistance programs. Further, to the best of our knowledge, none of the previous studies investigated whether these relationships are present when accounting for the variability of academic attrition (i.e., dropping out, transferring out). The present study set out to investigate whether the relationship between academic self-efficacy and procrastination with attrition intentions would differ for drop-out and transfer-out intentions. It was also hypothesized that students' tendency to procrastinate would mediate the relationship between academic self-efficacy and attrition intentions.

The general pattern of results is in line with previous research. Academic self-efficacy was negatively related to procrastination (Van Eerde, 2003; Steel and König, 2006; Klassen et al., 2008) and attrition intentions (Robbins et al., 2004; Willcoxson, 2010; Willcoxson et al., 2011). Further, procrastination showed a positive relationship with attrition intentions, as in the study by Bäulke et al. (2018). However, our

findings supplement Bäulke et al.'s (2018) results by indicating that procrastination is also related to transfer-out intentions. Importantly, when the relationships are considered separately for each category, our findings align with prior evidence on the greater importance of academic factors for dropping out (Tinto, 1993; Hovdhaugen, 2009, 2011; Quinn-Nilas et al., 2019). This is represented by the larger amount of variance accounted for by academic self-efficacy and procrastination in drop-out intentions and larger relationships between academic self-efficacy and drop-out compared to transfer-out intentions. Hence, universities aiming to reduce academic attrition should adjust their strategies accordingly. For example, providing academic mentoring programs focusing on academic skills to reduce transfer university rates may prove less effective than expected.

In addition, our study demonstrates that students' academic self-efficacy significantly relates to attrition intentions through academic procrastination. It has been traditionally assumed that students' pre-entry characteristics or previous experiences determine the nature of student-university interaction (Aljohani, 2016). Likewise, past experiences also determine students' academic self-efficacy beliefs that have a well-established relationship with students' academic success (Robbins et al., 2004; Richardson et al., 2012). In turn, students who enter university with low self-efficacy might be at a considerable disadvantage compared to students with firm beliefs in their abilities. In particular, students with low self-efficacy tend to devote less effort, persistence to a given task, and procrastinate (Bandura, 1986, 1997; Steel, 2007). According to Wäschle et al. (2014), low self-efficacy may be involved in a vicious circle of procrastination (low self-efficacy, procrastination → poor performance → low self-efficacy → procrastination). Over time, in the face of recurrent setbacks (i.e., low performance), students may start to question the desirability and feasibility of their degree attainment goal leading to subsequent goal disengagement or attrition (Brandstätter and Bernecker, 2021). Even if students enter university with firm self-efficacy beliefs, many students lack the required competencies or abilities to succeed at university such as critical thinking or information literacy (Dunlosky et al., 2013). Lack of such skills in a students' toolbox puts them at a disadvantage causing poor achievement and, as described, might lead to procrastination and academic attrition.

Nevertheless, procrastination partially mediated the relation of academic self-efficacy with drop-out and transfer-out intentions. Obviously, other mechanisms associated with academic self-efficacy should be explored in future studies. One of the candidates for the role of a mediator is academic performance. According to the Social Cognitive Theory, self-efficacy beliefs influence which course of action a person takes, the amount of effort devoted to a task, resilience, and perseverance in the face of obstacles (Bandura, 1986, 1997). Unsurprisingly, empirical evidence shows a medium-strong relationship between self-efficacy and academic performance (e.g., Robbins et al., 2004; Richardson et al., 2012; Schneider and Preckel, 2017). However, as discussed, performance comes up to be a non-significant determinant of transferring out while it does predict drop-out behaviors. Further, according to social cognitive theory (Bandura, 1997), self-efficacy influences behavior through motivational processes. In turn, Hovdhaugen's (2009) study shows a significant relationship between students' motivation (i.e., intrinsic and extrinsic) and transferring out and a non-significant association with dropping out. Finally, self-efficacy is related to students' effort and commitment (Bandura, 1986, 1997; Weng et al., 2015). Both factors have been found important for students' drop-out and transfer-out behaviors (Tinto, 1993; Hovdhaugen, 2009). Hence, students' effort and goal commitment might be additional contributors (i.e., omitted mediators) in explaining the observed relationship of academic self-efficacy with drop-out and transfer-out intentions.

Further, in the present study, we performed the exploratory analysis with a set of covariates to investigate their relationship with attrition intentions and their influence on mediation relationships. The investigated covariates have been found important in relation to actual attrition behaviors. Still, these factors have not been addressed in the context of students' intentions. Although intentions are good approximators of actual behaviors, still, they do not account for the whole variance in actual behaviors meaning that the factors are not identical (Webb and Sheeran, 2006). Hence, it can be assumed that differences found for actual attrition behaviors (e.g., gender differences) might be absent in the case of students' intentions. The results of the present study supported this assumption. As discussed, previous findings show that female students are less prone to drop out and switch universities than males (Hovdhaugen, 2009; Ishitani and Flood, 2018a) while they are more likely to switch majors (Astorne-Figari and Speer, 2018; Meyer et al., 2021). However, we did not find any significant gender differences in drop-out, transfer university, and transfer study field intentions. Among investigated covariates, only years studied, study field, and history of changing university were significantly related to attrition intentions in the present study. In line with the findings by Willcoxson (2010), Willcoxson et al. (2011), and Ishitani and Flood (2018b), we found that students' attrition intentions differed by year of study. In particular, the longer the students studied, the fewer attrition intentions they had. Hence, assisting and paying extra attention to students during their first year at university seems crucial (Willcoxson et al., 2011). Further, it was found that students reporting that they have previously changed university had more transfer university intentions. In addition, student assistance may be less of a concern for some study fields than others. In particular, it was found that medical students have fewer drop-out and transfer-out intentions than students from other study fields (see **Supplementary Tables 2**, **4**, **6**). This might be related to higher enrollment standards and programs' structure (e.g., same students, closer follow-up of the students) than it is the case for other study majors. In sum, our findings show that although academic self-efficacy and procrastination are related to the three types of attrition intentions, adressing the attrition issue should be tailored to specific study programs and student characteristics. Also, considering students' characteristics such as year of education and previous history of changing study place might be more relevant in the case of transfer-out

students based on the results showing a change in mediated relationships after the inclusion of covariates (indirect-only → complementary mediation).

To sum up, the present study shows that academic self-efficacy and procrastination are related to students' intentions to drop out, change their field of study, and change university. Our results indicate that procrastination might be detrimental not only to traditionally investigated academic performance but also to other aspects of academic success (i.e., persistence). Hence, procrastination might have much more extensive consequences considering the negative relationship of attrition with students' future economic success and well-being (Hout, 2012; Mayhew et al., 2016). In addition, the size of the relationships, the nature of mediation, and the amount of variance accounted for were dependent on the type of intentions being considered indicating the relevance of the distinction among students' attrition intentions. Hence, future studies and interventions should be cautious when defining and drawing conclusions about academic attrition and attrition intentions.

Finally, the present study contributes to the current research by investigating the factors that are malleable and may be influenced by universities. For example, Van Dinther et al.'s (2011) literature review shows that self-efficacy interventions based on social cognitive theory are the most effective in improving self-efficacy. Some researchers (e.g., Bartimote-Aufflick et al., 2016) provide research-based best practice suggestions on how students' self-efficacy can be improved via teaching, learning support, and curriculum design. Similarly, evidence shows that procrastination can be ameliorated, with self-efficacy being one of the proposed alternatives for intervention (Wäschle et al., 2014; Van Eerde and Klingsieck, 2018). Nevertheless, counselors and university staff might need to adjust assistance or intervention strategies. As discussed, evidence and result of the present study show that students switching to another university may do it less due to performance-related problems (Hovdhaugen and Aamodt, 2009; Quinn-Nilas et al., 2019). Thus, assisting students in improving their self-efficacy beliefs when they intend to change university might be a less effective or appropriate solution for these students. In this case, universities might be better of adjusting their strategy based on students' intentions and known reasons for why these intentions occur. Still, it is worth mentioning that external factors (e.g., work, child care, illness, finances) are also responsible for students' attrition (Bean, 1985; Bean and Metzner, 1985; Leveson et al., 2013; Hovdhaugen, 2015; Behr et al., 2021). For example, Behr et al. (2021) identified a separate cluster of students who left university for personal (e.g., illness, stay abroad) or family (e.g., child care) reasons. Family or personal reasons were rarely decisive for dropping out and were reported by a small proportion of participants. Still, universities can hardly address these student difficulties directly. Hence, institutional ability to reduce student attrition may be limited indicating the need for more complex state interventions such as financial support or child-care arrangements. Finally, although improving students' self-efficacy and reducing procrastination may be a prospective approach to tackle students' attrition, its effectiveness for the students leaving primarily due to external reasons can be questioned.

Limitations and Future Research

One of the main limitations of the current study is the validity of the attrition intentions scale. The measure of students' intentions used in the present study should be cautiously evaluated since it lacks validation other than face validity. Also, factors with only two indicators are prone to estimation problems when the sample size is small (Kline, 2015). Hence, future psychometric studies developing and validating the attrition intentions scale that is applicable irrespective of statistical analysis are required. Further, although intentions represent the closest antecedent of behavior, they cannot substitute students' actual behavior (Webb and Sheeran, 2006). Thus, examination of the mechanisms found in the current study when students' actual behavior is also considered represents a prospective line for future research. In this regard, measures of intentions that depict students' firm resolution or concrete action plan can be considered for inclusion since they may be more predictive of actual behaviors (Brandstätter et al., 2015; Achtziger and Gollwitzer, 2018; Gollwitzer, 2018). However, implementation intentions (i.e., concrete if-then plans) might be problematic to measure in the context of academic attrition considering the lack of measurement scales and ethical considerations related to experimental designs. Still, future studies might test whether less concrete measures such as action planning would serve as a substitute and better predictor of students' behaviors (Hagger and Luszczynska, 2014).

Second, the causality of the proposed mediatory mechanisms should be cautiously evaluated due to the correlational study design. In the present study, the directional relationships were derived from the available research literature and theory (Bandura, 1997; Steel and König, 2006; Wu and Fan, 2017; Bäulke et al., 2018). It is reasonable to assume that self-efficacy determines students' attrition intentions and not the other way around. The results of the meta-analysis of experimental evidence indicate that changes in self-efficacy beliefs lead to changes in health-related intentions and behaviors (Sheeran et al., 2016). Nevertheless, future studies should account for alternative models (Danner et al., 2015) since the relationship between self-efficacy and procrastination may be bi-directional (Wäschle et al., 2014).

Third, the non-probability based sampling method (i.e., convenience sampling) has been used for data collection purposes due to the exploratory nature of the present study. Thus, generalization of the results to the student population should be made with caution. Future studies should preferably acquire the probability-based sampling methods to make more valid inferences about the whole population of Norwegian students.

Fourth, the self-efficacy measure used in the present study can be questioned in terms of its validity. According to Marsh et al. (2019), relatively "pure" self-efficacy measures are characterized by the future orientation and purely descriptive nature of response items and clear frame-of-reference. In particular, the present measure lacks a clear frame of reference such as being

confident in obtaining a top grade in a certain course. Although achieving such a standard in the present context (i.e., participants from different study fields) was nearly impossible, future research should clarify this aspect of the present study and if the observed relationships are better explained by a more pure self-efficacy measure. In addition, investigating the role of students' social self-efficacy may be a prospective line for future research. Based on the classical perspective on academic attrition (i.e., Tinto, 1975, 1993), Bean and Eaton's (2000) model suggests that academic and social self-efficacy are important in explaining student attrition. Still, to the best of our knowledge, we are not aware of any study which addressed the role of students' social self-efficacy in explaining different types of academic attrition.

Finally, in the present paper, we investigated the relationships between academic self-efficacy and procrastination with *high-cost* transfer study field intentions. The high-cost transfer is described by Meyer et al. (2021) as situations when students switch between broad categories of academic disciplines (e.g., history → science). In contrast, a *low-cost* transfer means situations when students switch within the same academic discipline (e.g., sociology → political science). The distinction is worth noting since Meyer et al. (2021) found that two categories might be related to different factors. In particular, high-school final grades were related to switching across disciplines (i.e., high-cost transfer), while misfit between student's occupational interests and major's content was mainly related to switching within disciplines (i.e., low-cost transfer). Thus, the results of the present study are only applicable to the high-cost transfer intentions. Future studies are encouraged to investigate the generalizability of the present findings to low-cost transfer intentions.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://osf.io/k8ax4/?view_only=f8cf1a2b15ab4da7b552e4a20a79e125.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

EN wrote the draft and did the statistical analyses under the supervision of FS. TG-K assisted with data analyses. RS, TG-K, and FS edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was facilitated by funds from the UiT The Arctic University of Norway and publication charges were funded by a grant from the publication fund of UiT The Arctic University of Norway.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2022.768959/full#supplementary-material

REFERENCES

Achtziger, A., and Gollwitzer, P. M. (2018). "Motivation and volition in the course of action," in *Motivation and Action*, eds J. Heckhausen and H. Heckhausen (Berlin: Springer International Publishing), 485–527. doi: 10.1007/978-3-319-65094-4_12

Ajzen, I. (1991). The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* 50, 179–211. doi: 10.1016/0749-5978(91)90020-t

Ajzen, I. (2002). Perceived behavioral control, self—efficacy, locus of control, and the theory of planned behavior. *J. Appl. Soc. Psychol.* 32, 665–683. doi: 10.1111/j.1559-1816.2002.tb00236.x

Ajzen, I. (2020). The theory of planned behavior: Frequently asked questions. *Hum. Behav. Emerg. Technol.* 2, 314–324. doi: 10.1002/hbe2.195

Aljohani, O. (2016). A comprehensive review of the major studies and theoretical models of student retention in higher education. *High. Educ. Stud.* 6, 1–18. doi: 10.5539/hes.v6n2p1

Astorne-Figari, C., and Speer, J. D. (2018). Drop out, switch majors, or persist? The contrasting gender gaps. *Econ. Lett.* 164, 82–85. doi: 10.1016/j.econlet.2018.01.010

Aulck, L., and West, J. (2017). Attrition and performance of community college transfers. *PLoS One* 12:e0174683. doi: 10.1371/journal.pone.0174683

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Hoboken, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: Freeman.

Baron, R. M., and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Pers. Soc. Psychol.* 51, 1173–1182. doi: 10.1037/0022-3514.51.6.1173

Bartimote-Aufflick, K., Bridgeman, A., Walker, R., Sharma, M., and Smith, L. (2016). The study, evaluation, and improvement of university student self-efficacy. *Stud. High. Educ.* 41, 1918–1942. doi: 10.1080/03075079.2014.999319

Bäulke, L., Eckerlein, N., and Dresel, M. (2018). Interrelations between motivational regulation, procrastination and college dropout intentions. *Unterrichtswissenschaft* 46, 461–479. doi: 10.1007/s42010-018-0029-5

Bäulke, L., Grunschel, C., and Dresel, M. (2021). Student dropout at university: a phase-orientated view on quitting studies and changing majors. *Eur. J. Psychol. Educ.* 24, 1–24. doi: 10.1007/s10212-021-00557-x

Bean, J. (1982). Student attrition, intentions, and confidence: interaction effects in a path model. *J. Assoc. Inst. Res.* 17, 291–320. doi: 10.1007/bf00977899

Bean, J. P. (1985). Interaction effects based on class level in an explanatory model of college student dropout syndrome. *Am. Educ. Res. J.* 22, 35–64. doi: 10.3102/00028312022001035

Bean, J. P., and Eaton, S. B. (2000). "A psychological model of college student retention," in *Reworking the Student Departure Puzzle,* ed. J. M. Braxton (Nashville: Vanderbilt University Press).

Bean, J. P., and Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Rev. Educ. Res.* 55, 485–540. doi: 10.3102/00346543055004485

Behr, A., Giese, M., Kamdjou, H. D. T., and Theune, K. (2020). Dropping out of university: a literature review. *Rev. Educ.* 8, 614–652. doi: 10.1002/rev3.3202

Behr, A., Giese, M., Kamdjou, H. D. T., and Theune, K. (2021). Motives for dropping out from higher education—an analysis of bachelor's degree students in Germany. *Eur. J. Educ.* 56, 325–343. doi: 10.1111/ejed.12433

Brandstätter, V., and Bernecker, K. (2021). Persistence and disengagement in personal goal pursuit. *Annu. Rev. Psychol.* 73, 271–299. doi: 10.1146/annurev-psych-020821-110710

Brandstätter, V., Giesinger, L., Job, V., and Frank, E. (2015). The role of deliberative versus implemental mindsets in time prediction and task accomplishment. *Soc. Psychol.* 46, 104–115. doi: 10.1027/1864-9335/a000231

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*, 2nd Edn. New York, NY: Guilford Press.

Cabrera, A. F., Nora, A., and Castaneda, M. B. (1993). College persistence: structural equations modeling test of an integrated model of student retention. *J. High. Educ.* 64, 123–139. doi: 10.2307/2960026

Cross, K. W., and Hall, D. L. A. (1954). Survey of entrants to nurse training schools and of student-nurse wastage in the Birmingham region. *Brit. J. Prev. Soc. Med.* 8, 70–76. doi: 10.1136/jech.8.2.70

Danaher, P. A., Bowser, D., and Somasundaram, J. (2008). The student departure puzzle: Do some faculties and programs have answers? *High. Educ. Res. Dev.* 27, 271–280. doi: 10.1080/07294360802183820

Danner, D., Hagemann, D., and Fiedler, K. (2015). Mediation analysis with structural equation models: combining theory, design, and statistics. *Eur. J. Soc. Psychol.* 45, 460–481. doi: 10.1002/ejsp.2106

Demetriou, C., and Schmitz-Sciborski, A. (2011). "Integration, motivation, strengths and optimism: retention theories past, present and future," in *Proceedings of the 7th National Symposium on Student Retention*, (Charleston, SC), 300–312.

DesJardins, S. L., Kim, D.-O., and Rzonca, C. S. (2003). A nested analysis of factors affecting bachelor's degree completion. *J. Coll. Stud. Retent.* 4, 407–435. doi: 10.2190/bgmr-3ch7-4k50-b5g3

Diseth, Å (2011). Self-efficacy, goal orientations and learning strategies as mediators between preceding and subsequent academic achievement. *Learn. Individ. Differ.* 21, 191–195. doi: 10.1016/j.lindif.2011.01.003

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., and Willingham, D. T. (2013). Improving students' learning with effective learning techniques. *Psychol. Sci. Public. Interest* 14, 4–58. doi: 10.1177/1529100612453266

Eisinga, R., Grotenhuis, M., and Pelzer, B. (2013). The reliability of a two-item scale: pearson, Cronbach, or Spearman-Brown? *Int. J. Public Health* 58, 637—-642. doi: 10.1007/s00038-012-0416-3

Faas, C., Benson, M. J., Kaestle, C. E., and Salva, J. (2018). Socioeconomic success and mental health profiles of young adults who drop out of college. *J. Youth Stud.* 21, 669–686. doi: 10.1080/13676261.2017.1406598

Fishbein, M., and Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Boston, MA: Addison-Wesley.

Fishman, J., Lushin, V., and Mandell, D. S. (2020). Predicting implementation: comparing validated measures of intention and assessing the role of motivation when designing behavioral interventions. *Implement. Sci. Commun.* 1:81. doi: 10.1186/s43058-020-00050-4

Gollwitzer, P. M. (2012). "Mindset theory of action phases," in *Handbook of Theories of Social Psychology*, eds P. Van Lange, A. W. Kruglanski, and E. T. Higgins (London: Sage), 526–546. doi: 10.3389/fpsyg.2020.00922

Gollwitzer, P. M. (2018). The goal concept: a helpful tool for theory development and testing in motivation science. *Motiv. Sci* 4, 185–205. doi: 10.1037/mot0000115

Grau, J., and Minguillon, J. (2013). When procrastination leads to dropping out: analysing students at risk. *eLC Res. Pap. Ser.* 6, 63–74.

Hagger, M. S., and Luszczynska, A. (2014). Implementation intention and action planning interventions in health contexts: state of the research and proposals for the way forward. *Appl. Psychol. Health Well Being* 6, 1–47. doi: 10.1111/aphw.12017

Hardre, P. L., and Reeve, J. (2003). A motivational model of rural students' intentions to persist in, versus drop out of, high school. *J. Educ. Psychol.* 95, 347–356. doi: 10.1037/0022-0663.95.2.347

Herrmann, K. J., Bager–Elsborg, A., and McCune, V. (2017). Investigating the relationships between approaches to learning, learner identities and academic achievement in higher education. *High. Educ.* 74, 385–400. doi: 10.1007/s10734-016-9999-6

Herrmann, M., and Brandstätter, V. (2015). Action crises and goal disengagement: longitudinal evidence on the predictive validity of a motivational phase in goal striving. *Motiv. Sci.* 1, 121–136. doi: 10.1037/mot0000016

Honicke, T., and Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: a systematic review. *Educ. Res. Rev.* 17, 63–84. doi: 10.1016/j.edurev.2015.11.002

Hooper, D., Coughlan, J., and Mullen, M. R. (2008). Structural equation modelling: guidelines for determining model fit. *Electron. J. Bus. Res. Methods* 6, 53–60.

Hout, M. (2012). Social and economic returns to college education in the United States. *Annu. Rev. Sociol.* 38, 379–400. doi: 10.1146/annurev.soc.012809.102503

Hovdhaugen, E. (2009). Transfer and dropout: different forms of student departure in Norway. *Stud. High. Educ.* 34, 1–17. doi: 10.1080/03075070802457009

Hovdhaugen, E. (2011). Do structured study programmes lead to lower rates of dropout and student transfer from university? *Irish Educ. Stud.* 30, 237–251. doi: 10.1080/03323315.2011.569143

Hovdhaugen, E. (2015). Working while studying: the impact of term-time employment on dropout rates. *J. Educ. Work* 28, 631–651. doi: 10.1080/13639080.2013.869311

Hovdhaugen, E., and Aamodt, P. O. (2009). Learning environment: relevant or not to students' decision to leave university? *Qual. High. Educ.* 15, 177–189. doi: 10.1080/13538320902995808

Hoyt, J. E., and Winn, B. A. (2004). Understanding retention and college student bodies: differences between drop-outs, stop-outs, opt-outs, and transfer-outs. *NASPA J.* 41, 395–417. doi: 10.2202/1949-6605.1351

Hu, L.-T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6, 1–55. doi: 10.1080/10705519909540118

Iacobucci, D., Saldanha, N., and Deng, X. (2007). A meditation on mediation: evidence that structural equations models perform better than regressions. *J. Consum. Psychol.* 17, 139–153. doi: 10.1016/s1057-7408(07)70020-7

Ishitani, T., and Flood, L. (2018a). Reverse and horizontal transfer-out behavior at four-year institutions. *Coll. Univ.* 93, 16–28.

Ishitani, T., and Flood, L. (2018b). Student transfer-out behavior at four-year institutions. *Res. High. Educ.* 59, 825–846. doi: 10.1007/s11162-017-9489-4

Ishitani, T. T. (2008). How do transfers survive after "transfer shock"? a longitudinal study of transfer student departure at a four-year institution. *Res. High. Educ.* 49, 403–419. doi: 10.1007/s11162-008-9091-x

Jackson, B., Myers, N. D., Taylor, I. M., and Beauchamp, M. R. (2012). Relational efficacy beliefs in physical activity classes: a test of the tripartite model. *J. Sport Exerc. Psychol.* 34, 285–304. doi: 10.1123/jsep.34.3.285

Jones-White, D. R., Radcliffe, P. M., Huesman, R. L., and Kellogg, J. P. (2010). Redefining student success: applying different multinomial regression techniques for the study of student graduation across institutions of higher education. *Res. High. Educ.* 51, 154–174. doi: 10.1007/s11162-009-9149-4

Kehm, B. M., Larsen, M. R., and Sommersel, H. B. (2019). Student dropout from universities in Europe: a review of empirical literature. *Hung. Educ. Res. J.* 9, 147–164. doi: 10.1556/063.9.2019.1.18

Kim, K. R., and Seo, E. H. (2015). The relationship between procrastination and academic performance: a meta-analysis. *Pers. Indivd. Differ.* 82, 26–33. doi: 10.1016/j.paid.2015.02.038

Kirp, D. (2019). *The College Dropout Scandal*. Oxford: Oxford University Press.

Klassen, R. M., Krawchuk, L. L., and Rajani, S. (2008). Academic procrastination of undergraduates: low self-efficacy to self-regulate predicts higher levels of procrastination. *Contemp. Educ. Psychol.* 33, 915–931. doi: 10.1016/j.cedpsych.2007.07.001

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling*, 4 Edn. New York, NY: Guilford Press.

Klingsieck, K. B. (2013). Procrastination: when good things don't come to those who wait. *Eur. Psychol.* 18, 24–34. doi: 10.1027/1016-9040/a000138

Komarraju, M., and Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learn. Individ. Differ.* 25, 67–72. doi: 10.1016/j.lindif.2013.01.005

Korhonen, V., and Rautopuro, J. (2019). Identifying problematic study progression and "at-risk" students in higher education in Finland. *Scand. J. Educ. Res.* 63, 1056–1069. doi: 10.1080/00313831.2018.1476407

Leveson, L., McNeil, N., and Joiner, T. (2013). Persist or withdraw: the importance of external factors in students' departure intentions. *High. Educ. Res. Dev.* 32, 932—-945. doi: 10.1080/07294360.2013.806442

Li, D. (2010). They need help: transfer Students from Four-Year to Four-Year Institutions. *Rev. High. Educ.* 33, 207–238. doi: 10.1353/rhe.0.0131

Malouff, J. M., and Schutte, N. S. (2019). The efficacy of interventions aimed at reducing procrastination: a meta−analysis of randomized controlled trials. *J. Couns. Dev.* 97, 117–127. doi: 10.1002/jcad.12243

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., et al. (2019). The Murky distinction between self-concept and self-efficacy: beware of lurking jingle-jangle fallacies. *J. Educ. Psychol.* 111, 331–353. doi: 10.1037/edu0000281

Mashburn, A. J. (2000). A psychological process of college student dropout. *J. Coll. Stud. Retent.* 2, 173–190. doi: 10.2190/u2qb-52j9-ghgp-6lee

Mayhew, M. J., Rockenbach, A. N., Bowman, N. A., Seifert, T. A., Wolniak, G. C., Pascarella, E. T., et al. (2016). *How College Affects Students: 21st Century Evidence that Higher Education Works*. San Francisco, CA: Jossey-Bass.

Mccloskey, J., and Scielzo, A. S. (2015). *Finally!: The Development and Validation of the Academic Procrastination Scale.* Arlington, TX: University Of Texas. doi: 10.13140/RG.2.2.23164.64640

Melguizo, T. (2011). "A review of the theories developed to describe the process of college persistence and attainment," in *Higher Education: Handbook of Theory and Research*, eds J. Smart and M. Paulsen (Berlin: Springer), 395–424. doi: 10.1007/978-94-007-0702-3_10

Meyer, J., Leuze, K., and Strauss, S. (2021). Individual achievement, person-major fit, or social expectations: Why do students switch majors in German higher education? *Res. High. Educ.* 63, 222–247. doi: 10.1007/s11162-021-09650-y

Morwitz, V. G., and Munz, K. P. (2020). Intentions. *Consum. Psychol. Rev.* 4, 26–41. doi: 10.1002/arcp.1061

Muthén, L. K., and Muthén, B. O. (2017). *Mplus User's Guide*, 8th Edn. Los Angeles, CA: Muthén & Muthén.

OECD (2019). *Education at a Glance 2019: OECD Indicators*. Paris: OECD Publishing.

OECD (2021a). *Education at a Glance 2021: OECD Indicators*. Paris: OECD Publishing, doi: 10.1787/b35a14e5-en

OECD (2021c). *How much is Spent per Student on Educational Institutions?*. Paris: OECD Publishing, doi: 10.1787/be9806af-en

Pascarella, E. T., Duby, P. B., and Iverson, B. K. (1983). A test and reconceptualization of a theoretical model of college withdrawal in a commuter institution setting. *Sociol. Educ.* 56, 88–100. doi: 10.2307/2112657

Pintrich, P. R. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor, MI: National Center for Research to Improve Post-Secondary Teaching.

Puente−Diaz, R., and Cavazos−Arroyo, J. (2018). An exploration of some antecedents and consequences of creative self−efficacy among College Students. *J. Creat. Behav.* 52, 256–266. doi: 10.1002/jocb.149

Quinn-Nilas, C., Kennett, D. J., and Maki, K. (2019). Examining explanatory style for failure of direct entry and transfer students using structural equation modelling. *Educ. Psychol.* 39, 749–767. doi: 10.1080/01443410.2019.1574340

Richardson, M., Abraham, C., and Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol. Bull.* 138, 353–387. doi: 10.1037/a0026838

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., and Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychol. Bull.* 130, 261–288. doi: 10.1037/0033-2909.130.2.261

Rozental, A., Forsell, E., Svensson, A., Forsström, D., Andersson, G., and Carlbring, P. (2015). Differentiating procrastinators from each other: a cluster analysis. *Cogn. Behav. Therapy* 44, 480–490. doi: 10.1080/16506073.2015.1059353

Saele, R. G., Dahl, T. I., Sørlie, T., and Friborg, O. (2017). Relationships between learning approach, procrastination and academic achievement amongst first-year university students. *High. Educ.* 74, 757–774. doi: 10.1007/s10734-016-0075-z

Schneider, M., and Preckel, F. (2017). Variables associated with achievement in higher education: a systematic review of meta-analyses. *Psychol. Bull.* 143, 565–600. doi: 10.1037/bul0000098

Sheeran, P. (2002). Intention—behavior relations: a conceptual and empirical review. *Eur. Rev. Soc. Psychol.* 12, 1–36. doi: 10.1080/14792772143000003

Sheeran, P., Maki, A., Montanaro, E., Avishai-Yitshak, A., Bryan, A., Klein, W. M. P., et al. (2016). The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: a meta-analysis. *Health Psychol.* 35, 1178–1188. doi: 10.1037/hea0000387

Sirois, F. M. (2016). "Procrastination, stress, and chronic health conditions: a temporal perspective," in *Procrastination, Health, and Well-Being*, eds F. M. Sirois and T. Pychyl (Amsterdam: Elsevier), 67–92. doi: 10.1016/B978-0-12-802862-9.00004-9

Spencer, G. (2021). Off the beaten path: Can statewide articulation support students transferring in nonlinear directions? *Am. Educ. Res. J.* 58, 1070–1102. doi: 10.3102/0002831221999782

Steel, P. (2007). The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychol. Bull.* 133, 65–94. doi: 10.1037/0033-2909.133.1.65

Steel, P., and König, C. J. (2006). Integrating theories of motivation. *Acad. Manage. Rev.* 31, 889–913. doi: 10.5465/amr.2006.22527462

Svartdal, F., Dahl, T. I., Gamst-Klaussen, T., Koppenborg, M., and Klingsieck, K. B. (2020). How study environments foster academic procrastination: overview and recommendations. *Front. Psychol.* 11:540910. doi: 10.3389/fpsyg.2020.540910

Tinto, V. (1975). Dropout from higher education: a theoretical synthesis of recent research. *Rev. Educ. Res.* 45, 89–125. doi: 10.3102/00346543045001089

Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition*, 2nd Edn. Chicago, IL: The University of Chicago Press.

Tinto, V. (2017). Through the eyes of students. *J. Coll. Stud. Retent. Res. Theory Pract.* 19, 254–269. doi: 10.1177/1521025115621917

Van Dinther, M., Dochy, F., and Segers, M. (2011). Factors affecting students' self-efficacy in higher education. *Educ. Res. Rev.* 6, 95–108. doi: 10.1016/j.edurev.2010.10.003

Van Eerde, W. (2003). A meta-analytically derived nomological network of procrastination. *Pers. Indivd. Differ.* 35, 1401–1418. doi: 10.1016/s0191-8869(02)00358-6

Van Eerde, W., and Klingsieck, K. B. (2018). Overcoming procrastination? A meta-analysis of intervention studies. *Educ. Res. Rev.* 25, 73–85. doi: 10.1016/j.edurev.2018.09.002

Visser, L., Korthagen, F. A. J., and Schoonenboom, J. (2018). Differences in learning characteristics between students with high, average, and low levels of academic procrastination: students' views on factors influencing their learning. *Front. Psychol.* 9:808. doi: 10.3389/fpsyg.2018.00808

Visser, L., Schoonenboom, J., and Korthagen, F. A. J. (2017). A field experimental design of a strengths-based training to overcome academic procrastination: short- and long-term effect. *Front. Psychol.* 8:1949. doi: 10.3389/fpsyg.2017.01949

Wäschle, K., Allgaier, A., Lachner, A., Fink, S., and Nückles, M. (2014). Procrastination and self-efficacy: tracing vicious and virtuous circles in self-regulated learning. *Learn. Instr.* 29, 103–114. doi: 10.1016/j.learninstruc.2013.09.005

Webb, T. L., and Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychol. Bull.* 132, 249–268. doi: 10.1037/0033-2909.132.2.249

Weissberg, N. C., and Owen, D. R. (2005). Do psychosocial and study skill factors predict college outcomes? Comment on Robbins et al. (2004). *Psychol. Bull.* 131, 407–409. doi: 10.1037/0033-2909.131.3.407

Weng, F., Cheong, F., and Cheong, C. (2015). Modelling is student retention in Taiwan: extending Tinto and Bean's model with self-efficacy. *Innov. Teach. Learn. Inf. Comput. Sci.* 9, 1–12. doi: 10.11120/ital.2010.09020012

Willcoxson, L. (2010). Factors affecting intention to leave in the first, second and third year of university studies: a semester−by−semester investigation. *High. Educ. Res. Dev.* 29, 623–639. doi: 10.1080/07294360.2010.501071

Willcoxson, L., Cotter, J., and Joy, S. (2011). Beyond the first-year experience: the impact on attrition of student experiences throughout undergraduate degree studies in six diverse universities. *Stud. High. Educ.* 36, 331–352. doi: 10.1080/03075070903581533

Wolter, S. C., Diem, A., and Messer, D. (2014). Drop−outs from Swiss Universities: an empirical analysis of data on all students between 1975 and 2008. *Eur. J. Educ.* 49, 471–483. doi: 10.1111/ejed.12096

Wu, F., and Fan, W. (2017). Academic procrastination in linking motivation and achievement-related behaviours: a perspective of expectancy-value theory. *Educ. Psychol.* 37, 695–711. doi: 10.1080/01443410.2016.1202901

Yockey, R. D. (2016). Validation of the short form of the Academic Procrastination Scale. *Psychol. Rep.* 118, 171–179. doi: 10.1177/0033294115626825

Zajacova, A., and Lawrence, E. M. (2018). The relationship between education and health: reducing disparities through a contextual approach. *Annu. Rev. Public Health* 39, 273–289. doi: 10.1146/annurev-publhealth-031816-044628

Zhao, X., Lynch, J. G., and Chen, Q. (2010). Reconsidering Baron and Kenny: myths and truths about mediation analysis. *J. Consum. Res.* 37, 197–206. doi: 10.1086/651257

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Education

BRIEF RESEARCH REPORT
published: 03 June 2022
doi: 10.3389/feduc.2022.901319

What Kind of Students Attend Cyber Schools? Pandemic Enrollment as Evidence of Negative Selection

Ian Kingsbury[1], Dennis Beck[2] and Martha Bradley-Dorsey[2]*

[1] Educational Freedom Institute, Phoenix, AZ, United States, [2] Department of Curriculum and Instruction, University of Arkansas, Fayetteville, AR, United States

Fully online virtual schools have consistently underperformed academically compared to brick and mortar schools. Scholars debate the extent to which these differences are due school quality or the type of student that attends virtual schools. The large number of students who enrolled in virtual schools during the COVID-19 pandemic provides a unique opportunity to revisit this debate, as the phenomenon plausibly attenuates negative selection into virtual schools. Previous research concluded that a virtual school COVID cohort resembled prior groups demographically but reported greater success at their prior in-person schools and in cyber schools, however, it offered only limited insight into their academic performance at their virtual school. We use data from a large cyber charter network ("Countrywide Cyber") to assess whether students who enrolled in full time virtual schools due to COVID-related concerns performed better on entry diagnostic assessments. Results indicate that students who enrolled due to COVID-19 were stronger academically, corroborating recent descriptive research. The implications of these results for practice and policy are discussed.

Keywords: cyber schools, virtual schools, learning outcomes, negative selection, charter schools

OPEN ACCESS

Edited by:
Marcos Cupani, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

Reviewed by:
Zoe A. Morris, Monash University, Australia
Leticia Elizabeth Luque, National University of Córdoba, Argentina

****Correspondence:***
Ian Kingsbury
iansethkingsbury@gmail.com

Specialty section:
This article was submitted to Assessment, Testing and Applied Measurement, a section of the journal Frontiers in Education

Received: *21 March 2022*
Accepted: *18 May 2022*
Published: *03 June 2022*

Citation:
Kingsbury I, Beck D and Bradley-Dorsey M (2022) What Kind of Students Attend Cyber Schools? Pandemic Enrollment as Evidence of Negative Selection. Front. Educ. 7:901319. doi: 10.3389/feduc.2022.901319

BACKGROUND

Full-time, online schools, known as "cyber schools," have rapidly expanded in the last 20 years, making research a priority (Molnar et al., 2019). Student enrollment in these schools increased by nearly 30,000 students between 2017–2018 and 2019-2020 alone (Molnar et al., 2019). Early research by Clark (2000) first named these schools as "cyber schools," categorizing them among six other kinds of online learning programs. Since that point, this terminology has been used in the "Virtual Schools in the U.S." reports that are regularly published by the National Educational Policy Center (e.g., Molnar et al., 2019).

Cyber schools have consistently underperformed academically compared to traditional public schools (for summaries by charter school supporters see Finn et al., 2016; see Saultz and Fusarelli, 2017 for a summary by critics of cyber charter schools; also see Molnar et al., 2019 for a more balanced summary).The Center for Research on Educational Outcomes (CREDO) produced oft-cited descriptive research that indicates that these negative effects are statistically and practically significant. Still, some research hints at disproportionate negative selection into virtual charter schools (Beck et al., 2014; Bueno, 2020; Paul and Wolf, 2020). Recent research concludes that even controlling for prior achievement does not sufficiently account for this negative selection (Paul and Greene, 2022), a finding that supports the theory "that parents choose to enroll their children in

[virtual schools] because of problems or 'shocks' experienced in their previous school that might be connected to drops in student performance." (Lueken et al., 2015, p. 328).

The influx of students into virtual schools during the COVID-19 pandemic provides a unique opportunity to shed further light on what type of students typically select virtual charters. In theory, students enrolling due to COVID-19 should closely resemble the general population of brick and mortar students on observable and unobservable characteristics because they are enrolling due to extrinsic (i.e., a pandemic) rather than intrinsic forces (e.g., social emotional challenges). To that end, academic differences between the "COVID cohort" and other virtual charter students might offer clues regarding the differences between virtual charter students and public school students generally. In this study, we use data from a large cyber charter network ("Countrywide Cyber") to assess whether students who enrolled in full time virtual schools due to COVID-related concerns performed better on entry diagnostic assessments.

COVID-19 PANDEMIC AND CYBER SCHOOLS

The worst pandemic since the 1918-1920 flu outbreak, COVID-19 has killed over one million Americans and thoroughly disrupted all facets of life, including schooling. There is already significant literature addressing how the pandemic has affected in person public schools, the majority of which shifted to temporary remote learning. Yet there is scant literature regarding COVID-19's impacts on preexisting online schooling. Molnar et al. (2019) report that as of the 2017-2018 school year 501 full-time virtual schools enrolled approximately 297,000 full time students. Charter schools accounted for 79.1% of enrollment.

The rapidly growing research literature on the impact of COVID-19 on K-12 schooling has a noted gap: no prior work focuses on how COVID-19 affected cyber charter schools. Prior empirical work finds that cyber charter schools have lower academic value added than both charter and traditional public in person schools (e.g., Woodworth et al., 2015), and that artificial testing conditions only play a marginal role in explaining this gap (Beck et al., 2018; Kingsbury et al., 2020). Some work suggests that lower cyber charter performance may in part reflect student composition, with students whose needs are not met in in-person settings disproportionately choosing cyber options, but also relatively more in need of the in-person support which virtual schools have difficulty providing (Ahn and McEachin, 2017; Paul and Greene, 2022).

The COVID-19 pandemic provides an opportunity to better understand whether the observed performance of cyber charter schools is explained by school performance or negative selection of students into cyber schools. Rather than enrolling due to negative shocks or other circumstances that negatively predict achievement, the "COVID cohort" believed that dedicated virtual programs (district or charter) could deliver a better virtual education than brick and mortar schools that switched to emergency remote learning (Flanders, 2021). Recent research by Maranto et al. (2021) compared parent survey responses from newly enrolled students entering a large national cyber charter school network in Spring 2020, during the pandemic, with parents of students entering in 2019 and 2018, before the pandemic, and found that the COVID cohort resembled prior groups demographically but reported a substantially lower incidence of bullying, mental health issues, and physical health issues as reasons for enrollment. However, the only performance metrics that the study compared were grade point average and curriculum-based assessments completed. While both metrics hint at stronger performance from the COVID cohort, neither metric is standardized, and the study only reports averages, omitting analysis of where changes occurred along the distribution of outcomes. Our analysis fills a gap by using a standardized performance metric and assessing where changes occurred along the distribution of achievement outcomes. Our analysis can inform the degree to which cyber charter observed performance reflects school quality rather than student composition.

METHODS

Data was provided by a large education management organization (EMO) that manages tuition-free virtual charter schools across the United States. Student records contained standard demographic information as well as information about how students performed on their beginning of year 2020-21 STAR and NWEA tests, third-party computer-adaptive assessments used to diagnose learning levels. Both assessments are widely used in American public schools and have been deemed reliable and valid diagnostic tests (Bulut and Cormier, 2018; Institute for Education Sciences, n.d.). Student-level de-identified scores were provided as national proficiency ranks, a percentile rank for performance on each subject test compared to the universe of American students who participated in the same grade-level test. Data was provided for students who enrolled between March 14, 2020 and September 9, 2020. March 14 is one day after President Trump declared a national emergency in response to COVID-19, whereas the latter date represents the beginning of the 2020-21 school year for all schools served by the management organization. Overall, 95.2% of eligible students participated in 2020-2021 beginning of year NWEA and STAR assessments. English language arts (ELA) tests are dispensed to students in kindergarten through 12th grade (i.e., all primary years in American public education, from about age 5–18) whereas math tests are dispensed to students from 3rd through 12th grade. Though it is not clear why 4.8% of the eligible student body did not participate, missing data is unlikely to pose a threat to the validity of our analysis, as it would only bias our estimates if it was missing in a way that correlated with both whether enrollment was tied to COVID-19 and observed achievement, an unlikely scenario.

Critically, the EMO polled families in Fall 2021 asking them to assess the importance of several factors in their decision to enroll. Specifically, the survey prompts respondents that "Below are some reasons that parents have said they chose [School Name] for their child. For each, please indicate how well it describes why you

TABLE 1 | Demographics of those who enrolled in cyber charters depending on COVID as enrollment factor.

	Concerned about COVID (%)	Not concerned about COVID (%)	Full sample (%)
African American	18.7	19.7	19.3
Asian	2.4	2.8	2.6
Hispanic	10.0	10.4	10.2
White	61.0	60.0	60.4
Special Ed	16.5	17.3	17.0
FRL	59.1	59.0	59.0

chose [School Name] for [Student Name]." Respondents respond on a Likert scale with ranges from 1 (very unimportant or strongly disagree) to 5 (very important or strongly agree). Among the 16 potential reasons provided is "concerns related to COVID-19." Adjoined to student-level assessment data, these responses allow us to gauge the extent to which the "COVID cohort" (i.e., those who reported that COVID-19 was an important enrollment factor) profiled differently from the traditional population served by cyber charter schools. Overall, responses were collected for 9,091 students, representing 34.3% of students who enrolled during that time. Among those, math scores were available for 7,243 students and ELA scores for 8,981.

RESULTS

We begin by observing the degree to which the "COVID cohort" profiles differently from other cyber charter students by conducting two-sided *t*-tests on demographic variables, as seen in **Table 1**.

Overall, parents of students who responded that concerns about COVID-19 were "important" or "very important" regarding their decision to enroll profile very similarly in their demographic composition to parents who responded that COVID-19 was unimportant or very unimportant. Indeed, none of the *t*-tests indicated statistically significant differences.

Next, we employ the regression model below to gauge differences in the two groups of students according to academic preparedness.

$$Y_i = \beta 0 + \beta 1 COVID_i + \beta 2 X_i + \beta 3 D_i + \beta 4 Q_i + e_i$$

Y_i represents student percentile ranks on ELA or math test scores on their 2021 beginning of year NWEA or STAR assessment. Coefficient β1 denotes responses to the Likert scale question about the importance of COVID as an enrollment factor (1 = very unimportant, 5 = very important). X_i is a vector of indicator variables for each school managed by the EMO, D_i is a vector of demographic characteristics,[1] and Q_i is a vector of responses to other survey items that were asked alongside the question about COVID.[2]

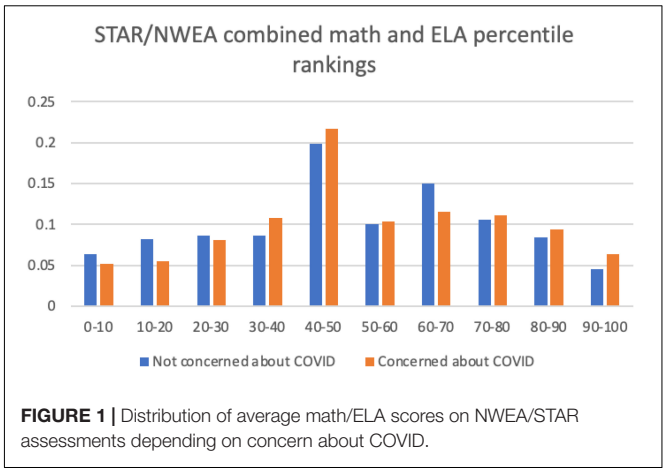


FIGURE 1 | Distribution of average math/ELA scores on NWEA/STAR assessments depending on concern about COVID.

Overall, results (**Table 2**) indicate that students whose enrollment was motivated by COVID were academically stronger, and that this difference was practically and statistically significant. Results were especially pronounced regarding ELA scores. For example, in the unadjusted model, a one-point increase on the Likert scale COVID question is associated with a 1.19 point increase in national percentile rank. Notably, the difference in the two groups of students is driven by differences on both tails of the distribution, as seen in **Figure 1**. That is, students enrolling due to COVID were more likely to be at the high-achieving end of the distribution and less likely to be at the low-achieving end. Illustratively, in averaging math and ELA percentile ranks, 14.6% of students not concerned about COVID scored between the 1st and 20th percentile compared to 10.6% of students concerned about COVID. Meanwhile, 12.8% of students not concerned about COVID were drawn from the top quintile compared to 15.7% of students concerned about COVID.

[1]The demographic control variables are the same variables that appear in **Table 1**.

[2]The EMO that provided data has requested that we not reproduce the full list of survey items, though generally speaking these 16 survey items cover a rich variety of factors and all responses are expressed on a 1–5 Likert scale reflecting responses from strongly disagree to strongly agree or very unimportant to very unimportant.

TABLE 2 | Concern about COVID as predictor of beginning of year percentile rank on NWEA/STAR assessments.

	ELA					Math				
COVID	1.19*** (0.22)	0.93*** (0.21)	1.25*** (0.26)	0.68*** (0.22)	0.49** (0.25)	0.74*** (0.23)	0.69*** (0.22)	0.67** (0.27)	0.33 (0.22)	0.12 (0.26)
School FE	N	Y	N	N	Y	N	Y	N	N	Y
Demographics	N	N	Y	N	Y	N	N	Y	N	Y
Surveys	N	N	N	Y	Y	N	N	N	Y	Y
n	8,981	8,981	6,452	8,981	6,452	7,243	7,243	5,207	7,243	5,207

****p* < 0.01; ***p* < 0.05; **p* < 0.10.

DISCUSSION

Our findings indicate that students who enrolled in cyber charters due to COVID-19 were better academically prepared than students who enrolled for other reasons. The pandemic and associated switches to emergency remote learning plausibly invited a representative cross-section of students to enroll in cyber charters, including families who anticipated higher-quality virtual learning in devoted cyber schools and others who had concerns about COVID-19 or associated mitigation measures in schools with in-person learning. To that end, the results support evidence that students are often negatively selected into cyber charter schools (Paul and Greene, 2022).

Still, our findings do not put to rest discussion surrounding the degree to which cyber charter performance reflects school composition rather than quality. To be sure, we cannot be certain of the true representativeness of the "COVID cohort." While our analysis of observed differences in learning levels between students who enrolled due to COVID versus other students hints at authentic differences, we do not have enough information about the representativeness of the COVID cohort to conclude that our estimates represent the discrete differences between students in the cyber charter sector versus students in brick and mortar schools.

Another notable limitation is that we did not have the data to assess differences in academic growth according to whether students enrolled due to COVID-19, as much of the criticism of the performance of virtual charter schools is directed not only at low test scores, but low observed year-over-year student growth. However, even with longitudinal assessment data, it's not clear that the COVID-19 enrollment surge provides an instructive counterfactual to assess student growth in virtual charters. To the degree that the disruption caused by COVID-19 is a shock that likely impacts the academic performance of the COVID cohort downstream, it is dubious whether the growth of the COVID cohort provides an instructive counterfactual for academic growth within virtual charters.

CONCLUSION: NORMALIZING CYBER SCHOOLS?

Prior research indicated that many parents who chose cyber schools did so out of serious dissatisfaction at their in-person schools, often due to social factors like bullying, but also due to concerns that the child's academic needs were unmet at their in-person school. This was particularly true for students with special education needs (Beck et al., 2014). These student-related factors may be among the drivers of the relatively weak cyber charter measured academic performance found in many studies (Lueken et al., 2015; Ahn and McEachin, 2017; Paul and Greene, 2022). The National Cyber data indicates that historical struggles characterized the COVID cohort to a lesser degree than prior cohorts.

In short, in the COVID era a greater percentage of new entrants to cyber schools may be those who were thriving rather than struggling at in person schools, but who were dissatisfied by their traditional public schools' adjustments to COVID. Survey research indicates that nationally, such parental dissatisfaction may reflect relatively ineffective implementation of hastily prepared online learning options in traditional schools (Henderson et al., 2021; Kingsbury, 2021).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

IK spearheaded data collection and analysis, and authored the results section. DB wrote the literature review. MB-D reviewed the analysis and revised the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

Ahn, J., and McEachin, A. (2017). Student enrollment patterns and achievement in Ohio's online charter schools. *Educ. Res.* 46, 44–57. doi: 10.3102/0013189X17692999

Beck, D., Egalite, A., and Maranto, R. (2014). Why they choose and how it goes: comparing special education and general education cyber student perceptions. *Comput. Educ.* 76, 70–79. doi: 10.1016/j.compedu.2014.03.011

Beck, D., Watson, A., and Maranto, R. (2018). Do testing conditions explain cyber charter schools' failing grades? *Am. J. Distance Educ.* 33, 46–58. doi: 10.1080/08923647.2019.1554989

Bueno, C. (2020). *Bricks and Mortar vs. Computers and Modems: The Impacts of Enrollment in K-12 Virtual Schools*. Available Online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3642969 (accessed March 1, 2022).

Bulut, O., and Cormier, D. (2018). Validity evidence for progress monitoring with Star reading: slope estimates, administration frequency, and number of data points. *Front. Educ.* 3:68. doi: 10.3389/feduc.2018.00068

Clark, T. (2000). *Virtual High Schools: State of the States - A Study of Virtual High School Planning and Preparation in the United States: Center for the Application of Information Technologies*. Macomb, IL: Western Illinois University.

Finn, C., Manno, B., and Wright, B. (2016). *Charter Schools at the Crossroads*. Cambridge, MA: Harvard Education Press.

Flanders, W. (2021). Opting out: enrollment trends in response to continued public school shutdowns. *J. Sch. Choice* 15, 331–343. doi: 10.1080/15582159.2021.1917750

Henderson, M., Houston, D., Peterson, P., and West, M. (2021). *What American Families Experienced When COVID-19 Closed Their Schools*. Education Next. Available online at: https://www.educationnext.org/wp-content/uploads/2022/01/ednext_XXI_1_poll_pandemic_henderson_et_al.pdf

Institute for Education Sciences (n.d.). *The Impact of the Measures of Academic Progress on Differentiated Instruction and Student Achievement*. Available Online at: https://ies.ed.gov/ncee/edlabs/projects/rct_245.asp?section=ALL

Kingsbury, I. (2021). Online learning: how do brick and mortar schools stack up to virtual schools? *Educ. Inf. Technol.* 26, 6567–6588. doi: 10.1007/s10639-021-10450-1

Kingsbury, I., Maranto, R., and Beck, D. (2020). Road weary? Testing whether long commutes to testing sites explain deficient cyber charter school academic

performance. *J. Sch. Choice* 14, 471–481. doi: 10.1080/15582159.2020.1845526

Lueken, M., Ritter, G., and Beck, D. (2015). Value-added in a virtual learning environment: an evaluation of a virtual charter school. *J. Online Learn. Res.* 1, 305–335.

Maranto, R., Beck, D., Clark, T., Tran, B., and Liu, F. (2021). Choosing cyber during COVID. *Phi Delta Kappan* 103, 30–33. doi: 10.1177/00317217211043622

Molnar, A., Miron, G., Elgeberi, N., Barbour, M. K., Huerta, L., Shafer, S. R., et al. (2019). *Virtual Schools in the U.S. 2019*. Boulder, CO: National Education Policy Center.

Paul, J., and Greene, J. (2022). *Investigating the Relationship between Negative Selection into Online Schooling and Achievement Growth. Education Reform Faculty and Graduate Students Publications*. Available Online at: https://scholarworks.uark.edu/edrepub/133 (accessed February 27, 2022).

Paul, J., and Wolf, P. (2020). *Moving on up? A Virtual School, Student Mobility, and Achievement. Annenberg Ed Working Paper No. 20-309*. Available Online at: https://www.edworkingpapers.com/sites/default/files/ai20-309.pdf (accessed February 21, 2022).

Saultz, A., and Fusarelli, L. (2017). Online schooling: a cautionary tale. *J. Sch. Choice* 11, 29–41. doi: 10.1002/hpja.203

Woodworth, J., Raymond, M., Chribas, K., Gonzalez, M., Negassi, Y., Snow, W., et al. (2015). *Online Charter School Study*. Stanford, CA: Center for Research on Education Outcomes.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Education

REVIEW
published: 16 June 2022
doi: 10.3389/feduc.2022.871973

Closing Achievement Gaps Through Preschool-to-Third-Grade Programs

*Judy A. Temple[1], Suh-Ruu Ou[2] and Arthur J. Reynolds[2]**

[1] Human Capital Research Collaborative, Humphrey School of Public Affairs, University of Minnesota, Twin Cities, Minneapolis, MN, United States, [2] HCRC, Institute of Child Development, University of Minnesota, Twin Cities, Minneapolis, MN, United States

Achievement gaps by family income, race, and ethnicity have persisted for decades. Yet only in recent years has this major social problem become a national priority in the United States and many other countries as concern rises over growing economic inequalities. In this article, we document gaps in school readiness and achievement in the United States and how they adversely affect the life course development of children and families from underrepresented groups. We emphasize the promising role of preschool-to-3rd grade (P-3) programs to reduce a variety of achievement gaps through comprehensive strategies that enrich educational and family experiences during most of the first decade of life. Implementation of the core elements of effective learning experiences, collaborative leadership, aligned curriculum, parent involvement and engagement, professional development, and continuity and stability in the Child-Parent Center (CPC) Program have shown relatively strong and sustained effects on school achievement, especially for Black children growing up in urban poverty. This evidence from the Chicago Longitudinal Study (CLS) and other projects suggests that broader scale up of truly comprehensive approaches that begin early, continue through most of the first decade, and are multilevel in scope can make a bigger difference than many existing strategies in reducing achievement gaps and their persistence.

Keywords: achievement gap, school readiness, poverty, child development, evaluation

OPEN ACCESS

Edited by:
Sebastian Urquijo, CONICET Institute of Basic, Applied and Technology Psychology (IPSIBAT), Argentina

Reviewed by:
Valeria Abusamra, University of Buenos Aires, Argentina
Jeffrey Roth, University of Florida, United States

****Correspondence:***
Arthur J. Reynolds
ajr@umn.edu

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

Received: *08 February 2022*
Accepted: *30 March 2022*
Published: *16 June 2022*

Citation:
Temple JA, Ou S-R and Reynolds AJ (2022) Closing Achievement Gaps Through Preschool-to-Third-Grade Programs. Front. Educ. 7:871973. doi: 10.3389/feduc.2022.871973

INTRODUCTION

Gaps in various measures of academic performance exist at school entry, whether researchers are measuring differences in student outcomes across racial or ethnic groups or comparing students with different levels of family income. In the United States, evidence from nationally representative analyses indicate that Black and Hispanic students have lower scores than white students in reading and math at school entry, and lower-income students from all races on average are less prepared for school compared to students from families with higher incomes. While much public attention has been paid to the existence and persistence of racial gaps in school readiness, the gaps between lower- and higher-income children may be even twice as large as racial gaps (Reardon and Portilla, 2016). A major report from the National Academies of Sciences, Engineering and Medicine in the

United States highlighted the existence of these gaps as indicative of significant societal inequities. The authors of the report called for more research to both track the inequities in outcomes and in access to opportunities for those most affected by the gaps (National Academies of Sciences, Engineering, and Medicine, 2019).

Differences in school performance by family income are found across the world. One strand of the intergenerational mobility literature has focused on how the magnitude of test score gaps for a given amount of income inequality varies across countries. Bradbury et al. (2019) compare differences in cognitive development occurring as early as age 5 by income across the United Kingdom, Australia, and the United States and report that for a given income difference, early test scores exhibit significantly larger gaps by income in the United States. Compared to these other countries, the same degree of inequality in incomes in the United States translates into larger inequities in child development outcomes at early ages.

In the past two decades, recognition of the importance of early childhood education has led governments in many countries including the United States and England to increase spending for preschool. In the United States, state governments have expanded access to preschool, mostly for 4 year olds but also for 3 year olds in some cases Friedman-Krauss et al. (2021). While in England the Childcare Act of 2006 requires local authorities to provide preschool to all children free of charge, in the United States the federally funded program Head Start is offered to children from low-income families. States and local governments also can choose to fund preschool, and where these programs exist, may be targeted (available to only those children from low-income families) or universal, depending on where a child resides. In this article, we focus primarily on the United States experience.

Political and economic constraints restrict full access to high- quality early education. In the United States, the National Institute for Early Education (NIEER) has presented a plan for a gradual roll-out eventually resulting in full access to state-funded preschool expansion by the year 2050 (National Institute for Early Educational Research, 2021). While school readiness gaps are large and persist over time, observers have noted that the modest reduction in the gaps by income or race that has occurred in recent years may be due in part to the expansion of publicly funded preschool (Reardon and Portilla, 2016; Bassok and Latham, 2017; Kuhfeld et al., 2020). But in their examination of early test score data by race and income over time, Reardon and Portillo's calculations suggest that closing the school readiness gap might take 100 years to achieve.

While the effects of preschool programs have been well studied (e.g., Cascio, 2022) and access to good-quality programs has expanded, some researchers and educators have turned to examining the topic of dosage. One question is whether publicly funded preschool should be offered for 1 vs. 2 years (Arteaga et al., 2014; Wasik and Snell, 2019). Other researchers and practitioners have focused on another question of duration. Should we consider the importance of incorporating the early years of formal schooling through a preschool through third grade model of education programming? In this article we focus on the potential for a more extended program of early intervention to have an impact on both school success in the shorter term in terms of reading scores and describe the potential for an extended early intervention to positively affect important long-term outcomes such as educational attainment and earnings.

NATIONAL EVIDENCE FROM UNITED STATES ON ACHIEVEMENT GAPS IN ELEMENTARY SCHOOL

While gaps in kindergarten readiness have remained significant for decades between children of different races and between children living in lower- and middle-income families, an examination of national data in the middle years of elementary school suggest that formal schooling in the early years has not reduced these gaps. As children make their way through elementary school, the existence of later gaps in achievement can be examined by looking at scores from the National Assessment of Educational Progress (NAEP). Sometimes called the Nation's Report Card, NAEP is a nationally representative assessment of reading, mathematics and science that public and private school students in the United States take starting in 4th grade (National Center for Education Statistics, 2019). While each state may have its own assessment tools for evaluating its public school students, the United States Congress has mandated since the 1960s that the United States Department of Education use a standard assessment nationwide to provide information on how well students in individual states as well as the nation are doing. The tests are administered to students in randomly selected schools in grades 4, 8, and 12. A common way of reporting results to show the percent of students in a particular grade who reach a proficiency threshold in reading, science or mathematics.

Figure 1 below reports the percent of United States students who have achieved a score in reading that is considered by the NAEP to represent proficiency. For comparison, the numbers are shown for students who are living in low-income families versus all other students. As school districts in the United States do not commonly collect detailed information on family income, the student's eligibility for a federal lunch subsidy (Free or Reduced Price Lunch or FRPL) is used as a proxy for low-income. According to the United States Digest of Education, the most recent data from 2016 to 2019 indicate that approximately 52% of public school students are eligible for the lunch subsidy (National Center for Education Statistics, 2020, Table 204.10).

In **Figure 1** it is clear that there are significant and persistent gaps in the percentage of children who are considered proficient in reading as of fourth grade. While the earlier discussion suggested that gaps at school entry might be modestly smaller in recent years (perhaps due in part to the expansion of publicly funded preschool education), here we see over the years 2005–2019 that the gap in fourth grade does not appear to have diminished and in every time period the share of children from lower-income families whose scores exceed the proficiency threshold is less than half of the proficiency rate of children from middle- and higher-income families. Importantly, the gap appears to have widened over time. While the proficiency rate

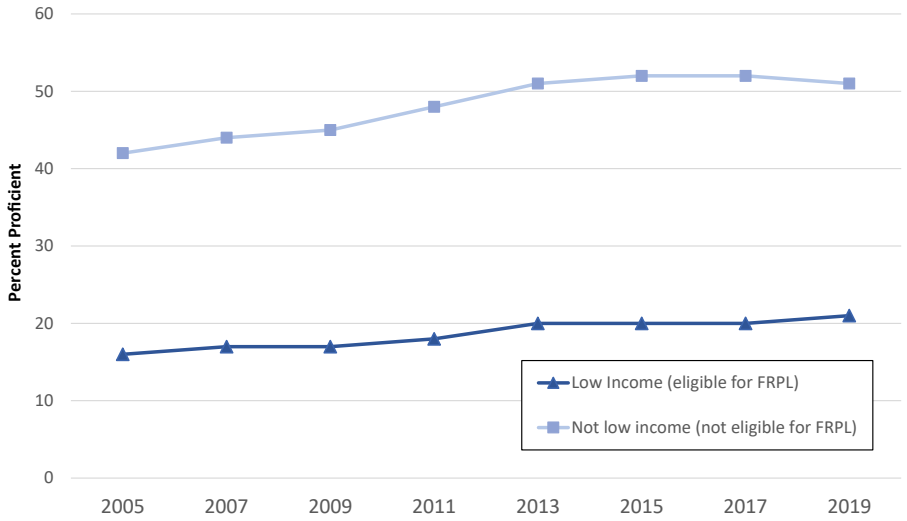


FIGURE 1 | Reading proficiency of United States 4th graders by year and family income status. Percent proficient refers to students who met or exceed the reading proficiency threshold in the National Assessment of Educational Progress (NAEP) by eligibility for free or reduced price lunch subsidy, an indicator of low-income status.

for students from families not eligible for a federal lunch subsidy increased from 42 to 51%, the improvement for students from low-income families improved from 16 to 21%. Although not shown here, these gaps across income groups also exist in the NAEP for mathematics and science.

In the remainder of the paper, we discuss the educational model of preschool to 3rd grade programming, a form of early intervention that provides enriched educational services beyond preschool into the early elementary grades (Reynolds, 2019). Preschool-to-3rd grade programs have been discussed in education practitioner and policy circles for several decades. One primary example of such a program is the Child-Parent Center (CPC) program offered in the Chicago Public Schools serving children and parents in high-poverty neighborhoods.

CHILD-PARENT CENTER PreK-3rd GRADE PROGRAM

The CPC are a preschool-to-3rd grade (P-3) program. Ideas for P-3 as a service continuum evolved during the early years of the War on Poverty/Great Society era of the mid 1960s (Zigler et al., 2006). Fundamental to P-3 is developmental continuity. This is the extent to which learning environments are consistent and predictable over time in promoting well-being, especially during transitions. This continuity provides a P-3 advantage, which is the added benefit of continued services above and beyond earlier experiences. Many studies show such an advantage (Reynolds, 1994; Ou and Reynolds, 2006; Zellman and Kilburn, 2015; Manship et al., 2016; Takanishi, 2016) and they demonstrate that P-3 can close achievement gaps and strengthen learning gains. Unique to the CPC model, however, is that the P-3 early education program model is a comprehensive model with key requirements on site leadership, class size, and parent involvement. Researchers have examined long-term effects into adulthood (Reynolds et al., 2017).

The CPC opened in 1967 in Chicago through funding from Title I of the Elementary and Secondary School Act of 1965, a key component of the War on Poverty. While the federal preschool program Head Start was just starting to roll out, many areas remained unserved, and the Chicago Public Schools chose to make use of a different federal funding stream to implement this program. Although the CPC program began as a comprehensive preschool program, children received continuing services in kindergarten and the early grades the following year, leading to the current configuration. Under the direction of a leadership team at each site and in collaboration with the Principal, CPC-P3 enhances school readiness skills, increases early school achievement, and promotes parent involvement (Ou and Reynolds, 2006; Reynolds et al., 2011a, 2016). The main criterion for CPC program enrollment is residence in a low-income neighborhood eligible for federal Title 1 funding (Kainz, 2019). Other enrollment criteria include family income, parent education, and previous experience in early childhood programs, with priority given to those with greater disadvantage (Reynolds, 2000). Over 90% of CPC and comparison group members resided in families with incomes below 185% of the federal poverty line (Reynolds et al., 2018).

Figure 2 shows the continuity inherent in the CPC preschool to third grade model in its equal emphasis on preschool, kindergarten, and the early grades. Early education provides the foundation and the next few grades build on this to promote achievement and well-being. Some dimensions of adult well-being investigated in a longitudinal study following a large cohort of CPC participants as well as a matched comparison group include income, employment, justice system involvement, and physical and mental health (e.g., Reynolds et al., 2019;

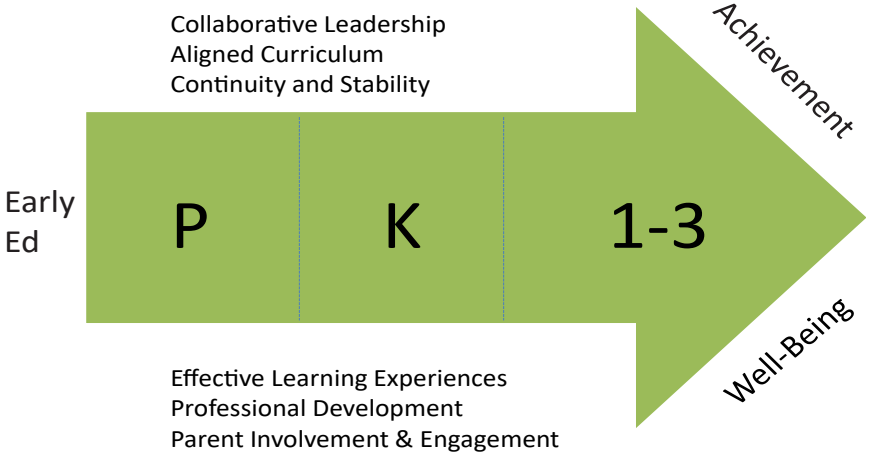


FIGURE 2 | CPC preschool to 3rd grade process. Key components of the school reform model with the dual goals of achievement and overall well-being. Early education begins at ages 3 or 4 with continuing services through 3rd grade. The six core elements are key to promoting developmental continuity.

Varshney et al., 2022). The figure also illustrates the inherent tension between the early timing of intervention versus duration of services. Early participation in programs is a frequent focus of impact efforts but usually does not address later stages of development. Longer duration programs rarely begin early enough in childhood. The CPC's preschool – third grade approach represent both dimensions, but a major focus of dissemination is promoting continuity in learning to realistically narrow achievement gaps.

The six elements noted in **Figure 1** are as follows:

1. *Collaborative leadership team* run by the Head Teacher in partnership with the Principal, Parent Resource Teacher, and School-Community Representative.
2. *Effective learning experiences* through small classes (17 or fewer), engaging instruction, and increased instructional time (e.g., full-day preschool).
3. *Aligned curriculum* is an organized and documented sequence of evidence-based instructional practices that build on prior learning and are supported by teacher collaboration across grades.
4. *Parent involvement and engagement* is a menu-based set of services led by the PRT and SCR and utilize a parent resource room. An involvement plan explicates the key elements.
5. *Professional development system* that combines on-site facilitation and on-line professional learning modules (e.g., STEM, thinking skills).
6. *Continuity and stability* includes co-located or close-by centers that provide year-to-year consistency in implementation.

Beginning at ages 3 and 4, children participate in small classes through 3rd grade and each class has an assistant for at least half of the day. The learning environment created by the principal and team provides an integrated context for improved achievement and sustained gains. Transitions from year to year are supported by the parent involvement team, site mentors, and school staff, who share instructional approaches and teaching practices across grades. Teachers are state-licensed and follow an instructional plan with a nearly equal mix of teacher-directed and child-initiated activities supported further by classroom assistants. Curriculum alignment and parent involvement plans are reviewed and updated annually. Professional development includes on-line teaching modules and on-site coaching of instructional practices to support a balance of teacher- and child-initiated instruction. Outreach services, including home visits and workshops, use a menu-based system informed by needs assessments conducted with the parents.

CHILD-PARENT CENTER PROGRAM BENEFITS

The positive effects of CPC are well documented in the Chicago Longitudinal Study (CLS), which tracks well-being over the life course of an early childhood cohort of 1,539 children growing up in high poverty neighborhoods (Reynolds, 2000; Reynolds et al., 2011a). They were born in 1979–1980 and entered kindergarten in 25 schools in the fall of 1985. In this matched-group, quasi-experimental design, 989 3- and 4-year-olds from in 20 CPCs were compared to 550 children of the same age who enrolled in the usual early childhood programs in five randomly selected schools also eligible for federal financial assistance due to high rates of neighborhood poverty. A broad range of measures of well-being have been collected for over three decades with over 90% of the original sample remaining in the study.

Child-Parent Center participants show consistent performance advantages in school achievement, need for remediation, delinquent behavior, and educational attainment through high school and college (Reynolds and Ou, 2011; Reynolds et al., 2018). Some evidence on reading achievement over time is presented in **Figure 3**. In this figure, developmental standard scores in reading achievement on the Iowa Test of

Basic Skills for CPC participants who participated in at least the preschool component of the intervention are compared to evidence on national norms as well as the reading scores from the no-preschool comparison group. For the CPC treatment and control groups, scores shown are adjusted for adjusted for child and family demographic attributions (e.g., family education, income, child gender and race; see Reynolds and Temple, 1998). While CPC participant and controls were well matched on many socio-economic characteristics, by kindergarten the CPC participants have an advantage over students in the comparison group. Participation in the CPC preschool program helps close the gap in test scores in the early years of schooling, although the evidence suggests that this gap then widens somewhat between the national average and the CPC students and non-CPC controls by age 15.

Although both the CPC participants and comparison group students came from low-income families residing in some of the city's poorest neighborhoods, the CPC participants are able to maintain their advantage over the comparison group members and are able to close half of the achievement gap between non-CPC preschool students and national norms during elementary school before the gap between the program participants and national norms start to diverge. While participation in the CPC program into the early grades helps sustain the initial gains from participation in a strong preschool program, compared to national norms the program participants continue to live in economically disadvantaged neighborhoods often with fewer available parental and school resources. **Figure 3** also includes estimates of the effect sizes associated with participation in CPC preschool. Students who had at least some CPC preschool perform better than the similar no-CPC preschool comparison group members after controlling for a rich set of family, neighborhood and child characteristics measured at the time of the child's birth. These effect sizes ranged from 0.62 to 0.25 as children progressed through school.

What if early interventions lasted longer, perhaps into third grade? **Figure 4** shows the growth in reading achievement from kindergarten entry (age 5) to 4th grade (age 10) for the CPC group participating in the entire program for 4–6 years (P-3) compared to the group without continuing services (P + K only). The comparison group from **Figures 2**, **3** is not included here. As in **Figure 3**, reading scores from the Iowa Tests of Basic Skills are shown and are adjusted for child and family demographic characteristics.

While as previously mentioned students in the United States from low-income households start kindergarten well below national norms, all the CPC students shown in **Figure 4** participated in the intervention in preschool and kindergarten. As a result, their scores resemble the national averages at the end of kindergarten. This figure also shows that participation in preschool through third grade portion of the CPC program cuts the test score gap in half by age 10 compared to students who participated in preschool and kindergarten only. The national norm scores shown have means of 60, 78, and 108, respectively, for kindergarten, first grade, and third grades. As shown, although growth during kindergarten was similar between groups and at/above national norms, the CPC-P3 group experienced greater growth between 1st and 4th grades. This translates to about a 6-month gain above and beyond earlier participation. This advantage in performance reduced the

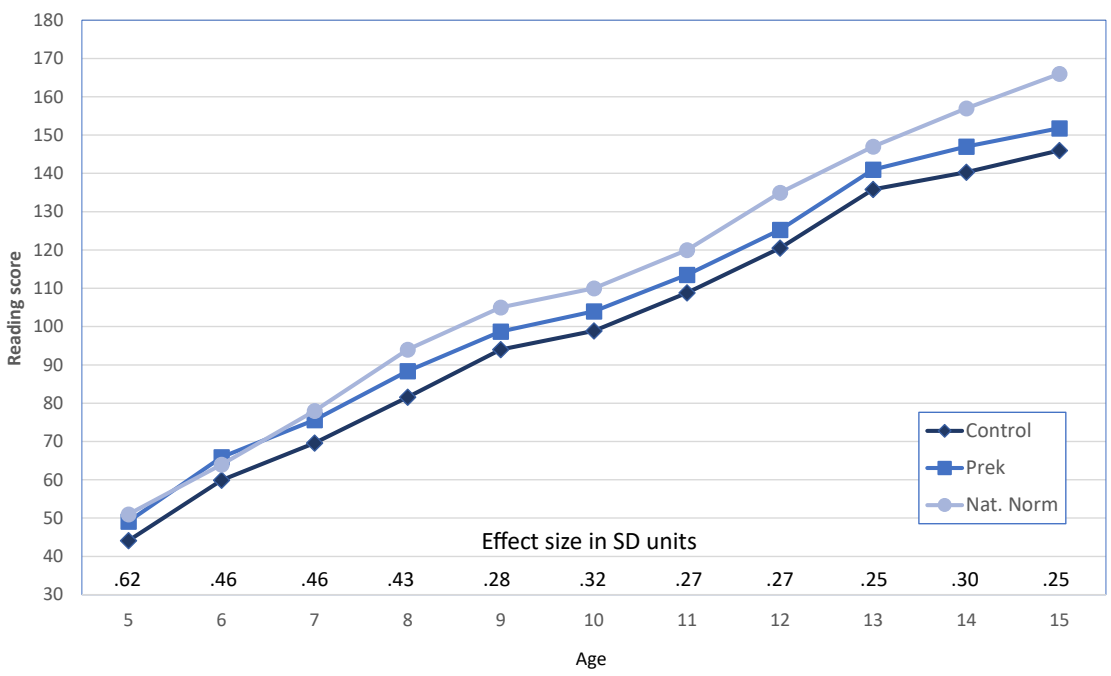


FIGURE 3 | Reading achievement for Child Parent Center (CPC) participants (with at least preschool) and others at ages 5–15. Reading test scores from Iowa Test of Basic Skills for CPC preschool program participants, non-CPC participants serving as controls, and national norms.

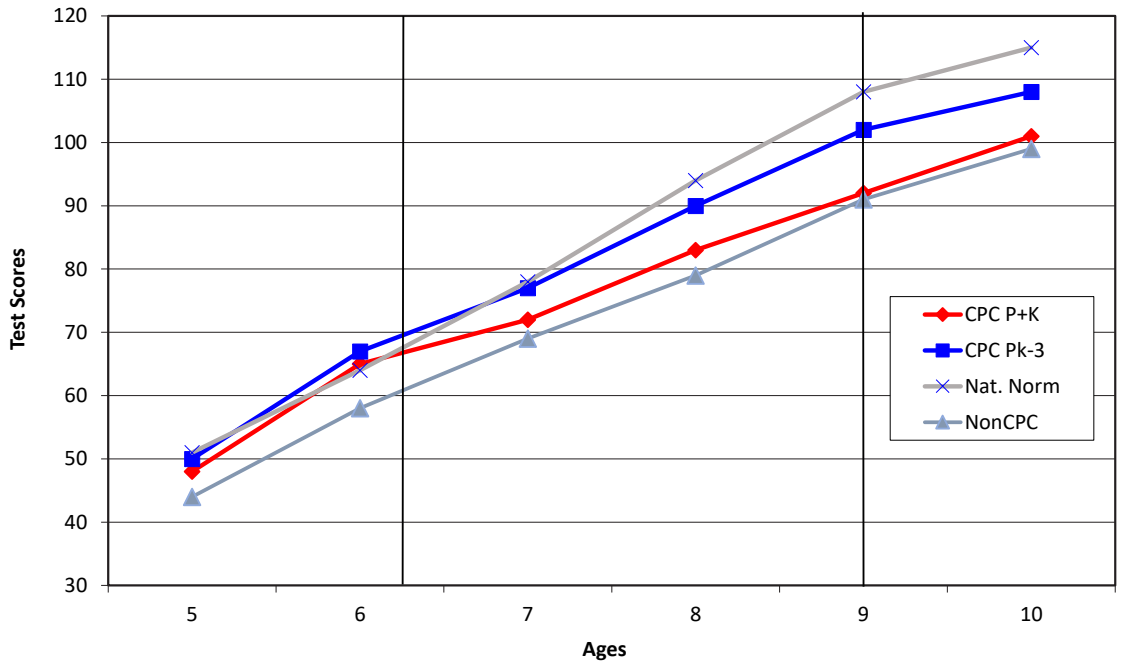


FIGURE 4 | Reading test scores for participants in the Child Parent Center program by length of participation at ages 5–10. Reading scores from Iowa Test of Basic Skills for CPC preschool plus extended intervention versus CPC preschool participants only and national norms.

gap with the national average by about 75% even though the average performance of the CPC-P3 group was not quite at the national average. The pattern of sustained effects in reading for CPC preschool is also evident. Students in the CLS who did not participate in the CPC program performed below all other groups after controlling for child and family demographic characteristics.

Paths through which gains are sustained over time has been documented in the Five-Hypothesis Model (Reynolds and Ou, 2011), including paths of cognitive advantage, motivational advantage, school quality and support, family support behavior, and socio-emotional adjustment. A distinctive feature of the CLS is the long history of investigating the validity and generalizability of findings. This includes analyses that account for participants lost to follow-up and differences in school experiences. For example, the original analyses of P-3 not only modeled growth in achievement due to the participation in the extended intervention with two pretests, but also assessed the impact of unobserved influences.

EVIDENCE ON LONGER-TERM OUTCOMES FOR EDUCATIONAL ATTAINMENT AND INCOME

The differential impacts of participation in extended preschool-to-third grade education programs versus participation in preschool and kindergarten only have been investigated in the CLS for educational attainment outcomes up to age 35 (Reynolds et al., 2018). The empirical analysis controlled for gender, race, and socio-economic characteristics observed from birth certificate records and early school administrative records. The use of inverse propensity score weighting resulted in some observations being weighted more heavily that others so that the analysis better resembled a randomized experiment. Similar corrections were also made to account for non-random attrition although sample retention remained high at over 80%. Overall, however, this reweighting had only a modest effect on estimation of effects due to the low attrition and the similarity of CPC participants to comparison group members on a rich set of background covariates including some measured at the time of birth. Some findings are reported in **Table 1** below where outcomes for CPC participants who participated in the preschool-to-third program for 4–6 years are compared to those who participated in CPC preschool and kindergarten only.

Overall, in this urban sample of minority students residing in high-poverty neighborhoods, high school graduation rates were low. But participation in the extended program of preschool into the early grades for 4–6 years seems to have made a difference above and beyond the effect of preschool and kindergarten by themselves. The adjusted rates of graduation (controlling for covariates and use of inverse propensity score weighting) indicate that almost 60% of students who participated in the CPC extended program graduated compared to 51% in both the preschool or no preschool comparison groups. In the United States, students are allowed to take the General Educational Development diploma (GED) at any time throughout their lives and in this sample many of those who did not formally graduate from high school

TABLE 1 | Predicted rates of educational attainment at age 35 in the Chicago Longitudinal Study.

	Preschool-to-third grade (4–6 years of intervention)	Preschool plus kindergarten only	Comparison group (0 years of participation in CPC program)
Educational outcomes by age 35%			
High school graduation %	59.1*#	50.7	50.6
Four-year high school graduation %	54.6*#	45.4	44.3
High school completion %	85.9#	84.5	81.4
Years of education (range 7–22)	12.9#	12.5	12.4
Any college attendance %	61.8#	57.3	54.1
Bachelors' degree or higher %	13.9*#	8.4	8.3

**Denotes statistical significance at 5% level for CPC enrollment for 4–6 years comparison to CPC preschool and kindergarten only. #Denotes significance at 5% level for CPC enrollment for 4–6 years comparison to students with 0 years of CPC enrollment. Rates shown are adjusted for gender, race and sociodemographic covariates. IPW is used to address non-random program assignment and attrition. GED holders are included in high school completion but not included in the high school graduation outcome. See Table 11 in Reynolds et al. (2018) for more detail.*

later acquired this credential. As a result, the high school completion rate, which includes the GED, is more similar for all groups although there was a statistically significance difference in the completion rate between the extended versus 0 years groups. A similar finding occurred for the outcome of "any college" attendance. In fact, a sizeable number of students entered the city college system, but persistence was an issue (Reynolds et al., 2011b). Importantly, however, the percentage of preschool-to-third grade participants who received a bachelors' degree was significantly higher than those with CPC preschool and kindergarten and those who did not participate in the intervention.

DISCUSSION AND CONCLUSION

Gaps in test scores and educational attainment are significant in large part because they help determine economic success throughout the life course. For students born into low-income families, education is the major opportunity for a child's upward mobility. Evidence on income mobility across generations indicates that many current young people in many countries may not attain their own parents' standard of living (Chetty et al., 2017; Manduca et al., 2020). Children born advantaged retain a large advantage at the end of early childhood, and the pattern persists in subsequent stages (Sawhill and Reeves, 2016). Evidence also indicates that the gaps in levels of educational attainment by parental income grow in the early years, through K-12, and into higher education (Duncan and Murnane, 2016). Black and economically disadvantaged populations stand out in the vulnerable populations because of the large and persistent gap in economic status between Blacks and Whites in the United States (Mazumder, 2014; Reardon and Portilla, 2016). Studies found that almost 50 percent of Black children born into the bottom 20 percent of the income distribution were in the same position as adults, but that only 23 percent of White children born in that quintile were (Timothy, 2016).

Research indicates that early cognitive and achievement advantages carry over to social and emotional competencies in middle childhood and adolescence, culminating in greater educational, economic, and social well-being in adulthood (Reynolds, 2000; Schweinhart, 2005; Karoly et al., 2006). A recent study (Ricciardi et al., 2021), for example, indicates that school readiness skills, both pre-academic readiness and socioemotional readiness, at age 4 have a long-term influence on academic performance through fifth grade and socioemotional readiness skills are an important component of school readiness. Improving school readiness skills at preschool can be an effective way to optimize students' chances of academic success although the positive effects on academic performance are not consistently found by late elementary grades.

While increasing access to high-quality preschool program is of great interest to educators and policymakers around the world, the current article explains and provides evidence relating to participation in an extended program of early intervention that provides continuity and alignment from preschool into the middle of elementary school. The test score trajectories shown in **Figures 3**, **4** suggest that early sustained intervention can help reduce the achievement gap between children living in poverty and the national average. Complete closing of the gap, no matter how good the intervention, is unlikely given the contributing role of poverty to children's daily experiences. Evidence suggests that post-preschool school quality can make a difference in sustaining early gains (Ansari and Pianta, 2018; Reynolds and Temple, 2019). The potential for school-based preschool programs to facilitate both collaborative leadership and curricula aligned vertically from preschool to kindergarten and beyond has been discussed in more detail by a number of authors including Bogard and Takanishi (2005); Kagan et al. (2006); Little (2020); and Justice et al. (2021). Evidence presented in this article suggests that preschool-to-third grade programs that intentionally combine the features of a collaborative leadership team, effective learning experiences, an aligned curriculum, parent involvement and engagement as well as an emphasis on a vibrant professional development system can make a difference in reducing achievement gains and helping to sustain gains from early interventions beginning in the preschool years.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

REFERENCES

Ansari, A., and Pianta, R. C. (2018). The role of elementary school quality in the persistence of preschool effects. *Child. Youth Serv. Rev.* 86, 120–127. doi: 10.1016/j.childyouth.2018.01.025

Arteaga, I., Humpage, S., Reynolds, A. J., and Temple, J. A. (2014). One year of preschool or two: Is it important for adult outcomes? *Econ. Educ. Rev.* 40, 221–237. doi: 10.1016/j.econedurev.2013.07.009

Bassok, D., and Latham, S. (2017). Kids today: the rise in children's academic skills at kindergarten entry. *Educ. Res.* 46, 7–20. doi: 10.3102/0013189x17694161

Bogard, K., and Takanishi, R. (2005). PK-3: an aligned and coordinated approach to education for children 3 to 8 years old. *Soc. Policy Rep.* 19, 1–24. doi: 10.1002/j.2379-3988.2005.tb00044.x

Bradbury, B., Waldfogel, J., and Washbrook, E. (2019). Income-related gaps in early child cognitive development: Why are they larger in the United States than in the United Kingdom, Australia, and Canada? *Demography* 367–390. doi: 10.1007/s13524-018-0738-8

Cascio, E. U. (2022). "Early Childhood Education in the United States: What, When, Where, Who, How, and Why," in *Handbook of the Economics of Education*, ed. B. McCall (London: Routledge), 30–72.

Chetty, R., Grusky, D., Hell, M., Hendren, N., Manduca, R., and Narang, J. (2017). The fading American dream: trends in absolute income mobility since 1940. *Science* 356, 398–406. doi: 10.1126/science.aal4617

Duncan, G. J., and Murnane, R. J. (2016). Rising inequality in family incomes and children's educational outcomes. *RSF* 2, 142–158. doi: 10.7758/rsf.2016.2.2.06

Friedman-Krauss, A. H., Barnett, W. S., Garver, K. A., Hodges, K. S., Weisenfeld, G. G., and Gardiner, B. A. (2021). *The State of Preschool 2020: State Preschool Yearbook*. New Brunswick, NJ: National Institute for Early Education Research.

Justice, L. M., Jiang, H., Purtell, K. M., Lin, T. J., and Ansari, A. (2021). Academics of the early primary grades: investigating the alignment of instructional practices from Pre-K to Third Grade. *Early Educ. Dev*. doi: 10.1080/10409289.2021.1946762.

Kagan, S. L., Carroll, J., Comer, J. P., and Scott-Little, C. (2006). Alignment: a missing link in early childhood transitions? *Young Child.* 61, 26–32.

Kainz, K. (2019). Early academic gaps and Title 1 programming in high poverty, high minority schools. *Early Child. Res. Q.* 47, 159–168. doi: 10.1016/j.ecresq.2018.08.012

Karoly, L. A., Kilburn, M. R., and Cannon, J. S. (2006). *Early Childhood Interventions: Proven Results, Future Promise*. Santa Monica, CA: Rand Corporation.

Kuhfeld, M., Soland, J., Pitts, C., and Burchinal, M. (2020). Trends in children's academic skills at school entry: 2010 to 2017. *Educ. Res.* 49, 403–414. doi: 10.3102/0013189x20931078

Little, M. (2020). Educators' views on the location of pre-k programs and its relations to features of P-3 alignment: an exploratory study. *Child. Youth Serv. Rev.* 18, 1–9.

Manduca, R., Hell, M., Adermon, A., Blanden, J., Bratberg, E., Gielen, A. C., et al. (2020). *Trends in Absolute Income Mobility in North America and Europe. IZA Institute of Labor Economics Discussion Paper 13456*. Bonn: IZA Institute of Labor Economics.

Manship, K., Farber, J., Smith, C., and Drummond, K. (2016). *Case Studies of Schools Implementing Early Elementary Strategies: Preschool through Third Grade Alignment and Differentiated Instruction*. Washington, DC: US Department of Education.

Mazumder, B. (2014). Black–white differences in intergenerational economic mobility in the United States. *Econ. Perspect.* 38, 1–18.

National Academies of Sciences, Engineering, and Medicine (2019). *Monitoring Educational Equity*. Washington, DC: National Academies Press.

National Center for Education Statistics (2019). *An Overview of NAEP. U.S Department of Education report NECS-2019-153*. Washington, DC: US Department of Education.

National Center for Education Statistics (2020). *Digest of Education Statistics, Table 204.10 Number and Percentage of Public School Students Eligible for Free or Reduced-Price Lunch, by State: Selected years, 2000-01 through 2018-19*.

National Institute for Early Educational Research (2021). *Funding High-Quality Preschool for All, March*. Available online at: https://nieer.org/wp-content/uploads/2021/04/Funding_High_Quality_PreK_for_All_.pdf (accessed April 21, 2022).

Ou, S.-R., and Reynolds, A. J. (2006). "School-age services: programs that extend the benefits of early care and education services," in *Evidence-Based Programs, Practices, and Policies for Early Childhood Care and Education*, eds C. J. Groark, K. E. Mehaffie, R. B. McCall, and M. T. Greenberg (Thousand Oaks, CA: Corwin press), 114–134. doi: 10.1111/j.1467-8624.2010.01563.x

Reardon, S., and Portilla, X. A. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *AERA Open* 2, 1–18.

Reynolds, A. J. (1994). Effects of a preschool plus follow-on intervention for children at risk. *Dev. Psychol.* 30, 787–804. doi: 10.1037/0012-1649.30.6.787

Reynolds, A. J. (2000). *Success in Early Intervention: The Chicago Child-Parent Centers*. Lincoln, NE: University of Nebraska Press.

Reynolds, A. J. (2019). The power of P-3 school reform. *Phi Delta Kappan* 100, 27–33. doi: 10.1177/0031721719834025

Reynolds, A. J., Hayakawa, M., Candee, A. J., and Englund, M. M. (2016). *CPC P-3 Program Manual: Child-Parent Center Preschool-3rd Grade Program*. Minneapolis, MN: University of Minnesota.

Reynolds, A. J., and Ou, S.-R. (2011). Paths of effects from preschool to adult well-being: a confirmatory analysis of the child-parent Center Program. *Child Dev.* 82, 555–582. doi: 10.1111/j.1467-8624.2010.01562.x

Reynolds, A. J., Ou, S.-R., Mondi, C. F., and Giovanelli, A. (2019). Reducing poverty and inequality through preschool-to-third-grade prevention services. *Am. Psychol.* 74, 653–672. doi: 10.1037/amp0000537

Reynolds, A. J., Ou, S.-R., and Temple, J. A. (2018). A multicomponent, preschool to third grade preventive intervention and educational attainment at 35 years of age. *JAMA Pediatr.* 172, 247–256. doi: 10.1001/jamapediatrics.2017.4673

Reynolds, A. J., Ou, S. R., Mondi, C. F., and Hayakawa, M. (2017). Processes of early childhood interventions to adult well-being. *Child Dev.* 88, 378–387.

Reynolds, A. J., and Temple, J. A. (eds) (2019). *Sustaining Early Childhood Learning Gains: Program, School, and Family Influences*. Cambridge: Cambridge University Press.

Reynolds, A. J., and Temple, J. A. (1998). Extended early childhood intervention and school achievement: age 13 findings from the Chicago Longitudinal Study. *Child Dev.* 69, 231–246. doi: 10.1001/jama.285.18.2339

Reynolds, A. J., Temple, J. A., Ou, S.-R., Arteaga, I. A., and White, B. A. B. (2011a). School-based early childhood education and Age-28 well-being: effects by timing, dosage, and subgroups. *Science* 333, 360–364. doi: 10.1126/science.1203618

Reynolds, A. J., Temple, J. A., White, B. A. B., Ou, S.-R., and Robertson, D. L. (2011b). Age 26 cost-benefit analysis of the child-parent center early education program. *Child Dev.* 82, 379–404. doi: 10.1111/j.1467-8624.2010.01563.x

Ricciardi, C., Manfra, L., Hartman, S., Bleiker, C., Dineheart, L., and Winsler, A. (2021). School readiness skills at age four predict academic achievement through 5th grade. *Early Child. Res. Q.* 57, 110–120. doi: 10.1016/j.ecresq.2021.05.006

Sawhill, I. V., and Reeves, R. V. (2016). Modeling EQUAL OPPORTUnity. *RSF* 2, 60–97. doi: 10.7758/rsf.2016.2.2.03

Schweinhart, L. J. (2005). *Lifetime Effects: the High/Scope Perry Preschool Study Through Age 40 (No. 14)*. Ypsilanti, MI: High/Scope Foundation.

Takanishi, R. (2016). *First Things First! Creating the New American Primary School*. New York, NY: Teachers College Press.

Timothy, M. S. (2016). Multiple barriers to economic opportunity for the "Truly" Disadvantaged and Vulnerable. *RSF* 2, 98–122. doi: 10.7758/rsf.2016.2.2.04

Varshney, N., Temple, J. A., and Reynolds, A. J. (2022). Early education and adult health: age 37 impacts and economic benefits of the Child-Parent Center preschool program. *J. Benefit-Cost Anal.*

Wasik, B. A., and Snell, E. K. (2019). "Synthesis of preschool dosage: how quantity, quality and content impact child outcomes," in *Sustaining Early Childhood Learning Gains: Program, School, and Family Influences*, eds A. J. Reynolds and J. A. Temple (Cambridge: Cambridge University Press).

Zellman, G. L., and Kilburn, M. R. (2015). *Final Report on the Hawai'i P-3 Evaluation*. Santa Monica, CA: RAND.

Zigler, E., Gilliam, W. S., and Jones, S. M. (eds) (2006). *A Vision for Universal Preschool Education*. Cambridge: Cambridge University Press.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 16 August 2022
DOI 10.3389/fpsyg.2022.943449

OPEN ACCESS

EDITED BY
Marcos Cupani,
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

REVIEWED BY
Angelica Moè,
University of Padua, Italy
Teresa Pozo-Rico,
University of Alicante, Spain

*CORRESPONDENCE
Chung-Ying Lin
cylin36933@gmail.com
Jeffrey Hugh Gamble
gamble@mail.ncyu.edu.tw;
mrjgamble@gmail.com

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

RECEIVED 24 May 2022
ACCEPTED 19 July 2022
PUBLISHED 16 August 2022

CITATION
Chen I-H, Chen X-m, Liao X-l, Zhao K-Y, Wei Z-H, Lin C-Y and Gamble JH (2022) Evaluating the immediate and delayed effects of psychological need thwarting of online teaching on Chinese primary and middle school teachers' psychological well-being.
Front. Psychol. 13:943449.
doi: 10.3389/fpsyg.2022.943449

Evaluating the immediate and delayed effects of psychological need thwarting of online teaching on Chinese primary and middle school teachers' psychological well-being

I-Hua Chen[1], Xiu-mei Chen[2], Xiao-ling Liao[3], Ke-Yun Zhao[4], Zhi-Hui Wei[5], Chung-Ying Lin[6,7,8]* and Jeffrey Hugh Gamble[9]*

[1]Chinese Academy of Education Big Data, Qufu Normal University, Qufu, China, [2]Faculty of Education, Qufu Normal University, Qufu, China, [3]International College, Krirk University, Bangkok, Thailand, [4]School of Communication, Qufu Normal University, Rizhao, China, [5]Department of Development and Research, Shanghai Open University, Shanghai, China, [6]Institute of Allied Health Sciences, College of Medicine, National Cheng Kung University, Tainan, Taiwan, [7]Department of Occupational Therapy, College of Medicine, National Cheng Kung University, Tainan, Taiwan, [8]Biostatistics Consulting Center, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan, [9]Department of English, National Changhua University of Education, Changhua, Taiwan

Recent studies on the effects of mandatory online teaching, resulting from the COVID-19 pandemic, have widely reported low levels of satisfaction, unwillingness to continue online teaching, and negative impacts on the psychological well-being of teachers. Emerging research has highlighted the potential role of psychological need thwarting (PNT), in terms of autonomy, competence, and relatedness thwarting, resulting from online teaching. The aim of this study was to evaluate the immediate and delayed (longitudinal) effects of PNT of online teaching on teachers' well-being (including distress and burnout), intention to continue online teaching, and job satisfaction. Moreover, data collected from both cross-sectional and longitudinal surveys allowed for a systematic validation of an important instrument in the field of teacher psychology, the Psychological Need Thwarting Scale of Online Teaching (PNTSOT), in terms of longitudinal reliability and validity. The data reveal the usefulness of the construct of PNT in terms predicting and explaining teachers' willingness to continue using online teaching as well as the degree of burnout after a period of 2 months, such that PNT is positively associated with burnout and negatively associated with willingness to continue online teaching. As such, the PNTSOT is recommended for future research evaluating the long-term psychological, affective, and intentional outcomes stemming from teachers' PNT. Moreover, based on our findings that the impact from PNT of online teaching is persistent and long-term, we suggest that school leaders provide flexible and sustained professional development, model respectful and adaptive leadership, and create opportunities for mastery for the development of community of practice that can mitigate the thwarting of teachers' autonomy, competence, and relatedness during

times of uncertainty. Additionally, in terms of the psychometric properties of the PNTSOT instrument, our empirical findings demonstrate internal reliability, test–retest reliability, measurement invariance, and criterion validity (concurrent and predictive) based on cross-sectional and longitudinal data.

KEYWORDS

psychological need thwarting, online teaching, longitudinal data, instrument validation, COVID-19 pandemic, teacher psychology, psychological well-being, burnout

Introduction

The COVID-19 pandemic has had a profound impact on the world, with pervasive effects in all aspects of life, including education. In fact, according to a survey by the United Nations Educational, Scientific and Cultural Organization (UNESCO), more than 180 countries had closed all school campuses during the pandemic, affecting the lives of 1.6 billion primary and secondary school students (UNESCO, 2020). This large-scale impact on a significant portion of the population has made the ongoing and lasting impacts of the COVID-19 pandemic of increasing concern to educational practitioners and researchers in terms of both short-term and long-term outcomes (Chen et al., 2021; Alhazmi et al., 2022; Haug et al., 2022). In response to the urgent demands and concerns of educators, several international organizations published special reports to provide the guidelines or frameworks for policy-makers, educators, and other stakeholders [e.g., The World Bank (Rodriguez et al., 2021), Organisation for Economic Co-operation and Development (Gouëdard et al., 2020), and UNESCO (2020)]. One common theme in these reports is an emphasis on the effective use of technology as a medium for teachers to shift instruction from physical to online learning environments as a primary and immediate adaptation to the crisis (Munoz-Najar et al., 2021; Statistics for National Statistics, 2021). Over time, researchers and practitioners began to share successful online teaching models (or experiences) adopted by schools in various countries during the closure of school campuses, with optimism that the pandemic may serve as an opportunity for reimagining the future of education by promoting teachers' innovative instruction and enhancing teachers' ability to effectively and appropriately integrate technology into teaching (Mishra et al., 2020; Shamir-Inbal and Blau, 2021).

Unexpected side effects of mandatory online teaching

The sudden onset of the pandemic required teachers to adopt online teaching, generally with very little training or background experience in distance learning (Silva et al., 2021; Trust and Whalen, 2021; Yi et al., 2021). Conceptually, the promotion of teachers' digital competencies brought about by the required use of both synchronous (SCMC) and asynchronous (ACMC) computer mediated communication held the potential to bring about the positive outcomes for teaching and learning associated with technology-integrated instruction (Bank, 2021; Hilger et al., 2021; Rodriguez et al., 2021). However, in reality, the outcomes of online teaching were perceived by teachers and other stakeholders to be largely negative, with empirical studies revealing that the hasty implementation of online courses by teachers with insufficient training or experience led to perceived high barriers related to the use of technology for teaching (Trust and Whalen, 2021) and resulting low teaching satisfaction (Fauzi and Khusuma, 2020), high workload (including work stress; Aperribai et al., 2020; Richmond et al., 2020; Jelińska and Paradowski, 2021a,b), and reduced willingness to use online teaching in the future (Zheng and Song, 2020). The title of a recent publication, "E-learning? Never again!" (Kulikowski et al., 2021), aptly describes the situation wherein instructors in higher education were required to implement online courses with unintended consequences of decreased job satisfaction, work motivation, and job involvement. In explaining how the online teaching required by COVID-19 impacted work motivation, core job characteristics were evaluated by Kulikowski et al. (2021); among these, four characteristics (task identity, task significance, task autonomy, and social interaction) decreased during online teaching, thereby leading to the unintended negative impacts on teachers' work motivation. Other empirical studies have revealed that mandatory online teaching contributed to detrimental side effects, particularly in terms of autonomy, due to work overload and inadequate working environments (Chan et al., 2021; Hilger et al., 2021; Soncini et al., 2021; Vargas Rubilar and Oros, 2021). More specifically, a significant decrease in work-related autonomy was found among Germany schoolteachers after the onset of pandemic (Hilger et al., 2021). Likewise, Italian teachers reported that, during mandatory online teaching, an increased workload and difficulties in carrying out teaching activities were main threats related to practical aspects of teaching (Soncini et al., 2021), while Argentine teachers reported that the work schedule for mandatory online teaching was disorganized (Vargas Rubilar and Oros, 2021), reflecting the lack of task autonomy (Kulikowski et al., 2021). In addition to threats to task autonomy, qualitative data from American elementary schoolteachers revealed that more freedom in the performance of tasks was required during mandatory online teaching and that the result of the standardization of distance

learning was characterized by "having their hands tied behind their backs" (Chan et al., 2021).

While the findings by Kulikowski et al. (2021) were from teachers in higher education, it is reasonable to infer that primary and middle school teachers may have encountered similar impacts to work motivation, since teachers at the primary and secondary level also had very little online teaching experience before the outbreak of the pandemic (Silva et al., 2021; Yi et al., 2021). Moreover, studies have shown that primary and middle school teachers may be even more reluctant to conduct online teaching or refuse to engage in online teaching while quarantine measures are in effect, as compared to teachers in higher education contexts (Jelińska and Paradowski, 2021a,b). As such, we contend that primary and middle school teachers represent a vulnerable population that deserves greater attention and research in terms of the psychological effects of online teaching during the pandemic, due to the several adverse effects reported in the literature (Chan et al., 2021; Ozamiz-Etxebarria et al., 2021b; Silva et al., 2021; Soncini et al., 2021; Vargas Rubilar and Oros, 2021).

Psychological need thwarting of online teaching

The emerging construct of psychological need thwarting (PNT) has been used to describe the effects of online teaching on primary and middle school teachers (Chen et al., 2020; Yi et al., 2021), finding significant effects from PNT of online teaching on psychological well-being. The construct of PNT was developed on the basis of Self Determination Theory (SDT), which includes three basic psychological needs – autonomy, competence, and relatedness – which strongly influence an individual's well-being (Ryan and Deci, 2000; Deci and Ryan, 2015). As such PNT serves as a more appropriate construct, as compared to Hackman and Oldham's Job Characteristics Theory (JCT), as applied in the study by Kulikowski et al. (2021). While JCT focuses on work motivation, the scale is largely descriptive in nature. In comparison, the variables of PNT have greater potential to interpret and elaborate the mechanisms behind the impact of online teaching on teachers' mental health. A better understanding of these underlying mechanisms can provide a more practical and nuanced contribution by evaluating the relationships among critical factors associated with teachers' psychological well-being during COVID-19 – an area of increasing importance to scholars in the field (Sahu, 2020; Palma-Vasquez et al., 2021; Yi et al., 2021; Jelińska and Paradowski, 2021a,b).

To extend our understanding of the dynamics of teacher psychology during online teaching, a more complete analysis of the long-term effects of online teaching on psychological well-being is required, beyond the relatively superficial findings concerning teachers' frustration (e.g., technological barriers, lack of willingness to engage in online teaching, and dissatisfaction) already covered sufficiently in the literature (Aperribai et al., 2020; Fauzi and Khusuma, 2020; Richmond et al., 2020; Zheng and Song, 2020; Trust and Whalen, 2021; Jelińska and Paradowski, 2021a,b). As such, while some preliminary investigations have demonstrated that schoolteachers were substantially impacted by mandatory online teaching in terms of psychological distress, the underlying mechanisms contributing to teacher psychology are still largely unknown. For example, while studies have shown that more than half (58.27%) of primary and secondary school teachers reported poor mental health during school closures (32) or that 20-26% of teachers reported mental health issues related to anxiety and depression occurred during quarantine (Yi et al., 2021), specific risk or protective factors related to teachers psychological well-being have yet to be clearly delineated in the extant literature.

The potential of PNT in explaining the mechanisms behind the impact of online teaching on teacher psychological well-being has some support from recent research (Yi et al., 2021). While most studies have focused on the satisfaction of individuals' psychological needs, research into environments which fail to satisfy (e.g., frustrate or block) these psychological needs has resulted in a complementary, but conceptually distinct, phenomenon to need satisfaction; namely, psychological need thwarting or PNT (Bartholomew et al., 2011; Gunnell et al., 2013; Cuevas et al., 2015). The literature has demonstrated that psychological need satisfaction and PNT involve different psychological processes, and importantly, that PNT exerts a stronger influence on an individual's negative affect as compared to psychological need satisfaction (Gunnell et al., 2013; Huyghebaert et al., 2018; Ebersold et al., 2019). As such, given the environment created through mandatory online teaching during the pandemic, it is reasonable to surmise that the perspective of PNT resulting from online teaching can better provide a more comprehensive explanation and elaboration of the influences of online teaching on teachers' motivation, willingness to conduct online teaching, and psychological well-being.

While the value of PNT in terms of psychological well-being has been reported in the literature, to date there have been few studies conducted during school closure which have examined the association between PNT related to online teaching (or subscales of PNT, such as autonomy-thwarting) and psychological well-being among primary and middle schoolteachers (Collie, 2021; Yi et al., 2021). One study adopted SDT as a framework to examine the effect of intrinsic motivation, extrinsic motivation, and TPK (technological and pedagogical knowledge) self-efficacy on the intention to continue using online teaching, mediated by the variable of burnout, but did not directly evaluate specific PNT factors (Panisoara et al., 2020). Thus, while these studies have provided preliminary support for the potential negative effects of PNT related to online teaching, data were collected only during school closure periods. This leads to an interesting, and heretofore unanswered, question: what are the lasting, long-term influences of PNT of online teaching on both school teachers' future intention to conduct online teaching and psychological well-being, particularly after online teaching is no longer required? The answer to this question is of vital importance, as the thwarting of

teachers' psychological needs may greatly reduce both immediate and long-term intention of using online teaching, potentially impacting teachers' overall attitudes towards the use of technology for teaching. Therefore, for a thorough evaluation of the impact of PNT from online teaching on both immediate and long-term intention and psychological well-being, data from different time points adopting a longitudinal design are necessary.

The measure of teachers' psychological need thwarting

In the specific context of the COVID-19 pandemic, PNT, particularly related to online teaching, has been developed and validated to some extent (Yi et al., 2021). However, compared to the construct of psychological need satisfaction, which has received greater attention, PNT is still an emerging construct (Huyghebaert et al., 2018; Wu et al., 2018; Lagios et al., 2022). As such, validation of instruments for the measurement of PNT is a prerequisite for conducting further studies in this area. According to a review of the literature, Bartholomew et al. (Bartholomew et al., 2011) first developed a scale to measure individual PNT in the context of sports (i.e., Psychological Need Thwarting Scale, PNTS) and later adapted this scale for the evaluation of a teaching environment; namely, a teachers' version of the Psychological Need Thwarting Scale (TPNTS; Bartholomew et al., 2014). The TPNTS has been translated into different languages, including a Spanish version (Cuevas et al., 2015) and a Chinese version (Chen et al., 2020). Recently, during the COVID-19 outbreak, in order to measure teachers' PNT related to the implementation of online teaching, Yi et al. (Yi et al., 2021) made necessary revisions to the Chinese version of the TPNTS to address the context of mandatory online teaching [i.e., Psychological Need Thwarting Scale of Online Teaching (PNTSOT)].

The TPNTS and PNTSOT instruments (Bartholomew et al., 2014; Cuevas et al., 2015; Chen et al., 2020; Yi et al., 2021) are based on Bartholomew et al.'s (2011) framework for PNT, and include 12 items equally divided into the three dimensions of psychological need thwarting: autonomy, competence, and relatedness. Regarding to the quality of the instruments, TPNTS and PNTSOT have suitable internal reliability with Cronbach's alphas ranging from 0.76 to 0.90. Moreover, factorial validity has been established by several studies through the use of confirmatory factor analysis (CFA; Cuevas et al., 2015; Chen et al., 2020; Yi et al., 2021). In support of the instruments' criterion validity, significant positive relationships of PNT with teacher burnout (Bartholomew et al., 2014; Cuevas et al., 2015; Chen et al., 2020; Franco et al., 2022) and psychological distress (Yi et al., 2021) have been found. Despite the above evidence, showing satisfactory reliability and validity for these scales, predictive validity testing has yet to be conducted in terms of the association between PNT of online teaching and both the intention to continue using online teaching in the future and the delayed effects on longer-term psychological well-being.

Aims of the present study

Online teaching has shown the potential to decrease task identity, task significance, autonomy, and social dimensions of teachers' job characteristics (Kulikowski et al., 2021). We believe this finding, of a negative impact of online teaching on teachers' job characteristics, may be best explained by an environment that thwarts individuals' psychological needs. Likewise, the thwarting of teachers' psychological needs arising from mandatory online teaching may also explain the moderate level of psychological distress (Palma-Vasquez et al., 2021; Yi et al., 2021) and very low satisfaction or intention to engage in online teaching (Fauzi and Khusuma, 2020; Zheng and Song, 2020) among school teachers reported in other COVID-19 studies. Considering the above points, we believe that it is essential to evaluate the PNT of online teaching and treat PNT as a warning (risk factor) for predicting potential burnout, considering the empirically demonstrated role of PNT as a mediator between a stressful environment and teacher burnout (Franco et al., 2022). As such, PNT can serve as a predictor of burnout, providing valuable insights for the design of interventions to address teachers' psychological needs before burnout occurs, preventing the associated negative effects on intention to continue online teaching and long-term psychological well-being. The contribution of this predictive role of PNT of online teaching can assist in overcoming the current difficulties in effectively evaluating teachers' immediate and long-term willingness to continue using online teaching after the pandemic ends, or if further waves (such as newer variants of the virus) of the pandemic require a return to distance learning.

Uniquely, this study evaluates the predictive influence of psychological need thwarting of online teaching, which itself is an emerging construct, from a longitudinal perspective. While the influence of the construct of psychological need thwarting and satisfaction has been simultaneously evaluated by some prior studies, the present study emphasizes the role of PNT in order to (a) address the lack of empirical studies evaluating PNT, as compared to psychological need satisfaction, particularly in the context of online teaching and (b) to avoid the inclusion of too many items in both the questionnaire and model, which would create a burden for respondents while also making validation of the PNTSOT instrument difficult. Some recent studies (Costa et al., 2019; Moè and Katz, 2020; Rodríguez-Meirinhos et al., 2020), have modelled and evaluated the complex interaction among elements of both psychological need thwarting and satisfaction or evaluated profiles including elements of both psychological need thwarting and satisfaction (Warburton et al., 2020), finding the expected positive impact of psychological need satisfaction as compared to frustration. Many of these studies have advocated for longitudinal research (Costa et al., 2019; Rodríguez-Meirinhos et al., 2020; Warburton et al., 2020). Thus, the present study, in the specific context of online teaching, seeks to better understand the role of PNT of online teaching in terms of teachers'

psychological well-being, intention to teach online, and satisfaction with online teaching, utilizing longitudinal data while, simultaneously, adopting a model for testing the immediate and delayed effects of PNT of online teaching.

Furthermore, in the present study, we adopt a research design that integrates cross-sectional and longitudinal elements to test the immediate and delayed effects of PNT of online teaching while simultaneously systematically evaluating the psychometric properties of an instrument for measuring PNT of online teaching (i.e., PNTSOT), originally developed by Yi et al. (2021). In addition to confirming the factorial validity and internal reliability, which have already been sufficiently evaluated by Yi et al. (2021), the core purpose of this study is to address a knowledge gap concerning the lack of longitudinal data required for more rigorous psychometric evaluation, such as test–retest reliability, longitudinal measurement invariance, and predictive validity regarding the association of PNT of online teaching (as measuring during school closure) with other relevant factors, such as future intention to continue online teaching and the long-term impacts on teachers' psychological well-being (measured when restrictions are removed and offline teaching is available).

Materials and methods

Participants

This study was conducted in a city in a province in central China which had implemented mandatory online teaching, with campuses closed due to multiple COVID-19 infections reported at the end of October 2021. It should be noted that although the Chinese government had relaxed COVID-19 restrictions since September 2020 (i.e., indoor activities were allowed and school campuses were reopened), once an infection was reported, restrictive measures were still untaken immediately in order to limit infections. As such, after an outbreak of the pandemic in the city in which the study took place, the city government decided to close all school campuses and fully implement online courses from November 3, 2021. Our research team has been engaged in long-term collaboration with the city's educational authorities, providing psychological counseling services and regularly holding mental health workshops for the teachers in the city. Thus, in order to monitor the mental health status of teachers during this quarantine period (i.e., 2 weeks after the full implementation of online teaching), our research team, with the assistance of local educational authorities, conducted an online survey of primary and secondary school teachers (Time 1: mid-November, 2021). A follow-up collection of data was conducted after a two-month interval (Time 2: mid-January, 2022). At Time 2, campuses had reopened for 2 weeks and mandatory online teaching was no longer being implemented, with schools returning to a face-to-face mode of instruction. The survey was administered through an online questionnaire, forwarded by the educational administration of each school district to each school in their jurisdiction for voluntary completion by teachers. A total of 9,554 (Time 1) and 4,176 (Time 2) teachers completed the online survey (cross-sectional portion). Participants completing the first survey were asked to leave their email if they would like to participate in a follow-up survey after 2 months. A total of 1,642 school teachers left their email information and participated in the longitudinal portion of the study. Written informed consent was obtained electronically on the first page of the online survey, providing participants with information on the purpose of the research, the affiliation of the researchers, and a guarantee of privacy and anonymity through appropriate storage and curation of the collected data. This study was approved by the Jiangxi Psychological Consultant Association (IRB ref.: JXSXL-2020-J013).

Measures

The design of the survey was purposefully arranged to minimize the burden to the respondents, providing questions which were relevant only to the current situation (mandatory online teaching due to the recent outbreak). Due to the sudden nature of the announcements related to school closure, teachers were required to conduct online teaching from home and, at this point in time, were required to create, prepare, and manage a great deal of instructional materials. As such, the data collected at Time 1 included measures related to online teaching (including the PNTSOT and a questionnaire assessing satisfaction with online teaching) and a measure of psychological distress (DASS-21).

For both theoretical and practical reasons, a measure of teacher burnout was not included at Time 1, but was evaluated at Time 2. Theoretically, the construct of burnout was utilized in our model as a predicted variable, indicative of the long-term effects of PNT and psychological well-being from a longitudinal perspective, and thus was not included in data collected a Time 1. From a practical perspective, this measure was used only for Time 2 in order to reduce the length of the survey at Time 1 and to avoid influencing teachers' attitudes towards the longer-term impacts of online teaching, with general attitudes towards online teaching and measures of psychological well-being collected using the PNTSOT and DASS-21 instruments.

For the assessment of satisfaction with online teaching, a single question was posed (Time 1): "How would you rate the effectiveness of your online teaching?" with possible responses varying from very dissatisfied (UNESCO, 2020) to very satisfied (Haug et al., 2022). To evaluate teachers' intention to continue using online teaching another question was included in the survey for Time 2: "Would you like to continue using online teaching in the future?" with possible responses varying was from very unwilling (UNESCO, 2020) to very willing (Shamir-Inbal and Blau, 2021).

Psychological need thwarting scale of online teaching

The PNTSOT developed by Yi et al. (2021) was used to assess the extent of psychological need thwarting during online teaching. As mentioned above, the PNTSOT includes three subscales (i.e., autonomy, competence, and relatedness thwarting) with means and standard deviations for each question by subscale provided in Table 1. The PNTSOT was rated on a seven-point Likert-type scale (ranging from 1 to 7) with higher scores indicating a greater degree of psychological need thwarting during online teaching. The PNTSOT has demonstrated good factorial validity among primary and middle school teachers (Yi et al., 2021). Specifically, the factor structure of the PNTSOT was consistent with the original structure (i.e., a three-factor structure) of the Chinese version of the TPNTS reported in Chen et al. (Chen et al., 2020), where the results of CFA demonstrated acceptable fit according to relevant indices (CFI = 0.97, NNFI = 0.95, RMSEA = 0.09, and SRMR = 0.05). The PNTSOT was administered twice in the present study. For the survey at Time 1, the instructions asked participants to respond while considering their current situation in implementing online teaching, while the instructions for Time 2 asked respondents to recall their experiences in implementing online teaching for the previous 2 months. The aim of this wording was to conduct a more systematic evaluation of the reliability and validity of the PNTSOT through a focus on the collection of longitudinal data which could be evaluated in terms of test–retest reliability. Details on the psychometric characteristics of the PNTSOT are presented in the results section.

Teacher burnout

To evaluate teacher burnout at Time 2, this study utilized the subscale of "Emotional Exhaustion" from the Chinese version of the Primary and Secondary School Teachers' Job Burnout Questionnaire (CTJBO). The questionnaire includes 22 items and was developed by Wu et al. (2016) as an adaptation of the Maslach Burnout Inventory (Maslach and Jackson, 1981) revised to fit the cultural and linguistic context of primary and middle school teachers in Mainland China. The CTJBO includes three dimensions: emotional exhaustion (8 items), depersonalization (8 items) and reduced personal accomplishment (6 items; Wu et al., 2016). The CTJBO scale has demonstrated satisfactory factorial validity (Wu et al., 2016) among primary and middle schoolteachers (i.e., RMSEA = 0.06, NFI = 0.950, and CFI = 0.960). Sample items from the Emotional Exhaustion sub-scale of the CTJBO are "I feel exhausted after a long day of work" and "I feel very tired when I wake up in the morning." The item adopted a Likert-type, ranging from 1 (strongly disagree) to 7 (strongly agree). The internal consistency was high for both longitudinal and cross-sectional data in the present study (Cronbach's alpha for both was 0.95).

Depression, Anxiety, and Stress Scale (DASS-21)

The Chinese version of the 21-item Depression, Anxiety, and Stress Scale (DASS-21) was adopted to measure teachers' psychological distress during school closure (Time 1) and after restrictions were lifted and face-to-face teaching resumed (Time 2). The survey instructions asked the participants to reflect on their current mental state during school closure (for Time 1) or their mental state in the recent 2 weeks (for Time 2). The DASS-21 is equally divided according to three emotional states: depression, anxiety, and stress (Lovibond and Lovibond, 1996). Items are rated on a four-point Likert-type scale (ranging from 0 to 3) with higher scores reflecting greater levels of these three emotional states. Recently, several studies have provided empirical evidence that the scores of the three subscales reflect general negative emotion (Yeung et al., 2020; Zanon et al., 2020). That is, the average score of all items from the three subscales can serve as an indicator of psychological distress. Moreover, according to Lovibond and Lovibond (1996), the summed score of the subscales multiplied by two can reflect clinical levels of psychological distress when exceeding the following cut-off values: depression (Shamir-Inbal and Blau, 2021), anxiety (Munoz-Najar et al., 2021), and stress (Bank, 2021). The DASS-21 has been translated into Chinese and has been widely used with Chinese samples, including studies by Chan et al. (2012) and Wang et al. (2015). In Chan et al.'s (2012) study with typical adults, the DASS-21 demonstrated satisfactory factorial validity (i.e., RMSEA = 0.08; SRMR = 0.06; CFI = 0.98; and NNFI = 0.97). In this present study, the Cronbach's alpha of the overall DASS-21 was 0.95 (Time 1) and 0.96 (Time 2) using the longitudinal data.

Data analysis

In order to thoroughly and systematically evaluate the impact of PNT of online teaching in terms of immediate effects (including psychological distress and satisfaction with online teaching) and delayed effects (including intention to continue online teaching and burnout), mean values and correlation coefficients for all variables were first analyzed. Subsequently, the reliability and factorial validity of PNTSOT were evaluated. Finally, structural equation modelling (SEM) was utilized in order to test the causal relationships among variables, including the effect of PNT of online teaching on psychological distress and satisfaction with online teaching (both measured at Time 1), as well as the delayed effects on intention to continue online teaching and burnout (both measured at Time 2). Furthermore, in order to evaluate the psychometric properties of the PNTSOT, data collected by Yi et al. (2021) was also included in the present study for analysis (i.e., descriptive statistics for the variables, tests of model fit, and measurement invariance). The procedures for data analysis conducted in this study are described in detail below.

Descriptive statistics were first used to analyze the characteristics of the participants and their responses on the PNTSOT and criterion variables (i.e., burnout, satisfaction with online teaching, intention to continue online teaching in the future, and psychological distress). Moreover, Pearson correlations among the observed variables (averaged values) for the PNTSOT,

TABLE 1 Demographics of the participants and the descriptive statistics of the observed variables.

Source	Yi et al. (12)[a]	The present study (cross-sectional portion)		The present study (longitudinal portion)	
Occasion	**2020.5-6**	**2021.11 (Time 1)**	**2022.01 (Time 2)**	**2021.11 (Time 1)**	**2022.01 (Time 2)**
Valid number	9,030	9,554	4,176	1,642	
School type (primary school); *n* (%)	5,838 (64.65%)	6,004 (62.8%)	2,580 (61.78%)	1,159 (70.6%)	
Sex (female); *n* (%)	6,563 (72.7%)	6,933 (72.6%)	3,190 (76.4%)	1,305 (79.5%)	
Age; mean (SD)	33.94 (8.81)	37.20 (9.63)	34.76 (10.04)	34.22 (8.72)	
A. Psychological need thwarting scale of online teaching; mean (SD)					
Subscale of autonomy					
1. In online courses during the pandemic, I cannot decide for myself how I want to teach	3.87 (1.38)	3.69 (1.48)	3.63 (1.64)	3.57 (1.48)	3.42 (1.62)
2. In online teaching work during the pandemic, I feel there is pressure that affects my behavior and requires me to comply in a certain way	4.00 (1.42)	3.96 (1.52)	3.69 (1.67)	3.94 (1.56)	3.54 (1.65)
3. I have to follow a prescribed online teaching style during the pandemic.	4.22 (1.43)	4.33 (1.48)	4.05 (1.62)	4.32 (1.49)	3.90 (1.59)
4. During the pandemic, I feel pressure from the external environment that limited me in choosing a particular online teaching style.	4.01 (1.43)	3.96 (1.49)	3.75 (1.62)	3.86 (1.52)	3.61 (1.58)
Overall mean (autonomy)	4.02 (1.09)	3.98 (1.15)	3.78 (1.37)	3.92 (1.19)	3.62 (1.35)
Subscale of competence					
5. There are some online teaching situations that make me feel incapable in my daily work environment during the pandemic.	4.41 (1.55)	4.45 (1.57)	4.15 (1.73)	4.45 (1.61)	4.07 (1.71)
6. I sometimes talk about the things that make me feel powerless to do my online teaching job during the pandemic.	4.25 (1.49)	4.16 (1.53)	3.96 (1.66)	4.11 (1.58)	3.87 (1.65)
7. Online teaching during the pandemic sometimes makes me feel powerless.	4.24 (1.53)	4.22 (1.56)	3.97 (1.69)	4.16 (1.61)	3.88 (1.68)
8. Due to the lack of training opportunities in my environment, I feel that I am capable of performing online teaching tasks.	3.18 (1.39)	3.09 (1.39)	3.10 (1.48)	3.01 (1.37)	2.95 (1.41)
Overall mean (competence)	4.02 (1.22)	3.97 (1.25)	3.79 (1.42)	3.93 (1.28)	3.69 (1.39)
Subscale of relatedness					
9. I feel disconnected from other colleagues and leaders when teaching online during the pandemic.	2.74 (1.32)	2.69 (1.34)	2.76 (1.42)	2.59 (1.30)	2.61 (1.33)
10. I do not feel that my colleagues and leaders care about me when teaching online during the pandemic.	2.93 (1.42)	2.82 (1.40)	2.89 (1.48)	2.70 (1.38)	2.76 (1.40)
11. I feel that my colleagues and leaders are jealous of me when I achieve good results in online teaching during the pandemic.	2.34 (1.17)	2.26 (1.17)	2.42 (1.31)	2.14 (1.14)	2.32 (1.22)
12. I feel that my colleagues and leaders do not like me when I conduct online teaching during the pandemic.	2.31 (1.16)	2.25 (1.15)	2.41 (1.29)	2.14 (1.13)	2.32 (1.23)
Overall mean (relatedness)	2.58 (1.09)	2.50 (1.10)	2.62 (1.23)	2.39 (1.07)	2.50 (1.16)
B. Psychological distress	15.04 (20.03)	20.15 (20.98)	22.44 (24.42)	19.62 (20.22)	19.12 (21.78)
C. Teacher burnout – emotional exhaustion	Not applicable		3.51 (1.54)	Not applicable	3.46 (1.52)
D. Satisfaction with online teaching (number of participants choosing satisfactory or very satisfactory); *n* (%) – measured at Time 1	Not applicable	6,595 (69.02%)	Not applicable	1,143 (69.61%)	Not applicable
E. Intention of adopting online teaching in the future (number or respondents responding with a score of 6 or higher, on a 10-point scale); *n* (%) – measured at Time 2	Not applicable	Not applicable	1,578 (37.78%)	Not applicable	615 (37.45%)

burnout, satisfaction of the online teaching, intention of continuing online teaching in the future, and psychological distress were computed for the longitudinal data. Following, McDonald's ω, the Intraclass correlation coefficient (ICC) with a 2-way mixed effects model combining a Bland–Altman plot, and CFA were used to evaluate internal reliability, test–retest reliability, and factorial validity. It should be noted that, during the development of the Chinese version of TPNTS, on which PNTSOT is based, item 8 ("due to the lack of training opportunities in the environment, I feel that I am not competent in my daily work tasks") was found to be cross-loaded for both the relatedness thwarting and competence thwarting factors. Thus, as the inclusion of item 8 may affect the overall measurement quality of the scale (Chen et al., 2020), we conducted CFA to evaluate whether this item should be included in the PNTSOT based on the data collected by our current study.

After scrutinizing the factorial validity of each sample, a multi-group and longitudinal invariance test was conducted to assess whether the PNTSOT possessed measurement invariance across different occasions. Finally, we constructed and tested a structural equation model (SEM) including a higher-order CFA of PNTSOT and a causal model to test criterion validity (see Figure 1). Specifically, in this model, the higher-order latent variable of PNTSOT (Time 1) served as an exogenous variable (with its three subscales serving as first order latent variables), with psychological distress (Time 1 and Time 2), satisfaction with online teaching (Time 1), intention for future online teaching (Time 2), and teacher burnout (Time 2) added as endogenous variables.

Due to violation of the assumption of a normal distribution for the longitudinal data from the PNTSOT (the values of Shapiro–Wilk test were 0.99 and 0.97, both $p < 0.01$), estimation utilizing diagonally weighted least squares (DWLS) was adopted for CFA, tests of measurement invariance, and SEM, as DWLS is more suitable for dealing with non-normally distributed data (Li, 2016). In terms of the evaluation of factorial and criterion validity, we adopted the following indices: Satorra-Bentler Scaled Chi-Square (SB χ^2), comparative fit index (CFI), non-normed fit index (NNFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). CFI and NNFI values of 0.95 or higher, RMSEA values of 0.06 or lower, and SRMR values of 0.08 or lower were considered acceptable (Hu and Bentler, 1999).

Finally, a series of model comparisons were conducted to evaluate the measurement equivalence of the PNTSOT across different occasions. More specifically, the following comparisons were made: (a) the configural model (i.e., baseline model) was compared with the factor-loading constrained equal model; (b) the factor-loading constrained equal model (a less constrained model) was compared with the factor-loading and item intercept constrained equal model (a more constrained model); (c) the factor-loading and item intercept constrained equal model (a less constrained model) was compared with the factor-loading, item intercept, and errors constrained equal model (a more constrained model); (d) the factor-loading and item intercept constrained equal model (a less constrained model) was compared with the factor-loading, item intercept, and factor variance, as well as the covariance constrained equal model (a more constrained model). The differences in CFI, RMSEA, and SRMR from the less constrained models to the more constrained model were used to judge whether or not measurement invariance was supported: ΔCFI > −0.01, ΔRMSEA <0.015, and ΔSRMR <0.03 (for factor loading) or ΔSRMR <0.01 (Chen, 2007).

Results

Participant characteristics, observed means, and relationships among variables

Table 1 presents the characteristics of the participants from data collected from the two cross-sectional surveys and the longitudinal study with mean observed scores for the variables of interest in this study. In terms of demographic characteristics, the participants in this study were mostly from primary schools (61.7–70.6%), mostly female (72.6–79.5%), with an average age between 34 and 37 years, which are similar to the participant characteristics reported by Yi et al. (2021). The above demographic variables were also close to the overall population statistics [i.e., from all primary and middle schoolteachers in mainland China; 2020 Education Statistics (Internet), 2021], in terms of age (i.e., population mean age of 37.78), school type (64% of schools nation-wide are primary schools), and gender (70% of teachers are female) which supports the representativeness of the participants. Supplementary Table 1 also reports the Pearson correlations among the observed variables of the three subscales of the PNTSOT.

Next, in order to evaluate potential differences between primary and middle school teachers, we evaluated the mean scores on all variables. The results of independent sample *t*-tests demonstrated very few significant differences (with trivial effect sizes) between primary and middle school teachers on the variables of interest (see Supplementary Table 2). Moreover, the results of CFA and measurement invariance also confirmed that PNTSOT is appropriate for use with both groups of teachers with no bias (see Supplementary Table 3). Consequently, it was not necessary to separately evaluate the two groups of teachers in the analysis and the data from both groups was pooled for subsequent analyses.

For the purpose of evaluating the relationships among variables, including the predictive role of PNT of online teaching on the delayed measures of intention to continue online teaching and burnout, PNTSOT scores from Time 1 were utilized. However, for the purpose of validation, differences in responses to the PNTSOT instrument were evaluated from Time 1 (assessing reported PNT of online teaching at that time) and Time 2 (where

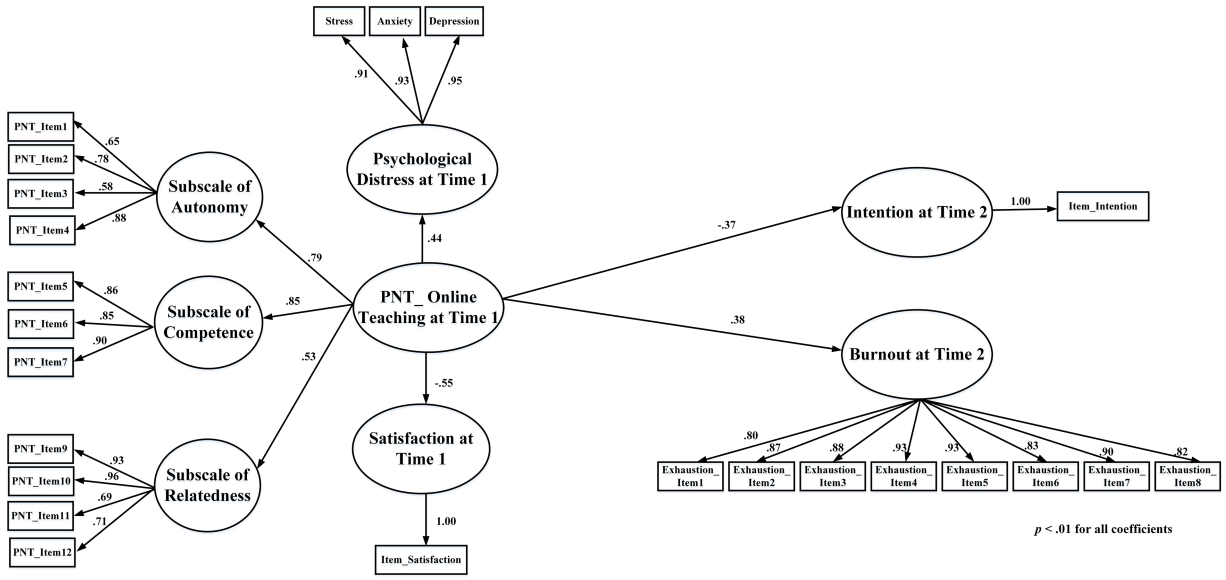


FIGURE 1
Criterion validity as examined by structural equation modeling (SEM; including second-order confirmatory factor analysis and a causal model; n=1,642); variables include psychological need thwarting of online teaching (PNT_Online Teaching), satisfaction with online teaching (Satisfaction), and intention of continuing online teaching the future (Intention).

respondents were asked to reflect on their previous 2 months of online teaching). Additionally, differences in perception between Time 1 and Time 2 were evaluated in order to better interpret the effects of PNT of online teaching in the context of other variables of interest.

Among the three dimensions of psychological need thwarting, autonomy and competence thwarting were higher during Time 1 (as compared to Time 2), with mean autonomy thwarting of 3.98 (cross-sectional data) and 3.92 (longitudinal data) and mean competence thwarting of 3.97 (cross-sectional data) and 3.93 (longitudinal data). The scores of these two subscales decreased when returning to offline instruction (Time 2) for both the cross-sectional and longitudinal data, with means for autonomy thwarting of 3.78 (cross-sectional data) and 3.62 (longitudinal data), and means for competence thwarting of 3.79 (cross-sectional data) and 3.69 (longitudinal data). Interestingly, relatedness thwarting increased from Time 1 to Time 2, for both cross-sectional and longitudinal data; with mean values at Time 1 of 2.50 (cross-sectional data) and 2.39 (longitudinal data); and mean values at Time 2 of 2.62 (cross-sectional data) and 2.50 (longitudinal data). It is noted that despite changes in the observed scores for PNTSOT over time, tests for longitudinal reliability and validity are still necessary to evaluate ICC and longitudinal measurement invariance.

The participants' overall psychological distress was 20.15 (cross-sectional data) and 19.62 (longitudinal data) at Time 1, while psychological distress scores were 22.44 (cross-sectional data) and 19.12 (longitudinal data) at Time 2. Regarding specific emotional states, the percentage of teachers with clinical depression increased in both cross-sectional (from 25.2 to 29.9%) and longitudinal data (from 22.41 to 23.51%); the percentage of anxiety and stress increased in the cross-sectional data (anxiety: from 35.5 to 40.35%; stress: from 15.7 to 19.0%) but decreased in the longitudinal study (anxiety: from 35.87 to 34.71%; stress: from 16.13 to 15.71%). Participants' burnout was considered moderate (cross-sectional data: 3.51; longitudinal data: 3.46) due to the value being close to the median of the scale (i.e., 3.50). Finally, more than half of the teachers were satisfied or very satisfied with their online teaching, as indicated by responses of either "satisfied" (3 out of 4 on the Likert-type scale) or "very satisfied" (4 out of 4 on the Likert-type scale); with 69.02% satisfaction based on the cross-sectional data and 69.61% based on the longitudinal data. However, only 37% of the participants responded they would like to continue using online teaching in the future (i.e., respondents who selected a value of 6 or more out of a total of 10).

Table 2 displays the Pearson correlations among variables of interest for the longitudinal data. PNT of online teaching (for both Time 1 and Time 2) was significantly correlated with all criterion variables. Specifically, a negative correlation was found between PNT of online teaching and satisfaction at Time 1 ($r = -0.39$; $p < 0.001$) and Time 2 ($r = 0.28$; $p < 0.001$); and a negative correlation between PNT of online teaching and intention to teach online found at Time 1 ($r = -0.28$; $p < 0.001$) and Time 2 ($r = -0.26$; $p < 0.001$). Moreover, PNT of online teaching was significantly and positively correlated with burnout and psychological distress, with correlation coefficients ranging from 0.27 to 0.44 (all $p < 0.001$). These observed associations were supported by the results of structural equation modelling (Figure 1), with the model demonstrating significant causal paths in line with the correlations

TABLE 2 Pearson correlations among the variables of psychological need thwarting (PNT) of online teaching, satisfaction, intention, burnout, and psychological distress for the longitudinal data (n=1,642).

	1	2	3	4	5	6	7
1. PNT of online teaching (Time 1)	1						
2. PNT of online teaching (Time 2)	0.56**	1					
3. Satisfaction (Time 1)	−0.39**	−0.28**	1				
4. Intention (Time 2)	−0.28**	−0.26**	0.27**	1			
5. Burnout (Time 2)	0.27**	0.44**	−0.20**	−0.15**	1		
6. Psychological distress (Time 1)	0.36**	0.32**	−0.20**	−0.11**	0.36**	1	
7. Psychological distress (Time 2)	0.27**	0.42**	−0.15**	−0.06*	0.52**	0.58**	1

Time 1 was measured at mid-November, 2021, Time 2 was measured at mid-January, 2022; **$p < 0.01$, *$p < 0.05$.

reported here. These findings are provided in the following sub-section.

Reliability and validity of the PNTSOT

The internal reliability of the PNTSOT was satisfactory, given that the McDonald's ω of three subscales were all higher than 0.80 for longitudinal participants (Time 1: Autonomy thwarting $\omega = 0.80$, Competence thwarting $\omega = 0.87$, and Relatedness thwarting $\omega = 0.88$; Time 2: Autonomy thwarting $\omega = 0.86$, Competence thwarting $\omega = 0.90$, and Relatedness thwarting $\omega = 0.91$). Similar findings were reported for the cross-sectional data, with coefficients exactly the same as with the longitudinal data (with the exception of an ω value of 0.78 for autonomy thwarting at Time 1). Moreover, the ICC for PNTSOT using a 2-way mixed effects model was 0.71, indicating acceptable test–retest reliability across the two-month interval. Analysis also demonstrated that 95% of the data points lied within ±1.96 SD of the mean difference in a Bland–Altman plot (see Figure 2), with only 88 data points outside of that range, suggesting appropriate consistency of PNTSOT between the two occasions.

The results of CFA demonstrated that an 11-item version of PNTSOT (i.e., excluding item 8) had a better model fit than the scale with 12 items (see Table 3). Specifically, in the 11-item version, CFI and NNFI ranged from 0.990 to 0.999, and RMSEA and SRMR ranged from 0.017 to 0.068, showing good fit. Subsequently, we further tested the measurement invariance of the PNTSOT (see Table 4). The results indicated that the 11-item PNTSOT demonstrated scalar invariance (i.e., with loadings and thresholds equal) across different occasions. Namely, the comparisons of pairs of data points collected in our present study matches the findings of Yi et al. (2021). However the 12-item version, which included item 8, failed to demonstrate scalar invariance. Furthermore, when online teaching was being launched (i.e., in the study of Yi et al. (2021) and at Time 1 of our study), the PNTSOT achieved the strictest measurement invariance, namely having error variance equivalence for each item (i.e., the measurement error of each item was equal across different occasions) and also had the same factor variance as well as covariance (i.e., the variance of the three factors and the covariance between factors in PNTSOT was equal across different occasions). However, for Time 2 of our study, the PNTSOT did not pass the measurement equivalence criteria between occasions (i.e., a comparison between the data from Yi et al. (2021) and Time 2 of our study and a comparison between Time 1 and Time 2 of our study) in terms of measurement error variance, factor variance, and factor covariance. However, this result must be interpreted with caution, since the instructions for the survey at Time 2 asked participants to consider their situation 2 months earlier, when online teaching was launched, which required recalling their memory of online teaching at that time.

Regarding criterion validity, concurrent and predictive validity were assessed by SEM (see Figure 1). Since the residual was too high when the scores for DASS-21 at Time 1 and Time 2 were both placed in the model simultaneously, ultimately only the scores for psychological distress at Time 1 were included for SEM. As the model demonstrated an acceptable fit (see Table 3), we further scrutinized the path coefficients of the model. The results supported both concurrent and predictive validity, since the PNT of online teaching at Time 1 was significant and positively associated with psychological distress at Time 1 ($\gamma = 0.44$, $t = 7.45$, $p < 0.001$), negatively associated with the satisfaction of online teaching at Time 1 ($\gamma = -0.55$, $t = -19.95$, $p < 0.001$), negatively associated with the intention to continue online teaching at Time 2 ($\gamma = -0.37$, $t = -11.45$, $p < 0.001$), and positively associated with burnout at Time 2 ($\gamma = 0.38$, $t = 12.39$, $p < 0.001$).

Discussion

Purpose and main findings

Recent literature on the effects of the COVID-19 pandemic on teachers has widely reported low levels of satisfaction with online teaching (Fauzi and Khusuma, 2020), an unwillingness to use online teaching in the future (Zheng and Song, 2020), and even poor psychological well-being (Aperribai et al., 2020; Richmond et al., 2020; Palma-Vasquez et al., 2021; Silva et al., 2021; Jelińska and Paradowski, 2021a,b). As such, the purpose of the present study was two-fold. First, the relationship between PNT of online teaching (characterized by the thwarting of teachers' autonomy,

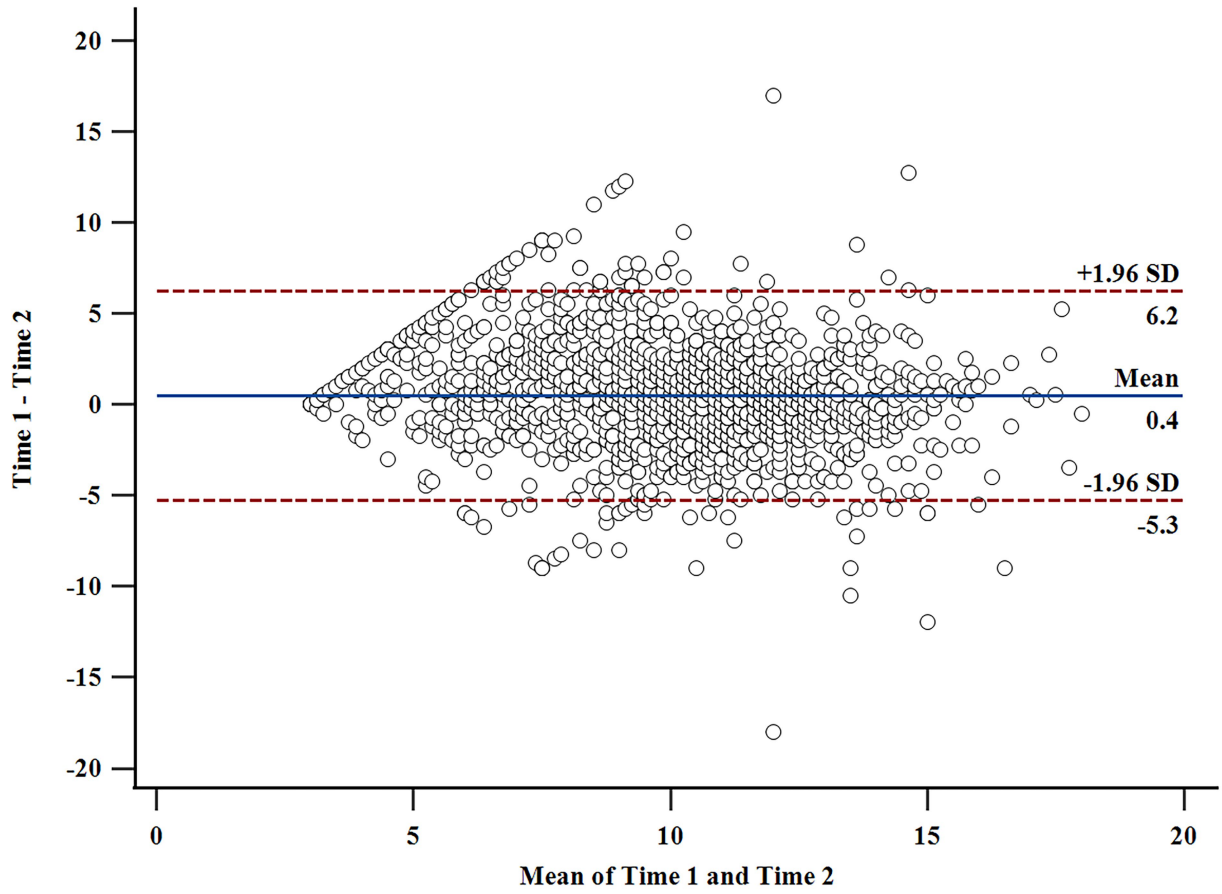


FIGURE 2
Bland–Altman plot of the Psychological Need Thwarting Scale of Online Teaching (PNTSOT) from Time 1 and Time 2 (n=1,642).

competence, and relatedness needs) and concurrent satisfaction with online teaching and psychological distress, as well as the predictive effect of PNT of online teaching on intention to continue online teaching after a return to regular classes and teacher burnout 2 months later. In terms of the first aim, the study found that PNT of online teaching is significantly and negatively associated with satisfaction with online teaching, such that increased need thwarting lowers teachers' satisfaction. At the same time, PNT of online teaching is positively associated with teachers' psychological distress, such that increased need thwarting is associated with increased depression, anxiety, and stress. Furthermore, need thwarting caused by online teaching is predictive of teacher burnout after 2 months as well as a decrease in teachers' intention to use online teaching when no longer required to do so. These findings extend those of previous studies – wherein psychological need thwarting, generally, have been shown to significantly and negatively impact teachers' psychological well-being (Cuevas et al., 2015; Chen et al., 2020) – by assessing delayed effects through the analysis of longitudinal data. The second aim was to validate an instrument for evaluating PNT of online teaching, the PNTSOT, which is proposed as a valuable measure for predicting teacher psychological and behavioral outcomes. The results of the present study convincingly support the internal reliability, test–retest reliability, measurement invariance, and criterion validity (concurrent and predictive) of the instrument, based on both cross-sectional and longitudinal data, suggesting that the PNTSOT instrument can be of value for both researchers and practitioners.

Theoretical implications

The lasting effect of PNT of online teaching

In terms of a theoretical contribution, although a few studies have reported a negative relationship between PNT of online teaching and the psychological well-being outcomes using cross-sectional data (Yi et al., 2021), to our knowledge, there have been no studies examining the long-term and delayed effects of mandatory online teaching on teachers' psychological well-being and intention to utilize online teaching during times when traditional face-to-face instruction can be adopted. Moreover, there has been a general lack of longitudinal data to support the impact of PNT, overall, on outcomes related to individual well-being (Cuevas et al., 2015; Chen et al., 2020; Lagios et al., 2022).

TABLE 3 Results of model fit.

	χ^2 (*df*)	CFI	NNFI	RMSEA (90% confidence interval)	SRMR
Yi et al. (2021) (12 items)	1795.06 (51)	0.983	0.978	0.062 (0.059–0.064)	0.088
Yi et al. (2021) (11items)	573.55 (41)	0.990	0.990	0.038 (0.035–0.041)	0.055
Time 1[a] (12 items)	594.64 (51)	0.996	0.994	0.033 (0.031–0.036)	0.089
Time 1[a] (11 items)	156.43 (41)	0.999	0.999	0.017 (0.014–0.020)	0.055
Time 2[a] (12 items)	1478.51 (51)	0.983	0.978	0.082 (0.078–0.085)	0.076
Time 2[a] (11 items)	431.05 (41)	0.990	0.990	0.048 (0.044–0.052)	0.046
Time 1[b] (12 items)	521.87 (51)	0.979	0.973	0.075 (0.069–0.081)	0.089
Time 1[b] (11 items)	166.95 (41)	0.990	0.990	0.043 (0.037–0.050)	0.060
Time 2[b] (12 items)	952.77 (51)	0.974	0.966	0.104 (0.098–0.110)	0.079
Time 2[b] (11 items)	356.12 (41)	0.990	0.990	0.068 (0.062–0.075)	0.047
Model of testing criterion validity (12 items)	1596.43 (268)	0.981	0.978	0.055 (0.052–0.058)	0.084
Model of testing criterion validity (11 items)	326.25 (245)	0.999	0.999	0.014 (0.009–0.018)	0.071

[a]Cross-sectional data;
[b]Longitudinal data. CFI, comparative fit index; NNFI, nonnormed fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.

TABLE 4 Tests of measurement invariance.

	Configural model		Loadings constrained as equal		Loadings and thresholds constrained as equal		Loadings, thresholds, and errors constrained as equal		Loadings, thresholds, factor variance, and covariance constrained as equal[a]	
Yi et al. (2021) and Time 1										
	12 items	11 items	12 items	11 items	12 items	11 items	12 items	11 items	12 items	11 items
X^2 (*df*) or ΔX^2 (Δdf)	1863.04 (102)	571.89 (82)	−57.32 (9)	−19.72 (8)	151.77 (9)	163.32 (8)	115.12 (12)	95.92 (11)	−280.42 (6)	−87.78 (Gouëdard et al., 2020)
CFI or ΔCFI	0.970	0.990	**0.001**	**0.001**	**−0.002**	**−0.003**	**−0.002**	**−0.002**	**0.005**	**0.001**
RMSEA or ΔRMSEA	0.043	0.025	**−0.002**	**−0.001**	**0.000**	**0.002**	**−0.001**	**0**	**−0.005**	**−0.003**
SRMR or ΔSRMR	0.050	0.035	**0.000**	**0.001**	**−0.038**	**−0.024**	**0.001**	**0.001**	**0.004**	**0.004**
Yi et al. (2021) and Time 2										
X^2 (*df*) or ΔX^2 (Δdf)	4479.91 (102)	1439.87 (82)	107.75 (9)	−9.29 (Munoz-Najar et al., 2021)	−208.01 (9)	117.47 (Munoz-Najar et al., 2021)	1779.46 (12)	1610.66 (11)	−82.94 (6)	329.05 (Gouëdard et al., 2020)
CFI or ΔCFI	0.896	0.961	−0.002	**0.001**	0.005	**−0.003**	−0.042	−0.046	0.002	**−0.009**
RMSEA or ΔRMSEA	0.081	0.050	−0.003	**−0.002**	−0.005	**−0.001**	0.01	0.018	−0.002	**0.004**
SRMR or ΔSRMR	0.036	0.023	0.002	**0.001**	−0.021	**−0.001**	0.004	**0.004**	0.0231	0.027
Time 1 and Time 2										
X^2 (*df*) or ΔX^2 (Δdf)	2614.71 (225)	1065.26 (183)	−38.16 (9)	−22.12 (8)	99.07 (9)	216.75 (Munoz-Najar et al., 2021)	382.94 (Yi et al., 2021)	295.24 (11)	156.15 (6)	454.16 (16)
CFI or ΔCFI	0.875	0.943	0.003	**0.002**	−0.005	**−0.010**	−0.019	−0.019	−0.008	−0.029
RMSEA or ΔRMSEA	0.081	0.054	−0.003	**−0.002**	0	**0.005**	0.004	**0.005**	0.002	**0.010**
SRMR or ΔSRMR	0.038	0.027	0	**0**	−0.020	**−0.011**	0.001	**0.001**	0.013	0.015

[a]The change was calculated with the model of Loadings and Thresholds Constrained as Equal. CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual. Supported measurement invariance values are in bold (i.e., ΔCFI > −0.01; ΔRMSEA <0.015; ΔSRMR <0.03 (for factor loading) or ΔSRMR <0.01) (for item threshold).

Therefore, our results bridge several existing gaps in terms of both theory and practice. Moreover, it should be noted that the effect (causal path) of PNT from online teaching at Time 1 on burnout at Time 2 was Gamma = 0.38, which is higher than effect reported by Huyghebaert et al. (2018) – one of the very few longitudinal studies in this field. In their study, prior PNT stemming from an individual's workplace had a significant effect on nurses' burnout 3 months later, with an effect of only 0.12. This difference may reflect the serious manner in which mandatory online teaching deeply frustrated teachers' psychological needs, resulting in an even more intense and lasting effect on teachers, as a particularly vulnerable population. As such, more longitudinal studies, related to the effects of PNT during mandatory job characteristics during pandemics, are needed to further examine the nature of the effect of PNT across different populations, including how PNT predicts and contributes to burnout or psychological distress. Given the ongoing challenges of the pandemic, and the likelihood of future waves of COVID, including newer variants, the application of the approach outlined in this study can yield important findings into the mechanisms behind the effects of pandemics on individuals in their workplace, including factors relevant to psychological needs and well-being (including distress), and intentions to continue work or, conversely, the potential for burnout.

Evaluating changes in PNT of online teaching

Given that measurement equivalence (scalar invariance) was supported for PNTSOT across different occasions, the values from the scale between two points in time (occasions) can be meaningfully compared and interpreted as reflecting a real change in the thwarting of psychological needs. Our results demonstrated that the three components of PNT of online teaching varied with context, depending on the conditions caused by the pandemic and the actions taken by local educational authorities. It is, therefore, not surprising that teachers' psychological need thwarting stemmed mainly from a stressful environment (Bartholomew et al., 2014; Franco et al., 2022), although the instructions provided in the administration of the PNTSOT asked the participants to respond based on the current situation (Time 1) when campuses were closed and mandatory online teaching was launched. However, given the need to establish test–retest reliability, this potential weakness is also a strength, in terms of the psychometric validation of the instrument. Nevertheless, although the instructions were intentionally worded in order to evaluate perceptions of the same experience (Time 1: mandatory online teaching) between occasions (i.e., for establishing longitudinal measurement invariance), differences were still found between the two occasions. We speculate that this may be due to the influence of changing circumstances (i.e., from mandatory online teaching to face-to-face teaching) on teachers' perceptions of previous experiences of psychological need thwarting related to online teaching. The existence of this unconscious influence, resulting in an evolving perception over time, is reasonable, as teachers' current environment (Time 2) resulted in less thwarting of their psychological needs for autonomy and competence, given the use of traditional instruction, thus softening their impressions of the negative feelings of thwarting regarding these two needs during the online teaching period (Time 1).

Unexpectedly, although the closure of schools reduced teachers' access to job-related resources, it also seemed to have reduced some demanding aspects of the teaching job. These findings highlight the double-edged nature of the teaching profession (Neves de Jesus and Lens, 2005). On the one hand, frequent and direct interaction with students constitutes a major motivator and source of job satisfaction for many teachers (Watt et al., 2012; Benita et al., 2019). On the other hand, the interpersonal and social nature of teaching can also serve as an occupational demand resulting in diminished well-being outcomes (Hilger et al., 2021) upon returning to an offline learning environment, as teachers' perceptions of thwarting in terms of relatedness (or interpersonal relationships during the online teaching period) were shown to have increased. This suggests that relatedness thwarting contributed more to the overall thwarting of psychological needs from Time 1 to Time 2. One possible reason is that although quarantine restrictions limited connections among teachers, these restrictions may also have reduced the frequency of negative interactions which could arise in the school or workplace. As such, in the early days of returning to the school campus, a certain amount of time to adapt and re-establish healthy interactions may be required. This effect has been noted in studies on return-to-work (Shaw et al., 2020). For example, there is a great deal of work for teachers to complete when they return to campus, including the enforcement of new, strict bio-safety protocols which can increase workload when teachers return to face-to-face environments (Silva et al., 2021). Moreover, increased requirements for group work may also affect teachers' perception of interpersonal relationships which can subconsciously influence their perceptions of previous online teaching experiences, wherein less group work was required. This result is consistent with the arguments mentioned by Yi et al. (2021) who stated that "lack of interaction with colleagues might have served as a buffer to reduce potential thwarting of relatedness needs due to less potentially negative interactions with colleagues or leaders." Furthermore, relatedness thwarting was more strongly correlated with the other two forms of PNT when schoolteachers returned to school (Time 2) as compared to during campus closure (Time 1), as is illustrated in Supplementary Table 1. We believe this finding also provides evidence to support the increased influence of relatedness thwarting on overall psychological distress, which can further and explain the unequal longitudinal measurement in terms of the constraint of variance and covariance of factors.

The measurement of PNT

In terms of the literature related to the measurement of PNT, starting from the initial assessment of athletes

(Bartholomew et al., 2011), an increasing number of studies and contexts have measured PNT, among adolescents populations (Hein et al., 2015), business environments (Lagios et al., 2022), nurses (Huyghebaert et al., 2018), as well as schoolteachers (Cuevas et al., 2015; Chen et al., 2020). Recently, in addressing the context of the COVID-19 pandemic, the focus of PNT was on the emergence of a specific demanding and unpredicted task, online teaching (Yi et al., 2021). In fact, among all contexts used for the evaluation of PNT, longitudinal designs have rarely been adopted. Therefore, our study primarily aimed to further evaluate the longitudinal reliability and validity of the PNTSOT instrument and, in this way, contribute to our understanding of the assessment of PNT which can be applied to different contexts. As such establishing the validity and reliability of the PNT instrument is fundamental to research and practice related to this emerging area of research, as compared to the emphasis prior research has largely placed on psychological need satisfaction (Huyghebaert et al., 2018; Wu et al., 2018; Lagios et al., 2022). In terms of the evaluation of the instrument, it should be noted that we found that item 8 in the competence thwarting subscale negatively affected overall measurement quality, not only making the 12-item version inferior to the 11-item version according each fit index, but also causing the 12-item version to fail to meet the criteria of measurement invariance. This result is consistent with findings related to the instrument as evaluated by Chen et al. (Chen et al., 2020), where item 8 lowered the quality of Chinese version of TPNTS.

Practical implications

The lasting harm from PNT of online teaching was demonstrated in this study, predicting teachers' willingness to continue using online teaching as well as the degree of burnout after a period of 2 months. These results highlight the delayed and long-term effect of mandatory online teaching which lasts beyond the period when online teaching is implemented – a finding which has been described by some studies (Kulikowski et al., 2021; Silva et al., 2021; Jelińska and Paradowski, 2021a,b). This finding is a warning that teachers' psychological need thwarting during online teaching is an issue of significance in terms of long-term psychological, affective, and intentional outcomes. Benefitting from the longitudinal design adopted in this study, we are able to suggest that school administrators provide teachers with effective relief of PNT during online teaching at different occasions. In terms of the three psychological needs that are potentially thwarted by online teaching, relatedness, autonomy, and competence are all relevant to teachers' psychological well-being and intention to continue online teaching. For example, in addition to the fact that teachers' need for relatedness may be thwarted due to home isolation, mandatory online teaching can further frustrate teachers' psychological needs of autonomy and competence.

In terms of competence, most teachers have not received sufficient training or experience in implementing online teaching (Trust and Whalen, 2021; Yi et al., 2021) which naturally leads to a higher perceived barrier related to technology use. This lack of technological pedagogical content knowledge also leads to teachers' frustration and loss of competence in respect to online teaching tasks. More specifically, it is more critical to address competence thwarting when online teaching, or other initiatives, are immediately enacted. In addition to providing training in technology-integrated instruction, as suggested by several studies as a means to increase teachers' experience with online teaching (Silva et al., 2021; Trust and Whalen, 2021), in order to address competence thwarting, school leaders should also be more flexible and lenient in their management, allowing more tolerance and flexibility to address (and avoid thwarting) teachers' autonomy needs during this period of uncertainty. In fact, autonomy is often considered a prerequisite for the development of competence, with scholars emphasizing the importance of "mastery experiences" in order to develop a sense of self-efficacy (Hagger et al., 2020). Mastery, in the context of the pandemic and online teaching involves several key factors, including conceptual technological pedagogical knowledge (TPK), actively engaging in tutorials to enhance digital competencies, and taking the initiative to embrace opportunities to learn (König et al., 2020). In fact, the role of TPK self-efficacy, in particular, during remote (online) teaching during the pandemic has been associated with perceived confidence as well as willingness to continue teaching online (Cahapay and Anoba, 2021) To avoid competence thwarting, and develop these key skills for developing mastery, school leadership must provide unique and complementary resources to help teachers in conducting online teaching, and most importantly, these resources should be provided continuously rather than as a short-term respond to the initial crisis (Matthews et al., 2022).

In terms of autonomy, in many cases teachers were required to use their school's designated platforms (including software) and follow prescribed course activities (including assessment methods) for online teaching which can result in perceived lack of autonomy in terms their teaching (Kulikowski et al., 2021). As stated in Kulikowski et al. (2021), restrictive rules and standards regarding online teaching methods harm teachers' autonomy. Thus, in order to address autonomy thwarting, school leaders should be more flexible and lenient in their management, allowing more tolerance and flexibility to address (and avoid thwarting) teachers' autonomy needs during this period of uncertainty, particularly as teachers voice their needs for receiving more flexibility from the administrators, such as relaxing state standards for curriculum content, flexibility with deadlines, and loosing requirements for evaluation during times of challenge (Chan et al., 2021). While organizations, such as the OECD, suggest setting school-based goals for promoting teacher autonomy, such as professional development on strategies for assisting teachers and parents in working together to implement online learning more smoothly, and the creation of professional communities of learning focused on promoting teacher autonomy (Benita et al., 2019), the cultural

context must also be considered. In the specific context of Chinese education, researchers have emphasized the potential thwarting role of specific school-level policies focused on extrinsic goal framing and a "controlling" approach in terms of student and teacher autonomy, requiring change at the organizational level (Yu et al., 2018), advocating for an awareness of the autonomy thwarting that exists within a controlling hierarchical system. As such, school administrators in Chinese schools must balance teachers' need for support with a climate that promotes voluntary participation and avoids conveying a sense of control over teachers or a reliance on extrinsic goals and rewards. Research on professional development for teachers has highlighted the importance of providing a rationale for teachers and accepting resistance, rather than forcing participation, giving teachers a chance to develop autonomy through hands-on learning supported by a warm and respectful environment in which positive feedback is provided (Aelterman et al., 2016). In sum, mitigating the risk of autonomy thwarting involves offering voluntary, flexible, and respectful opportunities to participate in communities of practice (Thornton, 2021).

From the point of view of relatedness – some literature has reported a separate spike in psychological distress among schoolteachers once schools reopen (72, 73) – the present study provides insights into the interpretation of this situation in terms of teachers' relatedness thwarting. Given the fact that teachers' perceptions of the PNT of online teaching were less severe when asked to evaluate online teaching in retrospect provides some hope of a "rebound" effect from the PNT of online teaching or negative experiences with other types of educational technologies, if teachers' psychological need thwarting can be averted. For example, there is potential in providing more opportunities for preventing relatedness thwarting through the establishing of communities of practice, simultaneously mitigating threats to competence and autonomy through targeted professional development that emphasizes not only skills and knowledge related to new technologies, but also takes into consideration potential thwarting of teachers' psychological needs. As such, the emotional care provided by school leaders is important during the early days of campus reopening. This kind of emotional care is characterized as warm and empathetic, led and modelled by front-line leaders, rather than enacted by means of an authoritarian style of leadership (Matthews et al., 2022). As such, avoiding thwarting of relatedness, and consequent turnover intention, involves leaders expressing their care for teachers and assisting teachers in maintaining balance of work and family responsibilities. From the experience of school leaders in the United Kingdom during the pandemic, several key leadership strategies were found in terms of relatedness: mitigating external pressures and expectation, adopting adaptive leadership, providing emotional guidance, working to build relationships, and maintaining resilience in an ever-changing and uncertain environment (Beauchamp et al., 2021). Thus, to avoid thwarting of teachers' relatedness needs, school leaders are encouraged to actively serve as models, maintaining relationships with teachers, parents, and students, assisting the most vulnerable, and adapting models of leadership based on present needs.

In light of the potential for thwarting of competency, autonomy, and relatedness needs, a recurring theme is the importance of professional development. Given the importance of teachers' psychological needs during challenging times, such as mandatory online teaching, the role of teacher training must also be considered. During teacher training, pre-service teachers can benefit from increased choice and freedom in pursuing individual goals (autonomy), positive feedback through coaching and mentorship which encourages student teachers to identify their unique personal qualities and incorporate these into their teaching (competence), and fostering a sense of the social environment of teaching with attention to individual students (relatedness; Evelein et al., 2008). In terms of in-service teacher training, particularly during online teaching, solutions such as personalized online workshops, characterized by "just-in-time learning," self-assessment, and flexibility in content and scheduling, should be considered to leverage the benefits of online learning wherein teachers are the students (Rhode et al., 2017). At the intersection of theory and practice, the importance of "quick response research," such as that conducted in the present study, can assist in understanding the needs and experiences of teachers as they transition from traditional to digital (online or blended) learning modes (Lockee, 2021).

Limitations and future studies

This study has several limitations. First, the major limitation of this present study is related to the sampling strategy, which relied upon with the assistance of the local government authorities, which may have unintentionally influenced teachers' responses, as well as their willingness to complete a follow-up survey after 2 months (i.e., the sample for the longitudinal portion of the study). It should be noted that, although the subjects of the longitudinal study and the subjects of the cross-sectional study shared a similar demographic background and reported similar levels of PNT of online teaching, the changes in psychological distress were different for the two samples. Whether or not this situation was due to official assistance provided through participation in the online survey is still uncertain. We suggest that future research should initiate longitudinal monitoring of teachers' mental health after they return to face-to-face teaching, and explore related factors which can influence teachers' psychological well-being and intention to continue in specific teaching tasks. Second, given that PNT of online teaching describes a perception toward one's working environment, exploring the effect of school management as a school-level variable to represent a work/environmental factor related to teachers' online teaching PNT is another potential area for future research. Third, while the present study focused on the thwarting of teachers' psychological needs as a risk factor, measures of psychological need satisfaction may be explored by future studies to examine its potential as a protective factor and its

relationship with other measures of psychological well-being regarding the use of online teaching during both times of distress, as well as under normal working conditions. Finally, while the present study was concerned with the evaluation of the direct effects of PNT of online teaching in terms of both immediate effects (including psychological distress and satisfaction with online teaching) as well as delayed effects (intention to continue using online teaching), future studies can further evaluate potential mediation and moderation effects. Moreover, future studies may test models that include alternative predictor and outcome variables in relation to the construct of PNT of online teaching.

Conclusion

In this present study, we systematically evaluated the psychometric properties of the PNTSOT instrument, as developed by Yi et al. (2021) focusing on establishing the longitudinal reliability and validity of the scale. The results demonstrated that the PNTSOT had ideal internal reliability and factorial validity among middle and primary school teachers. Moreover, the test–retest reliability was also acceptable and the tests of longitudinal measurement invariance further confirmed that the PNTSOT can be effectively used to compare the perceptions of PNT among different occasions. As such, we conclude that the PNTSOT can be relied upon as a valid instrument to predict teachers' PNT before or during the launching of online teaching, or other similar interventions. The results also remind us that, in addition to providing continual training related to online teaching, school administrators must provide more flexibility and autonomy during online teaching. Moreover, based on our findings, that the impact from PNT of online teaching is persistent and long-term, we suggest that school counselors could provide differentiated and personalized assistance to those teachers who express higher levels of psychological need thwarting during the mandatory online teaching or other similar interventions.

As noted in our results on the change in psychological distress between the two occasions, improvement in psychological well-being among schoolteachers was not found in either the cross-sectional or longitudinal data. This finding is consistent with the research from Spain (Ozamiz-Etxebarria et al., 2021a) and Denmark (Nabe-Nielsen et al., 2021). In the former study, high psychological distress still occurred when the school initially reopened, based on cross-sectional data (Ozamiz-Etxebarria et al., 2021a); while in the latter study, adopting longitudinal data, poor mental health outcomes increased extensively (from 27 to 84%) from the time of school closure to school re-opening (Nabe-Nielsen et al., 2021). From these studies, scholars identified teachers' concerns about infection and increased efforts to prevent contagion as major causes of poor mental health after campus reopening. Research into issues of teacher well-being and related influences must continue, even after vaccines become widely available, since pandemics are unpredictable and, along with other potential crises (e.g., sudden changes in the status quo for teachers) which require the adoption of new technologies or techniques, will certainly impact, and potentially thwart, teachers' psychological needs.

As we move through the next stages of the pandemic (a post-COVID-19 era), we can take advantage of the lessons learned during the pandemic as an opportunity to better evaluate the potential future of other innovations and evolutions in educational practice that promote online teaching and other interventions which can enhance teaching and learning (Gouëdard et al., 2020; UNESCO, 2020; Munoz-Najar et al., 2021; Rodriguez et al., 2021). School leaders are encouraged to acknowledge that mandatory online teaching is already very stressful for teachers, and that providing sufficient autonomy is key to maintaining teachers' willingness to continue using online teaching in the future, highlighting the role school leaders can play in such periods of uncertainty and psychological risk.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Jianxi Psychological Consultant Association. The patients/participants provided their written informed consent to participate in this study.

Author contributions

I-HC and C-YL: conceptualization, formal analysis and project administration. I-HC, X-MC, and C-YL: methodology. I-HC and X-MC: validation. I-HC, K-YZ, and Z-HW: investigation. I-HC and X-LL: resources. X-LL: data curation. I-HC and JG: supervision, writing—review and editing and writing—original draft preparation. I-HC: funding acquisition. All authors have read and approved the final manuscript.

Funding

This research was supported by the Anhui Province Philosophy and Social Science Planning Project "Evidence-Based Decision Making and Practice Research on Comprehensive College Entrance Examination Reform in Anhui Province" (Project No.: AHSKZ2021D12) and the 2022 Shanghai Philosophy and Social Science General Project for Educational Planning

"Research on the TPACK Development Mechanisms and Enhancement Paths of College Teachers in the Post-Epidemic Era" (Project No.: A2022014).

Acknowledgments

The authors thank all the participants who were involved in the present study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.943449/full#supplementary-material

References

2020 Education Statistics [Internet]. (2021). Available at: http://www.moe.gov.cn/jyb_sjzl/moe_560//quanguo/ (Accessed July 26, 2022).

Aelterman, N., Vansteenkiste, M., Van Keer, H., and Haerens, L. (2016). Changing teachers' beliefs regarding autonomy support and structure: the role of experienced psychological need satisfaction in teacher training. *Psychol. Sport Exerc.* 23, 64–72. doi: 10.1016/j.psychsport.2015.10.007

Alhazmi, R. A., Alghadeer, S., Al-Arifi, M. N., Alamer, A. A., Mubarak, A. M., Alwhaibi, A., et al. (2022). Prevalence and factors of anxiety during the coronavirus-2019 pandemic among teachers in Saudi Arabia. *Front. Pub. Health* 10:827238. doi: 10.3389/fpubh.2022.827238

Aperribai, L., Cortabarria, L., Aguirre, T., Verche, E., and Borges, Á. (2020). Teacher's physical activity and mental health during lockdown due to the Covid-2019 pandemic. *Front. Psychol.* 11:577886. doi: 10.3389/fpsyg.2020.577886

Bank, T. W. (2021). Remote Learning During Covid-19: Lessons from Today, Principles for Tomorrow. Available at: https://www.worldbank.org/en/topic/edutech/brief/how-countries-are-using-edtech-to-support-remote-learning-during-the-covid-19-pandemic (Accessed July 26, 2022).

Bartholomew, K. J., Ntoumanis, N., Cuevas, R., and Lonsdale, C. (2014). Job pressure and ill-health in physical education teachers: the mediating role of psychological need thwarting. *Teach. Teach. Educ.* 37, 101–107. doi: 10.1016/j.tate.2013.10.006

Bartholomew, K. J., Ntoumanis, N., Ryan, R. M., and Thøgersenntoumani, C. (2011). Psychological need thwarting in the sport context: assessing the darker side of athletic experience. *J. Sport Exerc. Psychol.* 33, 75–102. doi: 10.1123/jsep.33.1.75

Beauchamp, G., Hulme, M., Clarke, L., Hamilton, L., and Harvey, J. A. (2021). 'People miss people': A study of school leadership and management in the four nations of the United Kingdom in the early stage of the COVID-19 pandemic. *Educ. Manag. Adm. Leadersh.* 49, 375–392. doi: 10.1177/1741143220987841

Benita, M., Butler, R., and Shibaz, L. (2019). Outcomes and antecedents of teacher depersonalization: the role of intrinsic orientation for teaching. *J. Educ. Psychol.* 111, 1103–1118. doi: 10.1037/edu0000328

Cahapay, M. B., and Anoba, J. L. D. (2021). Technological pedagogical knowledge self-efficacy and continuance intention of Philippine teachers in remote education amid COVID-19 crisis. *J. Pedagog. Res.* 5, 68–79. doi: 10.33902/JPR.2021370614

Chan, M. K., Sharkey, J. D., Lawrie, S. I., Arch, D. A. N., and Nylund-Gibson, K. (2021). Elementary school teacher well-being and supportive measures amid Covid-19: an exploratory study. *Sch. Psychol.* 36, 533–545. doi: 10.1037/spq0000441

Chan, R. C. K., Xu, T., Huang, J., Wang, Y., Zhao, Q., Shum, D. H. K., et al. (2012). Extending the utility of the depression anxiety stress scale by examining its psychometric properties in Chinese settings. *Psychiatry Res.* 200, 879–883. doi: 10.1016/j.psychres.2012.06.041

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model. Multidiscip. J.* 14, 464–504. doi: 10.1080/10705510701301834

Chen, L., Gamble, J. H., Chen, I.-H., Lee, Z.-H., and Fu, Q.-L. (2020). Psychometric properties and measurement invariance for a Chinese version of a psychological need thwarting scale for teachers. *Int. J. Environ. Res. Public Health* 17:3247. doi: 10.3390/ijerph17093247

Chen, H., Zhang, M., Su, L., Cao, H., Zhou, X., Gu, Z., et al. (2021). Knowledge, attitudes, and practices toward Covid-19 among Chinese teachers, Shenzhen: an online cross-sectional study during the global outbreak of Covid-19. *Front. Public Health* 9:706830. doi: 10.3389/fpubh.2021.706830

Collie, R. J. (2021). Covid-19 and teachers' somatic burden, stress, and emotional exhaustion: examining the role of principal leadership and workplace buoyancy. *AERA Open* 7:6187. doi: 10.1177/2332858420986187

Costa, S., Gugliandolo, M. C., Barberis, N., Cuzzocrea, F., and Liga, F. (2019). Antecedents and consequences of parental psychological control and autonomy support: the role of psychological basic needs. *J. Soc. Pers. Relatsh.* 36, 1168–1189. doi: 10.1177/0265407518756778

Cuevas, R., Sánchez-Oliva, D., Bartholomew, K. J., Ntoumanis, N., and García-Calvo, T. (2015). Adaptation and validation of the psychological need thwarting scale in Spanish physical education teachers. *Span. J. Psychol.* 18:E53. doi: 10.1017/sjp.2015.56

Deci, E. L., and Ryan, R. M. (2015). Self-determination theory. *Int. Encyc. Soc. Behav. Sci.* 91, 486–491. doi: 10.1037//0003-066x.55.1.68

Ebersold, S., Rahm, T., and Heise, E. (2019). Autonomy support and well-being in teachers: differential mediations through basic psychological need satisfaction and frustration. *Soc. Psychol. Educ.* 22, 921–942. doi: 10.1007/s11218-019-09499-1

Evelein, F., Korthagen, F., and Brekelmans, M. (2008). Fulfilment of the basic psychological needs of student teachers during their first teaching experiences. *Teach. Teach. Educ.* 24, 1137–1148. doi: 10.1016/j.tate.2007.09.001

Fauzi, I., and Khusuma, I. H. S. (2020). Teachers' elementary school in online learning of Covid-19 pandemic conditions. *J Jurnal Iqra': Kajian Ilmu Pendidikan* 5, 58–70. doi: 10.25217/ji.v5i1.914

Franco, E., Cuevas, R., Coterón, J., and Spray, C. (2022). Work pressures stemming from school authorities and burnout among physical education teachers: the mediating role of psychological needs thwarting. *J. Teach. Phys. Educ.* 41, 110–120. doi: 10.1123/jtpe.2020-0070

Gouëdard, P., Pont, B., and Viennet, R. (2020). Education responses to Covid-19: implementing a way forward. *OECD Education Working Papers, No. 224, OECD Publishing*. doi: 10.1787/8e95f977-en

Gunnell, K. E., Crocker, P. R., Wilson, P. M., Mack, D. E., and Zumbo, B. D. (2013). Psychological need satisfaction and thwarting: a test of basic psychological needs theory in physical activity contexts. *Psychol. Sport Exerc.* 14, 599–607. doi: 10.1016/j.psychsport.2013.03.007

Hagger, M. S., Hankonen, N. E., Chatzisarantis, N. L., and Ryan, R. M. (2020). "Changing behavior using self-determination theory," in *The Handbook of Behavior change*. eds. M. S. Hagger, L. D. Cameron, K. Hamilton, N. Hankonen and T. Lintunen (Cambridge, UK: Cambridge University Press), 104–119.

Haug, E., Mæland, S., Lehmann, S., Bjørknes, R., Fadnes, L. T., Sandal, G. M., et al. (2022). Increased gaming during Covid-19 predicts physical inactivity among youth

in Norway—a two-wave longitudinal cohort study. *Front. Public Health* 10:812932. doi: 10.3389/fpubh.2022.812932

Hein, V., Koka, A., and Hagger, M. S. (2015). Relationships between perceived teachers' controlling behaviour, psychological need thwarting, anger and bullying behaviour in high-school students. *J. Adolesc.* 42, 103–114. doi: 10.1016/j.adolescence.2015.04.003

Hilger, K. J. E., Scheibe, S., Frenzel, A. C., and Keller, M. M. (2021). Exceptional circumstances: changes in teachers' work characteristics and well-being during Covid-19 lockdown. *Sch. Psychol.* 36, 516–532. doi: 10.1037/spq0000457

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6, 1–55. doi: 10.1080/10705519909540118

Huyghebaert, T., Gillet, N., Fernet, C., Lahiani, F.-J., and Fouquereau, E. (2018). Leveraging psychosocial safety climate to prevent ill-being: the mediating role of psychological need thwarting. *J. Vocat. Behav.* 107, 111–125. doi: 10.1016/j.jvb.2018.03.010

Jelińska, M., and Paradowski, M. B. (2021a). The impact of demographics, life and work circumstances on college and university instructors, well-being during quaranteaching. *Front. Psychol.* 12:643229. doi: 10.3389/fpsyg.2021.643229

Jelińska, M., and Paradowski, M. B. (2021b). Teachers' engagement in and coping with emergency remote instruction during Covid-19-induced school closures: a multinational contextual perspective. *Online Learn. J.* 25, 303–328. doi: 10.24059/olj.v25i1.2492

König, J., Jäger-Biela, D. J., and Glutsch, N. (2020). Adapting to online teaching during COVID-19 school closure: teacher education and teacher competence effects among early career teachers in Germany. *Eur. J. Teach. Educ.* 43, 608–622. doi: 10.1080/02619768.2020.1809650

Kulikowski, K., Przytuła, S., and Sułkowski, Ł. (2021). E-learning? Never again! On the unintended consequences of Covid-19 forced E-learning on academic teacher motivational job characteristics. *High. Educ. Q.* 76, 174–189. doi: 10.1111/hequ.12314

Lagios, C., Caesens, G., Nguyen, N., and Stinglhamber, F. (2022). Explaining the negative consequences of organizational dehumanization: the mediating role of psychological need thwarting. *J. Pers. Psychol.* 21, 86–93. doi: 10.1027/1866-5888/a000286

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. *Behav. Res. Methods* 48, 936–949. doi: 10.3758/s13428-015-0619-7

Lockee, B. B. (2021). Shifting digital, shifting context:(re) considering teacher professional development for online and blended learning in the COVID-19 era. *Educ. Technol. Res. Dev.* 69, 17–20. doi: 10.1007/s11423-020-09836-8

Lovibond, S. H., and Lovibond, P. F. *Manual for the depression Anxiety Stress scales*. Sydney: Psychology Foundation of Australia. (1996).

Maslach, C., and Jackson, S. E. (1981). The measurement of experienced burnout. *J. Organ. Behav.* 2, 99–113. doi: 10.1002/job.4030020205

Matthews, R. A., Wayne, J. H., Smith, C., Casper, W. J., Wang, Y. R., and Streit, J. (2022). Resign or carry-on? District and principal leadership as drivers of change in teacher turnover intentions during the COVID-19 crisis: a latent growth model examination. *J. Occup. Organ. Psychol.* 95, 687–717. doi: 10.1111/joop.12397

Mishra, L., Gupta, T., and Shree, A. (2020). Online teaching-learning in higher education during lockdown period of Covid-19 pandemic. *Int. J. Educ. Res. Open* 1:100012. doi: 10.1016/j.ijedro.2020.100012

Moè, A., and Katz, I. (2020). Self-compassionate teachers are more autonomy supportive and structuring whereas self-derogating teachers are more controlling and chaotic: the mediating role of need satisfaction and burnout. *Teach. Teach. Educ.* 96:103173. doi: 10.1016/j.tate.2020.103173

Munoz-Najar, A., Gilberto, A., Hasan, A., Cobo, C., Azevedo, J. P., and Akmal, M. *Remote Learning during Covid-19: Lessons from Today*, Principles for Tomorrow. Washington, DC. (2021). Available at: https://openknowledge.worldbank.org/handle/10986/36665 (Accessed July 26, 2022).

Nabe-Nielsen, K., Christensen, K. B., Fuglsang, N. V., Larsen, I., and Nilsson, C. J. (2021). The effect of Covid-19 on schoolteachers' emotional reactions and mental health: longitudinal results from the class study. *Int. Arch. Occup. Environ. Health* 95, 855–865. doi: 10.1007/s00420-021-01806-8

Neves de Jesus, S., and Lens, W. (2005). An integrated model for the study of teacher motivation. *Appl. Psychol.* 54, 119–134. doi: 10.1111/j.1464-0597.2005.00199.x

Ozamiz-Etxebarria, N., Berasategi Santxo, N., Idoiaga Mondragon, N., and Dosil, S. M. (2021a). The psychological state of teachers during the Covid-19 crisis: the challenge of returning to face-to-face teaching. *Front. Psychol.* 11:620718. doi: 10.3389/fpsyg.2020.620718

Ozamiz-Etxebarria, N., Idoiaga Mondragon, N., Bueno-Notivol, J., Pérez-Moreno, M., and Santabárbara, J. (2021b). Prevalence of anxiety, depression, and stress among teachers during the Covid-19 pandemic: a rapid systematic review with meta-analysis. *Brain Sci.* 11:1172. doi: 10.3390/brainsci11091172

Palma-Vasquez, C., Carrasco, D., and Hernando-Rodriguez, J. C. (2021). Mental health of teachers who have teleworked due to Covid-19. *Eur. J. Investig. Health Psychol. Educ.* 11, 515–528. doi: 10.3390/ejihpe11020037

Panisoara, I. O., Lazar, I., Panisoara, G., Chirca, R., and Ursu, A. S. (2020). Motivation and continuance intention towards online instruction among teachers during the Covid-19 pandemic: the mediating effect of burnout and technostress. *Int. J. Environ. Res. Public Health* 17, 8002. doi: 10.3390/ijerph17218002

Rhode, J., Richter, S., and Miller, T. (2017). Designing personalized online teaching professional development through self-assessment. *Tech. Trends* 61, 444–451. doi: 10.1007/s11528-017-0211-3

Richmond, G., Cho, C., Gallagher, H. A., He, Y., and Petchauer, E. (2020). The critical need for pause in the Covid-19 era. *J. Teach. Educ.* 71, 375–378. doi: 10.1177/0022487120938888

Rodriguez, M. B., Cobo, C., Munoz-Najar, A., and Ciarrusta, I. S. *Remote Learning during the Global school Lockdown*. Washington, DC: The World Bank Group. (2021). Available at: https://openknowledge.worldbank.org/handle/10986/36141 (Accessed July 26, 2022).

Rodríguez-Meirinhos, A., Antolín-Suárez, L., Brenning, K., Vansteenkiste, M., and Oliva, A. (2020). A bright and a dark path to adolescents' functioning: the role of need satisfaction and need frustration across gender, age, and socioeconomic status. *J. Happiness Stud.* 21, 95–116. doi: 10.1007/s10902-018-00072-9

Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78.

Sahu, P. (2020). Closure of universities due to coronavirus disease 2019 (Covid-19): impact on education and mental health of students and academic staff. *J. Cureus* 12:e7541. doi: 10.7759/cureus.7541

Shamir-Inbal, T., and Blau, I. (2021). Facilitating emergency remote K-12 teaching in computing-enhanced virtual learning environments during Covid-19 pandemic: blessing or curse? *J. Educ. Comput. Res.* 59, 1243–1271. doi: 10.1177/0735633121992781

Shaw, R., Kim, Y.-k., and Hua, J. (2020). Governance, technology and citizen behavior in pandemic: lessons from Covid-19 in East Asia. Progress in disaster. *Science* 6, 100090. doi: 10.1016/j.pdisas.2020.100090

Silva, D. F. O., Cobucci, R. N., Lima, S. C. V. C., and de Andrade, F. B. (2021). Prevalence of anxiety, depression, and stress among teachers during the Covid-19 pandemic: a prisma-compliant systematic review. *Medicine (Baltimore)* 100, e27684. doi: 10.1097/MD.0000000000027684

Soncini, A., Politi, E., and Matteucci, M. C. (2021). Teachers navigating distance learning during Covid-19 without feeling emotionally exhausted: the protective role of self-efficacy. *Sch. Psychol.* 36, 494–503. doi: 10.1037/spq0000469

Statistics for National Statistics (2021). Remote schooling through the coronavirus (Covid-19) pandemic, England, April 2020 to June 2021. Available at: https://www.gov.uk/government/statistics/remote-schooling-through-the-coronavirus-covid-19-pandemic-england-april-2020-to-june-2021 (Accessed July 26, 2022).

Thornton, K. (2021). Leading through COVID-19: New Zealand secondary principals describe their reality. *Educ. Manag. Adm. Leadersh.* 49, 393–409. doi: 10.1177/1741143220985110

Trust, T., and Whalen, J. (2021). K-12 teachers'experiences and challenges with using technology for emergency remote teaching during the Covid-19 pandemic. *Int. J. Educ. Technol.* 29, 10–25. doi: 10.17471/2499-4324/1192

UNESCO (2020). Global monitoring of school closures caused by Covid-19. Available at: https://en.unesco.org/covid19/educationresponse#schoolclosures (Accessed July 26, 2022).

UNESCO (2020). Covid-19 webinar: a new world for teachers, education's frontline workers: Covid-19 education webinar #2. Available at: https://en.unesco.org/news/covid-19-webinar-new-world-teachers-educations-frontline-workers-covid-19-education-webinar-2 (Accessed July 26, 2022).

Vargas Rubilar, N., and Oros, L. B. (2021). Stress and burnout in teachers during times of pandemic. *Front. Psychol.* 12:756007. doi: 10.3389/fpsyg.2021.756007

Wang, K., Shi, H.-S., Geng, F.-L., Zou, L.-Q., Tan, S.-P., Wang, Y., et al. (2015). Cross-cultural validation of the depression anxiety stress scale–21 in China. *Psychol. Assess.* 28:e88. doi: 10.1037/pas0000207

Warburton, V. E., Wang, J. C., Bartholomew, K. J., Tuff, R. L., and Bishop, K. (2020). Need satisfaction and need frustration as distinct and potentially co-occurring constructs: need profiles examined in physical education and sport. *Motiv. Emot.* 44, 54–66. doi: 10.1007/s11031-019-09798-2

Watt, H. M. G., Richardson, P. W., Klusmann, U., Kunter, M., Beyer, B., Trautwein, U., et al. (2012). Motivations for choosing teaching as a career: an international comparison using the fit-choice scale. *Teach. Teach. Educ.* 28, 791–805. doi: 10.1016/j.tate.2012.03.003

Wu, X.-C., Qi, Y.-J., Yu, R.-R., and Zang, W.-W. (2016). Revision of Chinese primary and secondary school teachers'job burnout questionnaire. *Chin. J. Clin. Psych.* 24, 856–860.

Wu, C.-Z., Rong, S., Zhu, F.-T., Chen, Y., and Guo, Y.-Y. (2018). Basic psychological need and its satisfaction. *Adv. Psychol. Sci.* 26, 1063–1073. doi: 10.3724/SP.J.1042.2018.01063

Yeung, A. Y., Yuliawati, L., and Cheung, S. H. (2020). A systematic review and meta-analytic factor analysis of the depression anxiety stress scales. *Clin. Psychol. Sci. Pract.*:e12362, 27. doi: 10.1037/h0101782

Yi, J., Chen, I.-H., Lin, C.-Y., Li, C.-C., Liao, X.-L., Wei, Z.-H., et al. (2021). The effect of primary and middle school teachers' problematic internet use and fear of Covid-19 on psychological need thwarting of online teaching and psychological distress. *Healthcare* 9, 1199. doi: 10.3390/healthcare9091199

Yu, S., Chen, B., Levesque-Bristol, C., and Vansteenkiste, M. (2018). Chinese education examined via the lens of self-determination. *Educ. Psychol. Rev.* 30, 177–214.

Zanon, C., Brenner, R. E., Baptista, M. N., Vogel, D. L., Rubin, M., Al-Darmaki, F. R., et al. (2020). Examining the dimensionality, reliability, and invariance of the depression, anxiety, and stress scale–21 (DASS-21) across eight countries. *Assessment* 28, 1531–1544. doi: 10.1177/1073191119887449

Zheng, Z., and Song, N. (2020). The real dilemma of online teaching in primary and secondary schools in Western China and its beyond path. *China Educ. Technol.* 12, 22–28.

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 13 September 2022
DOI 10.3389/feduc.2022.938078

OPEN ACCESS

EDITED BY
Jose Juan Gongora,
Tecnológico de Monterrey, Mexico

REVIEWED BY
Pradeep Kumar Choudhury,
Jawaharlal Nehru University, India
Laura Giusti,
University of L'Aquila, Italy
Stephen Adah,
Cross River University of Technology, Nigeria

*CORRESPONDENCE
Cecilia O. Nja
njacecilia@gmail.com

SPECIALTY SECTION
This article was submitted to Language, Culture and Diversity, a section of the journal Frontiers in Education

RECEIVED 06 May 2022
ACCEPTED 05 August 2022
PUBLISHED 13 September 2022

CITATION
Nja CO, Neji HA, Orim RE, Ukwetang JO, Ideba MA, Cornelius-Ukpepi B and Ndifon RA (2022) The socio-economic rank of parents and students' academic and cognitive outcomes: Examining the physical, psychological and social mediators.
Front. Educ. 7:938078.
doi: 10.3389/feduc.2022.938078

The socio-economic rank of parents and students' academic and cognitive outcomes: Examining the physical, psychological and social mediators

Cecilia O. Nja[1]*, Hope A. Neji[1], Richard E. Orim[1], John O. Ukwetang[2], Mary A. Ideba[1], Bernedette Cornelius-Ukpepi[2] and Rita A. Ndifon[2]

[1]Department of Science Education, University of Calabar, Calabar, Nigeria, [2]Department of Curriculum and Teaching, University of Calabar, Calabar, Nigeria

This study examined how the socioeconomic rank of parents correlates with students' academic and cognitive outcomes of science students in senior secondary school. Its objective was to examine the bedrock of physical and psychosocial mediators that influence students' learning and cognitive attitude. The sample comprised 548 science students drawn from 11 secondary schools in Calabar Municipality of Cross River State, Nigeria. A simple random sampling technique was used to select the sample from a population of 938 students. A cross-sectional observational type of survey design was used in this study. A self-reporting questionnaire labeled Socioeconomic Rank and Students Outcome Questionnaire (SERSOQ) was used for the study after validation and reliability. The results for reliability coefficients for SERSOQ range from 0.66 to 0.89 for Cronbach's alpha and 0.72–0.81 for Kuder Richardson's formula-20. Section "A" of SERSOQ was administered to the students in their schools by the research assistants, and students took section "B" home to their parents. Analysis of data collected was done using regression analysis, percentage, and mean. Results showed a great correlation between family income and academic achievement, cognitive attitude, and study habits. The study did not find a significant relationship between assignments with the variables under investigation. Importantly, the findings of this study found that parental control exhibited the greatest mediating function in providing family income impact on students' cognitive attitude. Other mediators like students' and peers' educational ambitions and mother-child verbal relationships were discovered as potent mediators. Findings also showed a slight impact of family income on parent-child and mother-father relationships. Parental control consists of an influential setting that is outside the school environment yet mounting a very powerful effect on determining school outcomes in teenagers. In conclusion, a

positive social environment is necessary for enhancing science students' cognitive ability, academic achievement, and study habits as money is not everything. Some of the recommendations made were that there should be an educative environment at home. Educators should encourage parents to provide the necessary means of academic success, such as a source of light, stationery, books, separate study rooms, and homework facilities in their respective homes.

KEYWORDS

cognitive ability, academic achievement, assignment, study habits, family income

Introduction

The type of family a child is born into is capable of influencing the overall development and growth of the child. This type of environment is recognized in terms of the social rank and the economic standing of the child's parents. Studies abound on how family background plays a vital role in the development of mental, emotional, physical, cognitive, and psychological, as well as their academic achievement and learning outcomes (Berkowitz et al., 2017; Lawson and Farah, 2017; Wen, 2017; Poulain et al., 2019; Nja et al., 2021). Families' dissimilitude in the developmental path shapes children differently and thereby resulting in man's capital formation and socioeconomic rank (SER) attainment in adulthood. This gives rise to a recurring generational rank that reproduces severe disparity (Duncan et al., 2010).

The basis of family background is SER of parents. The family structure that an individual comes from and its effect on learners' behaviors, as well as their academic achievement, has because a course of concern in studies that deal with social stratification (Black and Devereux, 2011), learners' growth (Wen and Lin, 2012), and academic outcomes (Bailey and Dynarksi, 2011).

The socioeconomic status of parents is represented by parents' social, economic, and cultural status index. It is made up of the occupation, level of education, family wealth, and culture of parents, as well as home educational resources (Organization for Economic Cooperation and Development [OECD], 2017). The SER comprised two factors: The social and the economic factors. The social status of an individual is the position a person occupies in society by acquisition, and the wealth of a person is the economic status (Miftahu and Melaiye, 2021).

A child's ability to excel in school is dependent on the extent to which the child was successfully managed by his/her parents in the home environment (Pant, 2020). **Many studies indicated** that the socioeconomic status of parents significantly contributed to learners' outcomes in the educational institution (Qasem, 2018; Fekadu et al., 2019; Maghra et al., 2019). Since low socioeconomic status families group tend not to have economic resources or do not have time to give their children, they needed academic support. Students from poor homes are most often exposed to feeding which is malnourish and thereby affects their cognitive functioning (Asiegbu and Ezeugbor, 2018). The level of vocabulary attained by students also influences academic ability, and openness to language is probably low in low socioeconomic cases (Pungello et al., 2009). Wadsworth and Raviv (2008) suggested that children from parents of low SER living in constant poverty grow up having physical, psychological, and educational health issues.

The problem with larger social digital inequality is that it hinders the implementation of distance learning as it is only the privileged few that can continue distance learning without dropping out of school (Aldama, 2020; Sindiani et al., 2020). The "homework gap" is very prominent as students are faced with no access to a high-speed connection in their homes, thereby not being able to perform their homework (Kelly, 2020). Investigation of the academic performance of students during the COVID-19 pandemic outbreak using distance education indicated that studying alone at one's parents' home instead of studying with friends increased the likelihood of poor academic performance. Students preferred the traditional face-to-face teaching method over the solo online teaching methods, implying that socialization is important in academic achievement (Alalawne and Tawalbeh, 2020; Giusti et al., 2021).

Academic and cognitive outcomes of secondary school science students are germane since they form the bedrock for man's capital and foretell an adulthood rank as well as the quality of life as gauged by various indicators like SER, family structure, and health (Hackman et al., 2010; Torr, 2011; Kell et al., 2013; Adler, 2013; Nja and Sampson, 2020). Erola et al. (2016) study indicated that more than half of the variance in the family level of children's SER is attributed to parents' SER. In another study by Chmielewski (2019), it was reported that the inequality between "they have" and they "have not" in terms of the academic achievement of low- and high-income SER origin has widened globally even though there is an increased opportunity to formal education.

Literature on the findings of different studies on the influence of the socioeconomic status of parents on the academic achievement of their wards among secondary schools

indicated a positive correlation between the socioeconomic status of parents and academic achievement (Mwariri et al., 2017; Onwukwe et al., 2017; Ovansa, 2017; Asiegbu and Ezeugbor, 2018; Esther et al., 2018; Osei-Owusu et al., 2018; Qasem, 2018; Fekadu et al., 2019; Maghra et al., 2019; Pant, 2020; Miftahu and Melaiye, 2021). Fekadu et al. (2019) study on the SER of parents' influence on their secondary school student's academic achievement indicated that parent income, occupation, and educational level made a significant contribution to students' academic achievement. Parents' educational level impacted more on the academic performance of secondary school students than their parents' occupation and income (Mwariri et al., 2017). In a related study by Pant (2020), the findings on the relationship between parental socioeconomic status and academic achievement of students showed that most of the students from low socioeconomic status have poor academic achievement. Miftahu and Melaiye (2021) study indicated that parents' occupation did not influence their children's academic achievement in secondary school, but their income affected their students. The income of parents is needed to pay the necessary levy and fees needed for their education. Parental care, good home parental practices, adequate facilities at home, involvement in the education of their students, and income enhanced their children's academic achievement (Mwariri et al., 2017; Osei-Owusu et al., 2018; Qasem, 2018).

The findings in Sirin (2005) research on socioeconomic status and academic achievement indicated that many studies combine one or more factors including parents' education, occupation, and income; others include parental expectations. This paper looked at SER in terms of parents' education, income, home facilities, and educational resources.

The obvious fact of the dissimilitude in socially relevant attributes that are related to family background prompted the curiosity of the researchers to investigate this inclination as they underscore social justice and impede growth (Adler, 2013; Jackson, 2013). Hitherto, studies centered on the description that is related to family SER and students' academic achievement. Recently, SER's influence on adulthood and investigation of the mechanisms through which these relationships occur are being studied. The majority of the studies were done with a sample drawn from international countries. The culturally based home environment is the fundamental pathway that connects family SER to students' academic and cognitive outcomes and, therefore, should be studied to ascertain this relationship in their unique setting (Lareau, 2002).

The majority of the studies carried out earlier are multifaceted and dealt with one dependent variable even though learners are enveloped in diverse ecological systems that are concurrently affected by external variables in many environments (Bronfenbrenner, 1979). The ability to provide relevant, specific areas of family SER background is the first step in the right direction to develop efficient remediation aimed at the intervening route to minimize perjured diverseness.

This paper addresses these questions in Nigeria. The study attempted to fill the gaps by the analysis of the relationships between family income and learners' academic achievement, cognitive attitude, students' study habits, and multidimensional routes as the key to the relationships. This paper also focused on the design and mediators of the intergenerational dissemination of the merit or demerit on the whole, and how family income affects senior secondary school students' academic and cognitive outcomes in the Nigeria setting.

Theoretical underpinning. A child is reared in a family, and the family is a multidimensional system that is made of the very near social environment. Theories of sociology and psychology of development have furnished good conceptual frameworks on how a family impacts children's growth. It may not be out of place to say that the family's economic power enhances students' development since their parents purchase whatever they need (Kaushal et al., 2011). Students whose parents earn high income most probably will live in affluent environments and will have all their educational materials like computers books, reading tables, internet services, and so on at home. Such children will attend the best schools and will have home teachers for extra tutelage (Chin and Phillips, 2004), these activities that learners engaged in at home stimulates cognitive growth which enhances children's academic achievement.

In the meantime, impalpable benefits in the home, although cannot be directly consumed or measured by money, are crucial in a child's development (Heckman, 2006). Social resources of the family, which parents practice, and the cultures presented in terms of the beliefs and values system, as well as the characters exhibited in the home environment, can also affect students' learning in school (Bourdieu, 1984). Developmental theories of children have enumerated the advantages of a democratic family setting. In this parenting style, the environment is such that there is a combination of warmth, responsiveness is high, and children only make reasonable demands. The implication of using this style of parenting is that parents provide their children with love, support, and self-governance, as well as they set realistic goals for their children (Pinquart, 2016, 2017; Kuppens and Ceulemans, 2019).

Studies have indicated the relationship between SER and children's upbringing. It has been reported that lower-SER parents are more likely to be harsh and punitive compared to higher-SER parents. Roubinov and Boyce (2017) study on parental SER and parenting practices indicated that parents with low SER are not happy parents and, as such, are harsh and bully their children more than the high-SER parents. Family conflicts are more prevalent among the low SER, giving rise to low levels of support for their children, and, also, the risk of exposure of children to family violence is high (Repetti et al., 2002).

The socioeconomic rank of parents has the capability of influencing the development of their children's outcomes

through a student's agency. For instance, studies have reported that students' academic ambition is a propelling factor in their eventual educational attainment and academic achievement (Khattab, 2015). Burger and Walk (2016) study on students' agency evaluated by their self-control, self-concept, and work value positively correlated with the social class of the students and their academic performances. Hitherto, studies have targeted majorly on the influence of external and contextual factors in the intergenerational distribution of ranks and little or less focus on students' role in the distribution.

Bandura's social learning theory (Bandura, 1977) has intensified the influence of other factors that play a vital role in the development of children. Other than parents and all others living in the home, peers make up a significant key group. This is relevant for adolescents because, at that stage in life, they have begun yearning for their own identity and rank. Studies on the effect of peer influence on students' learning, educational pursuit, and educational outcomes abound (Wilkinson et al., 2000). It can be contended that the impact of peer influence and family impact are interwoven; this is so as families most times influence the formation of children's peers through identifying which schools their children should attend, the type of neighborhoods, as well as extracurricular activities outside the school environment. Despite this, studies indicated that peers' educational achievement ambition influences one's performance without recourse to family and school impacts (Hoxby, 2000).

Many studies have not been done to concurrently examine the intervening role of both physical and material resources of the home environment, socialization patterns in the home, students' agency, and peers' impact and their influence on students' academic outcomes. The literature reviewed so far is majorly from studies conducted in developed countries and the Western world and, therefore, there is a need to carry out this type of research in the third world and African cultural contexts.

The Nigeria families scenario

Nigerian families today are faced with frail merits of education that have disconnected students from economic recourses. In Nigeria, instead of people getting access to education, they get access to poverty. Unequal educational opportunities and children's poverty are like Siamese twins. The disadvantaged family is seen in their children's educational prospects. Oftentimes, children whose parents have low qualifications or low-status jobs, living in dilapidated houses and poor neighborhoods, are more likely not to gain good qualifications themselves at school (Reay, 2019).

For a nation like Nigeria to advance, the education of its citizenry is a propeller for the development of individuals, society, and Nigeria in general (Olusegun, 2010). The relationship with all elements in the society in terms of social, economic, and political gains is harnessed through education as it is an important tool for social growth and capacity building and the acquisition of skills (Osonwa et al., 2013; Dagbo, 2014; Olayanju, 2014).

Studies have shown that there exists a relationship between parents' SER, parenting style, and academic achievement of secondary school students. Results obtained from the research showed that parents' SER and methods of parenting were significantly correlated to their children's academic achievement (Abdu-Raheem, 2015; Usman et al., 2016). Inasmuch as there is information in foreign countries on the effect of mediators on students' learning outcomes, little or nothing has been done in Nigeria as regards to mediators between the SER and the students' learning. This, therefore, informed this study to specifically examine potential means that are involved as the mediator variables in the investigation of students' academic and cognitive outcomes in science.

The purpose of this study was to investigate the correlation of the SER of a parent with multiple cognitive and academic outcomes of senior secondary school science students. Its objective was to examine the bedrock of physical and psychosocial mediators that influences students' learning and cognitive attitude. This paper examined four learners' outcomes: academic achievement, cognitive attitude, study habit, and assignment. This was done using secondary school science students. The null hypothesis stated that family income does not influence children's cognitive and academic achievement through better resources at home, friendly family socialization patterns, positive child ambition in terms of having higher academic aspiration, as well as peer influence investigated through peer university pursuits. This research also sought the effect of family income on learners' outcomes either directly or indirectly through many routes. These routes included a net of socio-demographic variables. The strengths of relative mediating effects were not hypothesized because the conflicting theoretical perspectives and findings from earlier work did not agree.

Materials and methods

A cross-sectional observational type of survey design was used for the study. It was suitable for this research as it enabled the researchers to analyze data across a sample population at a particular point in time and also a host of many variables at a time (Mahmutovic, 2021).

Participants and data collection

The research was conducted in Calabar Municipality Local Government Area of Cross River State, Nigeria during the 2020/2021 academic session. There are 11 public secondary schools in Calabar Municipality, with a total population of

938 students opting for science. The choice of science students was because the performance of science students in external examinations has been very poor (Nja et al., 2021). Secondary school students were used in this study as the study attempted to meet the scope of Frontiers in Education Journal, which emphasizes PreK-16 education that leads to the flourishing of all human beings. To obtain a representative sample for this study, a simple random sampling procedure was adopted in selecting the subjects for the study. One of the criteria for a student to belong to this research was that the student should have both parents living as the research required the father and the mother to respond to some items in the questionnaire.

For an equal spread of the sample across the 11 schools, 58% of the population in each school was selected for the study. This was done by writing numbers as appeared on the students' register on pieces of paper folded and put in a bucket. The research assistant blindfolded a student who was not part of the study and asked him/her to pick one at a time the folded papers from the bucket. Any paper picked was returned to the bucket after recording. This was also done until 58% of the respondents were selected; if a number was picked and it was discovered that the student's both parents are not living, the number was dropped in the bucket and another was picked. Only students whose numbers were picked were used for the research. The sample for the study was 548 senior secondary science students. Parents whose children were picked to form the sample of the research automatically became part of the sample as the research involved students with their parents.

Data collection

A questionnaire labeled SER and Students Outcome Questionnaire (SERSOQ) was the instrument that was used for data collection in this study. SERSOQ was an instrument developed by the researchers for data collection. It was made up of two sections. Section A comprised of questions for the students to respond to and Section B comprised of questions for the students' parents to respond to. SERSOQ was face and content validated by experts in test and measurement; they examined the items in the instruments and checked for their appropriateness, relevance, and coverage of the traits under consideration before carrying out reliability. Five items were deleted because they were not suitable. Ten items were modified/revised to arrive at the final number. The outcome of this study was made up of two cognitive abilities and academic achievement outcomes. Academic achievement was investigated by a student reporting his or her academic achievement scores. The questionnaire for academic achievement was made up of 3 items and had 4 responses on the Likert scale strongly agreed (SA), agreed (A), disagreed (D), and strongly disagreed (SD). SA = 4 points, A = 3 points, D = 2 points, and SD = 1 point. Students' responses to SA indicated better scores in school. The highest score a student should have was 12, and the lowest was 3; this was divided by the number of items in the questionnaire to get the actual value. The reliability test for SERSOQ was carried out with 30 science students and their parents in Calabar South Local Government Area of Cross River State, who were equivalent to the students that were used for the study but were not part of the study. This test aimed to ascertain the reliability of the instrument. The Cronbach's alpha reliability coefficient for academic achievement, during the trial test, was 0.85, which is appropriate. A reliability coefficient of 0.50 and above is good and high enough to justify the usage of an instrument (Joshua, 2005). The mean score was 2.05, just an "average" academic achievement score.

The Socio-Economic Rank and Students Outcome Questionnaire section for cognitive abilities that had 20 items was divided into two sections: Critical thinking and problem-solving abilities. It was made on a 4-point response Likert scale of SA, A, D, and SD. This was used to evaluate students' cognitive abilities. The highest score for the 20 items was 80, and the least was 20. The score was divided by the number of items; the mean was 2.88, slightly higher than average. The problem-solving questionnaire was adopted from Pandit (2011). The original questionnaire had 20 items, but this study used 10 items. The highest score for the ten items was 40, and the least was 10. The critical thinking questionnaire was adopted from Castle (2006). The questionnaire had 12 items originally, but 10 items were adapted and used in this study (**Supplementary Appendix A**). The Cronbach's alpha reliability coefficient for cognitive abilities during the trial test was 0.75.

Attitude outcomes categorized into two groups were also examined; study habit and assignment. Study habits of the students were investigated through their parents' responses on a 3-item questionnaire that used a 4-point Likert scale of SA, A, D, and SD. SA = 4, A = 3, D = 2, and DA = 1. The questions were my child is very serious with his/her study. My child does not joke with his/her studies. I will rate my child as a very serious scholar. A score of 12 (12/3 items = 4) is the highest score, and a score of 3 (3/3 items = 1) is the least score. A score of 4 indicates the most serious study habit. The mean score was 3.07, well above average.

The assignment was examined by students' ticking the 3-item questionnaire on a 4-point Likert point of SA = 4, A = 3, D = 2, and *SD* = 1, with the statement: I do my best on an assignment even when I do not like it. I do my assignment before anything else when I get back from school. Doing my assignment is not a burden to me. A higher value indicated greater assiduousness in doing an assignment. The mean score was 3.18, which was well above average. The Cronbach's alpha reliability coefficient for study habit and assignment, during the trial test, was 0.82 and 0.88, respectively.

Parents responded to the family income by responding to this statement; Tick the statement below that appeals to your income: "very difficult," "pretty difficult," "average," "pretty

affluent," and "very affluent." This was used to assess parents' absolute income. Very difficult score = 1, pretty difficult score = 2, average score = 3, pretty affluent score = 4, and very affluent score = 5. The relative income of parents was investigated through parents' responses to the statement: when you compare your income with that of others where you are resident, what would you rate your income? Low, somehow low, average, somehow high, and high. Their scores were Low = 1, low = 2, Average = 3, somehow high = 4, and high = 5. A high score implies that parents had high absolute or relative family income. The means of the two income variables were 2.26 for absolute income and 2.39 for relative income. All the income variables were about the average level. The Cronbach's alpha reliability coefficients for absolute income and relative income, during the trial test, were 0.66 and 0.71, respectively.

This mediators section of SERSOQ had nine categories. First, home superfluity took into cognizance home essentials. The questions were 9, and the respondents were requested to give either a yes or no answer to the questions. The statements were as follows: in my home, there is electricity for studying. There is pipe-borne water running in the house. I have my private toilet. I have my private bathroom. My bathroom is modern. I have a reading table. I have a computer. I have internet facilities. I have educational videos. The score on 9 items on superfluity ranges from 0 to 9. The mean score was close to 2.47. Kuder Richardson's formula-20 analysis of the reliability test of dichotomously scored data of home superfluity had a reliability coefficient of 0.81.

The second category of mediators was a family association type, which was measured using six variables: parental control, verbal relationship with mother, verbal relationship with father, affinity to mother, affinity to father, and father-mother relationship. The parental control section of SERSOQ had eight items on a 4-point Likert scale of SA = 4, A = 3, D = 2, and DA = 1. The items were my parents are strict with me on my homework and exams. My parents insist that I go to school every day. My parents monitor the time I come back from school. My parent checked who should be my friend. My parents check my dress and my appearance. My parents check the time I will be on the internet. My parents have TV watching time. My parents check my performance in school. My parents insist that I get to school before morning assembly. The lowest score for the 9/9 items was 1, and the highest was 36/9 = 4. Reliability was done using Cronbach's alpha, and the coefficient was 0.86. The mean score was 3.15, corresponding to a bit more than the third level of strictness.

Parent-child verbal relationship of SERSOQ had five items on a 4-point Likert scale of SA = 4, A = 3, D = 2, and *SD* = 1 for the student to respond. The items were My father/mother often discuss occurrences at school. My father/mother often discuss my relationships with friends. My father/mother often discuss my relationships with teachers. My father/mother often discuss my mood. My father/mother often discuss my worries or concerns. The highest score for this section of this questionnaire, which was made up of five items, was 20/5 items, and the lowest was 5/5 items. The reliability of the parent-child verbal relationship of SERSOQ was good, with a 0.80 Cronbach's alpha coefficient.

The Socio-Economic Rank and Students Outcome Questionnaire also examines students' affinity with their mothers/fathers. One question was used to check parents' affinity with their children, and it was "My association with my father is" "not close," "somehow," and "very close," and "not close," scored 1 point; "somehow," scored 2 points; and "very close," scored 3 points. "My association with my mother is" "not close," "somehow," "very close," and "not close," scored 1 point;

TABLE 1 Sample descriptive statistics.

Variable	Mean%	SD
Outcome		
Students' academic score	2.05	2.404
Students' cognitive attitude score	2.88	31.616
Students' study habit	3.07	2.861
Students' assignment	3.18	2.731
Principal predictors		
Family absolutely income	2.26	1.102
Family relative income	2.39	1.037
Mediators		
Home superfluities	2.47	0.542
Parental control	3.15	6.667
Mother-child verbal relationship	3.06	1.889
Father-child verbal relationship	2.95	2.363
Daddy-child affinity	2.18	0.757
Mummy-child affinity	2.00	0.829
Parents' relationship (good)	52.4%	0.784
Student academic pursuit	83.6%	
Not above SSS 3	19.3%	
Bachelor's degree,	40.5%	
Master's degree	38.5%	
Ph.D. degree	1.4%	
Students' close peers' academic ambition		
Control variables		
Age	2.00	0.390
Male	46.5%	
Female	53.5%	
Rural	36.70%	
Urban	63.30%	
Parental education (mother/father)		
No schooling	5.3%	
Not above SSS 3	36.6%	
Bachelor's degree	28.2%	
Master's degree	19.6%	
Ph.D. degree	10.3%	
Parents education	2.95	1.092

"somehow," scored 2 points; and "very close," scored 3 points. The results indicated that children's verbal relationship was more toward their mothers than their fathers, and their affinity to their mothers was also more than their fathers. The reliability of the parent-child verbal relationship of SERSOQ was good, with a 0.89 Cronbach's alpha coefficient. The mother-father relationship variable required students to respond to a yes or no answer to the two items; my parents do not quarrel most times (yes/any). My parents are like friends (yes/no). For the yes answer, the score is 1, and, for a no answer, the score is zero. Students responded that parents had a 52.4% good relationship. This result is a pointer to the level of parental disagreement in the home. Kuder Richardson's formula-20 analysis of the reliability test of dichotomously scored data of mother-father relationship had a reliability coefficient of 0.72.

The students' ambition section of SERSOQ was investigated by students' educational pursuit concerning students' responses to the question: "Tick the peak of educational attainments you desire" "The response categories were?" "Not above SS3," (coded 1), "bachelor's degree," (coded 2) "Master's degree," (coded 3) and "Ph.D. degree" (coded 4). The students who responded that they would like to go to university were about 80.7%.

Peer influence in this SERSOQ was examined through SSS students' ambition among their close friends in their classes. The statement for the students to respond was "How many of your best friends at school want to go to the university?" "few/none" or "many"; For "many," it was (coded 1), and few/none was (coded 0). The students responded that 83.6% of their close friends at school have the ambition of studying up to the university level.

Four demographic moderator variables were involved in this study: Location (urban/rural), age (measured in years), gender (male or female), and parents' highest educational attainment as responded by their children. These included five response levels: No school, secondary school certificate, first degree, master's degree, or Ph.D. degree.

The sample of this study was made up of students aged 12–17, having a mean age close to 14 years. Gender distribution was

TABLE 2 Regression statistics of the relationship between family income, academic achievement, cognitive ability, and assignment.

	Absolute family income				Relative family income			
	Academic achievement	**Cognitive attitude**	**Study attitude**	**Assignment**	**Academic achievement**	**Cognitive attitude**	**Study attitude**	**Assign ment**
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Absolute family income	1.669* (0.086) Beta = 638 *t* = 19.36 Sig = 0.000	8.166* (1.177) Beta = 0.285 *t* = 6.94 Sig = 0.000	0.391* (0.110) Beta = 151 *t* = 3.563 Sig = 0.000	0.042 (0.106) Beta = 0.017 *t* = 0.398 Sig = 0.691				
Relative family income					1.565* (0.098) Beta = 0.565 *t* = 16.000 Sig = 0.000	8.262* (1.250) Beta = 0.272 *t* = 6.611 Sig = 0.000	0.295* (0.117) Beta = 0.108 *t* = –2.527 Sig = 0.012	0.151 (0.112) Beta = 0.057 *t* = 1.345 Sig = 0.179
Age	0.294 (0.707) Beta = 0.017 *t* = 0.416 Sig = 0.677	0.072 (0.209) Beta = 0.015 *t* = 0.343 Sig = 0.732	130 (0.81) Beta = 0.067 *t* = 1.603 Sig = 0.110	0.095 (0.089) Beta = 0.045 *t* = 1.067 Sig = 0.287	0.048 (0.041) Beta = 0.051 *t* = 1.184 Sig = 0.237	0.083 (0.011) Beta = 0.011 *t* = 0.332 Sig = 0.740	4.453 (3.402) Beta = 0.055 *t* = 1.309 Sig = 0.191	0.042 (318) Beta = 0.006 *t* = 0.133 Sig = 0.894
Gender	–0.183 (0.198) Beta = 0.033 *t* = –0.927 Sig = –0.354	1.609 (2.700) Beta = 0.026 *t* = 0.596 Sig = 0.552	1.239* (0.246) Beta = 5.028 *t* = 5.028 Sig = 0.000	0.796* (0.241) Beta = 0.151 *t* = 3.305 Sig = 0.001	–0.293 (0.214) Beta = –0.052 *t* = –1.370 Sig = 0.171	1.594 (2.733) Beta = 0.026 *t* = 0.583 Sig = 0.560	1.351* (0.249) Beta = 0.244 *t* = 5.424 Sig = 0.000	0.910* (0.242) Beta = 0.172 *t* = 3.760 Sig = 0.000
Location	0.110 (0.196) Beta = –0.019 *t* = –0.563 Sig = 0.574	3.757 (2.673) Beta = 0.058 *t* = 1.405 Sig = 0.160	1.159* (0.245) Beta = 0.196 *t* = 4.731 Sig = 0.000	0.963* (0.238) Beta = –0.171 *t* = 4.050 Sig = 0.000	0.204 (0.210) Beta = 0.034 *t* = 0.971 Sig = 0.332	4.277 (2.684) Beta = 0.066 *t* = 1.594 Sig = 0.112	1.146* (0.247) Beta = 0.194 *t* = 4.643 Sig = 0.000	0.977* (0.237) Beta = 0.173 *t* = 4.115 Sig = 0.000
Parent education	0.046 (0.098) Beta = 0.017 *t* = 0.468 Sig = 0.640	6.210* (1.314) Beta = 0.215 *t* = 4.727 Sig = 0.000	0.137 (0.125) Beta = –0.052 *t* = 1.098 Sig = 0.272	0.010 (0.121) Beta = 0.004 *t* = 0.086 Sig = 0.932	0.099* (0.239) Beta = 0.091 *t* = 2.407 Sig = 0.016	6.735* (1.245) Beta = 0.233 *t* = 5.411 Sig = 0.000	0.218 (0.119) Beta = 0.083 *t* = 1.827 Sig = 0.068	0.027 (0.115) Beta = 0.011 *t* = 0.236 Sig = 0.813

Sample size = 548; Coefficients presented; Standard errors in parentheses; * $p < 0.05$.

TABLE 3 Regression statistics of the relationship between absolute family income and mediators hypothesized.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
	Home super fluity	**Parental control**	**Discussion with father**	**Discussion with mother**	**Affinity to mother**	**Affinity to father**	**Parents relation ship**	**Students pursuit**	**Peer influence**
A									
Absolute family income	0.101* (0.021) Beta = 0.205 t = 4.903 Sig = 0.000	2.058* (0.243) Beta = 0.340 t = 8.456 Sig = 0.000	0.072 (0.073) Beta = 0.042 t = 0.976 Sig = 0.329	0.365* (0.090) Beta = 0.170 t = 4.037 Sig = 0.000	0.170* (0.208) Beta = 0.248 t = 5.983 Sig = 0.000	0.258* (0.030) Beta = 0.343 t = 8.533 Sig = 0.000	0.259* (0.028) Beta = 0.362 t = 9.077 Sig = 0.000	0.175* (0.029) Beta = 0.249 t = 5.999 Sig =	0.016 (0.014) Beta = 0.048 t = 1.124 Sig = 0.262
Gender	0.280* (0.046) Beta = 0.267 t = 6.126 Sig = 0.000	5.72* (0.502) Beta = 0.444 t = 11.403 Sig = 0.000	0.103 (0.168) Beta = 0.028 t = 0.613 Sig = 0.540	0.729* (0.205) Beta = 0.160 t = 3.553 Sig = 0.000	0.145* (0.065) Beta = 0.099 t = 2.224 Sig = 0.027	0.415* (0.067) Beta = 0.259 t = 6.180 Sig = 0.000	0.101 (0.065) Beta = 0.066 t = 1.541 Sig = 0.124	0.059 (0.067) Beta = 0.039 t = 0.884 Sig = 0.377	0.238* (031) Beta = 0.335 t = 7.650 Sig = 0.000
Location	0.184* (0.046) Beta = t = Sig =	0.789 (0.553) Beta = 0.057 t = 1.428 Sig = 0.154	0.009 (0.167) Beta = 0.002 t = 0.055 Sig = 0.956	0.497* (0.205) Beta = 0.102 t = 2.427 Sig = 0.016	0.110 (0.065) Beta = 0.070 t = 1.695 Sig = 0.091	0.158* (0.068) Beta = 0.092 t = 2.304 Sig = 0.022	0.770* (0.056) Beta = 0.474 t = 13.792 Sig = 0.000	0.040 (0.066) Beta = 0.025 t = 0.597 Sig = 0.551	0.009 (0.032) Beta = 012 t = 0.275 Sig = 0.784
Parents education	0.284* (0.020) Beta = 0.572 t = 14.141 Sig = 0.000	1.033* (0.274) Beta = 0.169 t = 3.773 Sig = 0.000	0.028 (0.083) Beta = 0.016 t = 0.331 Sig = 0.741	0.595* (0.100) Beta = 0.275 t = 5.960 Sig = 0.000	0.082* (0.032) Beta = 0.118 t = 2.527 Sig = 0.012	0.282* (0.32) Beta = 0.371 t = 8.729 Sig = 0.000	0.379* (0.028) Beta = 0.526 t = 13.484 Sig = 0.000	0.34 (0.033) Beta = 0.049 t = 1.039 Sig = 0.299	0.111 (0.016) Beta = 0.330 t = 7.150 Sig = 0.000
R squared	0.042	0.116	0.072	0.029	0.062	0.118	0.131	0.062	0.002
B									
Relative family income	0.105* (0.022) Beta = 0.202 t = 4.819 Sig = 0.000	1.863* (0.262) Beta = 0.291 t = 7.108 Sig = 0.000	0.079 (0.078) Beta = 0.044 t = 1.018 Sig = 0.309	0.164 (0.097) Beta = 0.072 t = 1.695 Sig = 0.091	0.166* (0.030) Beta = 0.229 t = 5.489 Sig = 0.000	0.194* (0.033) Beta = 0.244 t = 5.877 Sig = 0.000	0.243* (0.31) Beta = 0.322 t = 7.939 Sig = 0.000	0.166* (0.031) Beta = 0.223 t = 5.351 Sig = 0.000	0.008 (0.015) Beta = 0.023 t = 0.527 Sig = 0.598
Gender	0.282* (0.046) Beta = 0.269 t = 6.102 Sig = 0.000	6.014 (0.512) Beta = 0.467 t = 11.744 Sig = 0.000	0.928* (0.208) Beta = 0.203 t = 4.463 Sig = 0.000	0.111 (0.170) Beta = 0.030 t = 0.656 Sig = 0.512	0.152* (0.066) Beta = 0.104 t = 2.305 Sig = 0.022	0.482* (0.069) Beta = 0.301 t = 6.961 Sig = 0.000	0.086 (0.067) Beta = 0.057 t = 1.293 Sig = 0.196	0.069 (0.068) Beta = 0.046 t = 1.020 Sig = 0.308	0.249 (0.031) Beta = 0.351 t = 7.982 Sig = 0.000
Location	0.191* (0.046) Beta = 0.171 t = 4.134 Sig = 0.000	0.682 (0.563) Beta = 0.050 t = 1.210 Sig = 0.227	0.004 (0.167) Beta = 0.001 t = 0.024 Sig = 0.016	120 (0.065) Beta = 0.077 t = 1.846 Sig = 0.065	0.501* (0.207) Beta = 0.103 t = 2.419 Sig = 0.016	0.168* (0.071) Beta = 0.098 t = 2.376 Sig = 0.018	0.786* (0.057) Beta = 0.484 t = 13.865 Sig = 0.000	0.050 (0.067) Beta = 0.031 t = 0.744 Sig = 0.457	0.009 (0.032) Beta = 0.012 t = 0.282 Sig = 0.778
Parents education	0.269* (0.019) Beta = 0.542 t = 14.078 Sig = 0.000	1.303* (0.262) Beta = 0.214 t = 4.975 Sig = 0.000	0.667* (0.095) Beta = 0.309 t = 7.042 Sig = 0.000	0.018 (0.079) Beta = 0.010 t = 0.222 Sig = 0.825	0.100* (0.031) Beta = 0.144 t = 3.248 Sig = 0.001	0.316* (0.031) Beta = 0.416 t = 10.214 Sig = 0.000	0.385* (0.027) Beta = 0.532 t = 14.464 Sig = 0.000	0.060 (0.032) Beta = 0.084 t = 1.885 Sig = 0.060	0.106* (0.015) Beta = 0.314 t = 7.154 Sig = 0.000

Regression statistics of the relationship between relative family income and the mediators hypothesized. Sample size = 548; Coefficients presented; Standard errors in parentheses; * $p < 0.05$.

perfectly balanced. About average parental education was at the secondary school certificate level.

Procedure for data collection

In carrying out this research, approval was received from the ethical committee of the Secondary School Education Board of Cross River State. The participants were intimated about the aim of the research; they were told that the exercise was purely for research purposes, and it was highly confidential and anonymous in terms of data collection and analysis. The respondents willingly gave their consent and participated in the research. This research was carried out during the first semester of the 2020/2021 academic year. Non-science teachers were used as research assistants and administered SERSOQ to the students

TABLE 4 Regression statistics of the relationship between the hypothesized mediators and cognitive and academic outcomes.

Mediators	Academic achievement	Cognitive attitude	Study habit
Home superfluity	1.005* (0.224) Beta = 0.189 t = 4.493 Sig = 0.000	5.285 (2.486) Beta = 0.091 t = 2.125 Sig = 0.034	0.239 (0.226) Beta = 0.045 t = 1.059 Sig = 0.290
Parental control	0.103* (0.018) Beta = 0.237 t = 5.705 Sig = 0.000	0.023 (0.203) Beta = 0.005 t = 0.115 Sig = 0.908	0.024 (0.018) Beta = 0.056 t = 1.300 Sig = 0.194
discussion with mother	0.189* (0.052) Beta = 0.154 t = 3.653 Sig = 0.000	0.671 (0.716) Beta = 0.070 t = 0.937 Sig = 0.349	0.019 (0.065) Beta = 0.013 t = 0.296 Sig = 0.767
Affinity with mother	1.038* (0.157) Beta = 0.272 t = 6.617 Sig = 0.000	1.087 (1.786) Beta = 0.026 t = 0.609 Sig = 0.543	0.363* (0.147) Beta = 0.105 t = 2.470 Sig = 0.014
Affinity with father	1.311* (0.138) Beta = 0.377 t = 9.508 Sig = 0.000	3.012 (1.627) Beta = 0.079 t = 1.851 Sig = 0.065	0.008 (0.162) Beta = 0.002 t = 0.052 Sig = 0.959
Parents relationship	0.759* (0.153) Beta = 0.207 t = 4.946 Sig = 0.000	15.494 (1.586) Beta = 0.386 t = 9.769 Sig = 0.000	131 (0.155) Beta = 0.036 t = 0.842 Sig = 0.400
Academic pursuit	1.142* (0.152) Beta = 0.307 t = 7.525 Sig = 0.000	1.831 (1.746) Beta = 0.045 t = 1.049 Sig = 0.295	0.096 (0.158) Beta = 0.026 t = 0.606 Sig = 0.545
Peers influence	0.484 (0.335) Beta = 0.062 t = 1.445 Sig = 0.149	7.322* (3.669) Beta = 0.085 t = 1.996 Sig = 0.046	0.932* (0.331) Beta = 0.120 t = 2.818 Sig = 0.005
Control			
Age	0.886* (0.315) Beta = 0.120 t = 2.815 Sig = 0.005	0.430 (3.476) Beta = 0.005 t = 0.124 Sig = 0.902	0.183 (0.314) Beta = 0.025 t = 0.583 Sig = 0.560
Gender	1.473* (0.230) Beta = 0.264 t = 6.399 Sig = 0.000	7.820* (2.594) Beta = 0.128 t = 3.015 Sig = 0.000	1.378* (0.229) Beta = 0.249 t = 6.010 Sig = 0.000
Parent education	0.743* (0.108) Beta = 0.281 t = 6.852 Sig = 0.000	8.693* (1.181) Beta = 0.300 t = 7.358 Sig = 0.000	0.291 (0.111) Beta = 0.111 t = 2.608 Sig = 0.009
Location	0.058 (0.225) Beta = 0.010 t = 0.230 Sig = 0.818	3.502 (2.787) Beta = 0.054 t = 1.257 Sig = 0.209	1.171* (0.248) Beta = 0.198 t = 4.730 Sig = 0.000

Sample size = 548; Coefficients presented; Standard errors in parentheses; * $p < 0.05$.

in the Assembly hall during a break period for 40 min. Science teachers were not used as their presence can elucidate biased responses from the students. Non-science teachers were used in this survey as their presence provided a familiar atmosphere for responses from the students as against the use of total strangers.

The same students took Section B-required responses from their parents at home and were brought back to school the next day. A total of 552 SERSOQs were administered, and 548 were retrieved.

Statistical analysis

The data analysis process was done by first coding the result obtained from the participants. Data analysis was done using Statistical Package for the Social Sciences software (26). A trial test was done using 30 science students and their parents who were not part of the research but were equivalent to the science students used for the research. This was used for the analysis of reliability using Cronbach's alpha coefficient for Likert scales and Kuder Richardson formula 20 for dichotomous scales. Counterfactual Variable Control (CVC) was conducted using two different counterfactual control: (i) Principal variables control only and (ii) mediator variable control only. The thought of CVC was to preserve only the strong predictions (Morgan and Winship, 2015).

Data obtained from SERSOQ were analyzed using inferential and descriptive statistics. Descriptive statistics used mean and percentage. Inferential statistics used were linear regression statistics and Sobel-mediating test analysis.

Results

Descriptive Statistics: The study involved 4 learning outcomes; the family variable was in two levels: Absolute income and relative income; these were used as major predictors, nine mediators, and six control variables.

Table 1 presents sample statistics of all the variables included in the analysis.

Family income and cognitive and academic outcomes

The regression statistics in Table 2 show that the correlation between absolute income, academic achievement, cognitive ability, and study habits was statistically significant. However, on the regression statistics in Table 2, the relationship between relative income and academic achievement, cognitive ability, and study habits was statistically significant. The regression statistics showed that the interaction between parent education and absolute income on cognitive ability is significant. Table 2

also indicates that the interaction between parent education and relative income on students' academic achievement and cognitive ability was statistically significant.

The regression statistics in Table 3 show the interaction between parents' education and absolute income on mediating variables of home superfluity), parental control discussion with father affinity to mother, affinity to father, father/mother relationship, and peer influence was statistically significant. The regression statistics in Table 3 also show that the correlation between absolute income and the father/mother relationship was statistically significant. Table 3 also shows that the interaction between parents' education and relative income on mediating variables of parental control was statistically significant.

The regression statistics in Table 4 show that the relationship between home superfluities with academic achievement was statistically significant. The regression statistics in Table 4 show that the correlation of parental control with academic achievement was statistically significant. Table 4 also indicates that the relationship between discussion with the mother and academic achievement was statistically significant. Same on that regression statistics in Table 4, the correlation between affinity with mother and academic achievement and study habits was positively correlated.

The same regression statistics in Table 4 indicate that affinity with father for academic achievement was statistically significant. The regression statistics in Table 4 indicate that the father/mother relationship and academic achievement were statistically significant. Table 4 also indicates that academic pursuit was positively correlated with academic achievement. The regression statistics in Table 4 also indicate that parent education was related to academic achievement, cognitive ability, and study habits of students. All non-significant variables were deleted from the regression analysis.

Mediating effects

The regression statistics for mediating effects in Table 5 show that, while controlling for the independent variable (absolute family income), the mediating variable (home superfluity, parental control, discussion with mother, affinity to mother, affinity with father, parents relationship, and academic pursuit) significantly predicted the dependent variable (academic achievement). When absolute family income was controlled in Table 5, the mediating variable (peers' influence) was a significant predictor of the dependent variable (cognitive ability). Table 5 also shows that, when absolute family income was controlled, parent education was a significant predictor of the dependent variable of academic achievement, cognitive ability, and study habit. The Sobel mediation test was also done individually for each mediator, and the result is presented in Table 6.

TABLE 5 Regression statistics of the mediating effects.

	School grades	Cognitive attitude	Study habit
	Model 1	Model 2	Model 3
Absolute family income	0.623* (0.119) Beta = 0.179 t = 5.228 Sig = 0.000	0.587* (0.124) Beta = 0.158 t = 4.723 Sig = 0.000	0.258* (0.030) Beta = 0.343 t = 8.533 Sig = 0.000
Home superfluity	1.005* (0.224) Beta = 0.189 t = 4.493 Sig = 0.000		
Parental control	0.103* (0.018) Beta = 0.237 t = 5.705 Sig = 0.000		
discussion with mother	0.189* (0.052) Beta = 0.154 t = 3.653 Sig = 0.000		
Affinity to mother	1.038* (0.157) Beta = 0.272 t = 6.617 Sig = 0.000		0.363* (0.147) Beta = 0.105 t = 2.470 Sig = 0.014
Affinity with father	1.311* (0.138) Beta = 0.377 t = 9.508 Sig = 0.000		
Parents relationship	0.759* (0.153) Beta = 0.207 t = 4.946 Sig = 0.000		
Academic pursuit	1.142* (0.152) Beta = 0.307 t = 7.525 Sig = 0.000		
Peers influence		7.322* (3.669) Beta = 0.085 t = 1.996 Sig = 0.046	0.932* (0.331) Beta = 0.120 t = 2.818 Sig = 0.005
Gender	1.473* (0.230) Beta = 0.264 t = 6.399 Sig = 0.000	7.820* (2.594) Beta = 0.128 t = 3.015 Sig = 0.000	1.378* (0.229) Beta = 0.249 t = 6.010 Sig = 0.000
Parent education	0.743* (0.108) Beta = 0.281 t = 6.852 Sig = 0.000	8.693* (1.181) Beta = 0.300 t = 7.358 Sig = 0.000	0.082* (0.032) Beta = 0.118 t = 2.527 Sig = 0.012
Location	1.171* (0.248) Beta = 0.198 t = 4.730 Sig = 0.000	3.502 (2.787) Beta = 0.054 t = 1.257 Sig = 0.209	1.165 (0.235) Beta = 0.078 t = 1.601 Sig = 0.126

Sample size = 548; Coefficients presented; Standard errors in parentheses; * $p < 0.05$.

TABLE 6 Proportions of total effect mediated.

Mediators	Academic achievement	Cognitive ability	Study attitude
Home superfluity	44.7%	NIL	NIL
Parental control	47.6%	NIL	NIL
Mother-child verbal relationship	22.8%	NIL	NIL
Mother-child affinity	9.3%	NIL	9.3%
Father-child affinity	8.9%	NIL	5.6%
Parental closeness	42%	NIL	NIL
Educational pursuit	NIL	15.8%	NIL
Peer influence	NIL	18.0%	12.7%
Gender	23.5%	20.2%	18.6%
Parent education	19.5%	29.5%	18.9%
Location	31.2%	NIL	23.2%

Sobel mediation test results

The mediating test results using Sobel mediation test results in **Table 6** showed that the indirect effect *via* parental control was the strongest among all the mediating effects of the total effect of family income on academic achievement. Educational aspiration and peer college aspiration were the two mediators that significantly mediated between absolute income and cognitive ability. Parent education was the largest indirect effect on cognitive ability. The location had the highest indirect effect on study habits when absolute income was controlled.

Discussion and conclusion

This study investigated the influence of the SER of parents on students' academic and cognitive outcomes in senior secondary school science in Nigeria. The findings from the analysis of data received collaborated with contemporary foreign research on this topic similar to the case of Nigeria. Broberg et al. (1997) and Reynolds and Temple (1998) investigation of the United States and Sweden indicated that previous outcomes of the children are very important. This study indicated the same result for Nigeria. The results of this research indicated that both absolute and relative income had a positive influence on cognitive ability, academic achievement, and study habits of senior secondary school science students. It also showed the indirect effect of mediating variables (home superfluities, parental control, mother-child verbal relationship, father-child verbal relationship, daddy-child affinity, mummy-child affinity, parents' relationship, a student academic pursuit, and peers' influence) on students' outcomes (students' cognitive ability, academic achievement, and study).

In addition, the influence of the home environment (mediating variables) seems even more crucial for learners' outcomes. These findings showed the essence of controlling for mediating variables to have positive learner outcomes. Hence, a study on development should permanently be longitudinal to control for this. Without that, the impact of some variables could be exaggerated. The findings indicated that social inequality exists in PreK-16 school years. Parents who are highly educated may upbring their children more positively than the low level of educated parents. This would have provided a relaxed atmosphere to encourage students' academic outcomes. In line with previous studies (Reynolds et al., 2014), friendly parent-child and mother-father relationships contributed to positively affecting students' learning outcomes, not minding the effect of SER factors. It is obvious from this study that money is not everything in child upbringing.

Becker and Tomes (1986) study in agreement with this study indicated that basic educational materials like desks, computers, and the Internet are necessary for students' academic achievement. Evans (2006) has emphasized the important role a good physical environment and good housing conditions play in the overall development of students. Furthermore, the SER of parents has a strong influence on cognitive abilities, academic achievement, and study habits. Hence, parents with a low education level should be advised to be actively involved with their children; this is so as combined activities motivate the cognitive enhancement of their wards. This study collaborated with earlier studies by Cole-Henderson (2000) and Hornby and Blackwell (2018) whose works indicated that parental involvement was associated with greater academic achievement. Parents' SER determines the type of association and the style of interaction that occurs between siblings in the family. The upper-class and middle-class children are given the freedom to decide on the home. Children are expected to take responsibility for their actions (Usman et al., 2016).

On the other hand, low-income families may not have an interest in education and, therefore, will have low educational aspirations for their children. The findings also indicated the strength of absolute income as it correlated more strongly to learners' outcomes than relative outcomes. This may be connected to the fact that what matters is if the income can provide the basic family needs, not if your income is comparable to your colleagues around the place of residence. With a good family income, children's fees are paid on time, and students will not be driven out of school. This is so as students who do not attend classes because school fees are not paid promptly do miss classes and, as such, do not do well in academic achievement. The cognitive ability of students is enhanced because, if students

come from homes where parents' income can provide the basic need, hunger which is a threat to cognitive ability is eliminated. Thus, when a child is well fed, the mental processes are active to get involved in cognitive ability. Absolute and relative incomes were positively and significantly related to parent-child verbal relationships and mother-father relationships. It can be said that a family's economic rank brought about the calmness in the minds of parents and, as such, enables the constant exchange of thoughts and feelings, increasing the affinity of the parent to the child. This would have provided a relaxed atmosphere to encourage students' cognitive and academic outcomes. Hitherto, researchers have indicated that absolute, unlike relative income, is more formidable in indicating socioeconomic predictors both in the physical or objective development of children's outcomes (Joseph et al., 2018). Zhou et al. (2019) study in agreement with this study emphasized relative income as it has prominent effects on the emotional/subjective outcomes of students. The explanation is that, while an absolute income takes care of material benefits with an emphasis on monetary gain, a relative income takes care of emotion, which is the psychological aspect of children. It is related to how children feel satisfied or deprived and can lead to diverse levels of learning and cognitive outcomes (Adler, 2013).

The inferences, which were drawn from the findings of this research, are connected to the positive effect of family income on science students' cognitive ability, academic achievement, and study habits. Family income affected students' learning directly, but there were also indirect variables that affected students' outcomes. Mediators' variables like parental control, friendly parent-child, and parental closeness contributed positively to students' learning outcomes, not minding the effect of SER factors. It is obvious from this study that money is not everything in child upbringing. Looking at the strongest and consistent mediating impact of parental control, students and peer educational ambition, as well as the mother-child verbal relationship, it can be concluded that both economic and a positive social environment are necessary to enhance science students' cognitive ability, academic achievement, and study habits.

The inferences, which were drawn from the findings of this research, are connected to the positive effect of family income on science students' cognitive ability, academic achievement, and study habits. Family income affected students' learning directly, but there were also indirect variables that affected the students' outcomes. Mediators' variables like parental control, friendly parent-child, and parental closeness contributed positively to the students' learning outcomes, not minding the effect of SER factors. It is obvious from this study that money is not everything in child upbringing. Looking at the strongest and consistent mediating impact of parental control, students and peer educational ambition, as well as the mother-child verbal relationship, it can be concluded that both economic and a positive social environment are necessary to enhance science students' cognitive ability, academic achievement, and study habits.

Implications for further research

The following implications for further research might be suggested, given the results of the study; both absolute and relative incomes were statistically significant for academic achievement and cognitive attitude. Absolute income was also significant for study habits. The result also indicated that absolute and relative income were not significant for assignment. This result implies that, when one's parents are rich and or richer than others, it enhances the academic achievement and the cognitive abilities of adolescents. When this occurs, intergenerational transmission of ranks is facilitated.

The crux of the matter in this paper is that, even though teaching and learning take place in a school setting, the home atmosphere plays a vital role in influencing learners' outcomes. Research should be conducted to investigate youth irrational behavior on what is more influential, "the school or the youth environments." This paper aims to request that studies be conducted to investigate the impact of students' wellbeing in relation to schools, families, peers, and communities on socio-psychological, cognitive advancement, and socioeconomic outcomes in Nigeria. The cause, extent, composition, and predictors of peer impact should be investigated in future studies in the Nigeria setting. This study indicated that higher family SER is positively and significantly related to higher educational ambition, yet the route-joining SER to educational ambition is not known. Therefore, a study should focus on investigating the predictors of students' outcomes, such as educational ambition, locus of control, and self-concept to supply proof of how education, family, and community can support students' knowledge that helps their socio-emotional wellbeing and ranks actualization when they become adult.

Recommendations

Keeping in view the findings of this research, the following recommendations are given below: students should be provided with a serene home environment for studies, which could help to control mediators' variables and promote students learning and cognitive outcomes. This can be done by giving proper time to children and having an educative environment at home. Educators should encourage parents to provide the necessary means of academic success, such as a source of light, stationery, books, separate study rooms, and homework facilities in their respective homes. Children should be given enough space and opportunity to air out their views on family issues. Parents should raise their children in a loving, caring, secure, consistent, and stable home environment as this will make them

develop well socially, psychologically, physically, emotionally, and morally to cope with learning outcomes.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

CN: substantial contributions to the conception and design of the work; or the acquisition, analysis or interpretation of data for the work. HN: substantial contributions to the acquisition of data for the work. RO: substantial contributions to the interpretation of data for the work. JU: substantial contributions to the coding of data for the work. MI: substantial contributions to the analysis of data for the work. BC-U: drafting the work to critically assess the important intellectual content. RN: provide approval for publication of the content of this manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2022.938078/full#supplementary-material

References

Abdu-Raheem, B. O. (2015). Parents' Socio-economic rank as predictor of secondary school students' academic performance in Ekiti State, Nigeria. *J. Educ. Pract.* 6, 123–128.

Adler, N. E. (2013). Health disparities: Taking on the challenge. *Perspect. Psychol. Sci.* 8, 679–681.

Alalawne, T., and Tawalbeh, L. I. (2020). Distance education during the COVID-19 outbreak: A cross-sectional study among medical students in North of Jordan. *Ann. Med. Surg. (Lond)* 59, 186–194.

Aldama, P. (2020). *[OPINION] What will happen to poor students when schools go online?*. Available online at: https://rappler.com/voices/ispeak/opinionpoorstudents-schools-online-coronavirus (accessed July 24, 2021).

Asiegbu, C. E., and Ezeugbor, C. O. (2018). Relationship between the socio-economic status of parents and academic performance of students in Onitsha north local government area of Anambra state. *J. Plus Educ.* 19, 166–175.

Bailey, M. J., and Dynarksi, S. M. (2011). "Inequality in postsecondary education," in *Whither opportunity? Rising inequality, schools, and children's life chance*, eds J. Greg, Duncan, J. Richard, and Murnane (New York, NY: Russell Sage Foundation), 117–132.

Bandura, A. (1977). Social cognitive theory: An agentic perspective. *Annu. Rev. Psychol.* 52, 1–26.

Becker, G. S., and Tomes, N. (1986). Human capital and the rise and fall of families. *J. Labour Econ.* 4, 1–39. doi: 10.1086/298118

Berkowitz, R., Moore, H., Astor, R. A., and Benbenishty, R. (2017). A research synthesis of the associations between socioeconomic background, inequality, school climate, and academic achievement. *Rev. Educ. Res.* 87, 425–469.

Black, S. E., and Devereux, P. J. (2011). "Recent developments in intergenerational mobility," in *Handbook of labor economics*, eds O. Ashenfelter, and D. Card, Vol. 4B, (North-Holland: Elsevier), 1487–1541.

Bourdieu, P. (1984)). *Distinction: A social critique of the judgement of taste*. Cambridge: Harvard University Press.

Broberg, A. G., Wessels, H., Lamb, M. E., and Hwang, C. P. (1997). Effects of daycare on the development of cognitive abilities in 8-year-olds: A longitudinal study. *Dev. Psychol.* 33, 62–69. doi: 10.1037//0012-1649.33.1.62

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge: Harvard University Press.

Burger, K., and Walk, W. (2016). Can children break the cycle of disadvantage? Structure and agency in the transmission of education across generations. *Soc. Psychol. Educ. Int. J.* 19, 695–713.

Castle, A. (2006). Assessment of the critical thinking skill of the student. *Sci. Direct* 12, 88–95.

Chin, T., and Phillips, M. (2004). Social reproduction and child-rearing practices: Social class, children's agency, and the summer activity gap. *Sociol. Educ.* 77, 185–210.

Chmielewski, A. K. (2019). The global increase in the socioeconomic achievement gap, 1964 to 2015. *Am. Sociol. Rev.* 84, 517–544. doi: 10.1097/OPX.0000000000000590

Cole-Henderson, B. (2000). Organizational characteristics of schools that successfully serve low-income Urban African American students. *J. Educ. Stud. Placed Risk (JESPER)* 5, 77–91.

Dagbo, S.S. (2014). Deaf children: their right to education through sign language. *J. Special Gen. Educ.* 3, 103–107.

Duncan, G. J., Morris, P. A., and Rodrigues, C. (2010). Does money matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments. *Dev. Psychol.* 47, 1263–1279. doi: 10.1037/a0023875

Erola, J., Jalonen, S., and Hannu, L. (2016). Parental education, class, and income over the early life course and children's achievement. *Res. Soc. Stratif. Mobil.* 44, 33–43.

Esther, A. E., Ruffina, A. N., and Anastecia, I. N. (2018). Influence of parents' socioeconomic status on the academic achievement in the biology of public senior secondary school students in Orlu LGA Imo state. *Br. J. Educ. Learn. Dev. Psychol.* 1, 66–73.

Evans, G. (2006). Child development and the physical environment. *Annu. Rev. Psych.* 57, 423–451.

Fekadu, A. A., Negassa, G. M., and Tegegne, M. A. (2019). The impact of parents' socioeconomic status on academic performance of grade twelve students: In case of Dire Dawa city, Ethiopia. *Int. J. Math. Stat. Stud.* 7, 28–39.

Giusti, L., Mammarella, S., Salza, A., and Vecc, S. D. (2021). Predictors of academic performance during the covid-19 outbreak: Impact of distance education on mental health, social cognition and memory abilities in an Italian university student sample. *BMC Psychol.* 9:142. doi: 10.1186/s40359-021-00649-9

Hackman,D. A., Farah, M. J. and Meaney, M. J. (2010). Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nat. Rev. Neurosci.* 11, 651–659.

Heckman, J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science* 312, 1900–1902.

Hornby, G., and Blackwell, I. (2018). Barriers to parental involvement in education: An update. *Educ. Rev.* 70, 109–119.

Hoxby, C. (2000). *Peer effects in the classroom: Learning from gender and race variation.* Available online at: https://www.nber.org/papers/w7867 (accessed April 2022).

Jackson, M. (2013). *Determined to succeed? Performance versus choice in educational attainment*. Stanford: Stanford University Press.

Joseph, G., Inácio, C. M., Günther, F., Aluisio, J. D., Barros, and Cesar, G. V. (2018). Absolute income is a better predictor of coverage by skilled birth attendance than relative wealth quintiles in a multicountry analysis: Comparison of 100 low- and middle-income countries. *BMC Pregnancy Childbirth* 18:104. doi: 10.1186/s12884-018-1734-0

Joshua, M. T. (2005). *Fundamentals of test and measurement in education*. Calabar: The university of Calabar press.

Kaushal, N., Katherine, M., and Waldfoge, l. J (2011). "How is family income related to investments in children's learning?," in *Whither opportunity? Rising inequality, schools, and children's life chances*, eds J. Greg, Duncan, J. Richard, and Murnane (New York, NY: Russell Sage Foundation), 187–205.

Kell, H. J., Lubinski, D., and Benbow, C. P. (2013). Who riser to the top? Early indicators. *Psychol. Sci.* 24, 648–659. doi: 10.1177/0956797612457784

Kelly, A. (2020). *The digital divide 'isolates and endangers' millions of the UK's poorest*. Available online at: https://www.theguardian.com/world/2020/apr/28/digital-divide-isolates-andendangers-mill (accessed July 17, 2020).

Khattab, N. (2015). Students' aspirations, expectations, and school achievement: What matters? *Br. Educ. Res. J.* 41, 731–748.

Kuppens, S., and Ceulemans, E. (2019). Parenting styles: A closer look at a well-known concept. *J. Child Fam. Stud.* 28, 168–181. doi: 10.1007/s10826-018-1242-x

Lareau, A. (2002). Invisible inequality: Social class and childrearing in black families and white families. *Am. Sociol. Rev.* 67, 747–776.

Lawson, G. M., and Farah, M. J. (2017). Executive function as a mediator between SER and academic achievement throughout childhood. *Int. J. Behav. Dev.* 41, 94–104. doi: 10.1177/0165025415603489

Maghra, H. G., Galadima, A., and Rufus, M. O. (2019). Effects of socioeconomic status of parents on the academic performance of senior secondary school students in agricultural science in Taraba State, Nigeria. *World J. Innov. Res. (WJIR)* 7, 05–10.

Mahmutovic, J. (2021). *What is a cross-sectional survey and why should I use one?*. Available online at: https://www.surveylegend.com/survey-examples/cross-sectional-surveys (accesssed March 18, 2022).

Miftahu, I., and Melaiye, R. O. (2021). Influence of parental socioeconomic status on student's academic performance among secondary schools in Taraba State, Nigeria. *Jalingo J. Soc. Manage. Sci.* 3, 239–255.

Mwariri, M. W., Marete, J. B., and Mwenda, E. G. (2017). Impact of parents' social economic status on students' academic performance in public secondary schools in Kieni East Subcounty, Nyericounty, kenya. *IJRDO-J. Educ. Res.* 2, 38–60.

Nja, C. O., and Sampson, S. R. (2020). Teaching practice experience and teaching effectiveness of science students- teachers. *Interdiscip. J. Sci. Educ.* 2, 125–133.

Nja, C. O., Uwe, U. E., and Ekpen, M. O. (2021). Influence of scientific attitudes on students' academic performance in chemistry in Bekwarra local government area of Cross River State. *Prestige J. Educ.* 4, 229–238.

Olayanju, O. J. (2014). Historical analysis of the implementation of teacher education policy in Nigeria: 1896- 2013. *Int. J. Spec. Gen. Educ.* 3, 131–154.

Olusegun, O. (2010). Education and national development. *Paper presented at the 2012 graduation ceremony* (Nsukka: University of Nigeria).

Onwukwe, E., Anyanwu, E., and Agommuoh, P. (2017). Influence of parents' socio-economic status on academic achievements of secondary school students in Owerri Imo State, Nigeria. *IOSR J. Res. Method Educ. (IOSR-JRME)* 7, 42–47.

Organization for Economic Cooperation and Development [OECD] (2017). *Scaling procedures and construct validation of context questionnaire data.Ch. 16*. Paris: OECD Publishing.

Osei-Owusu, B., Ampofo, E. T., Akyina, K. O., Ampomah, R., and Osei-Owusu, E. (2018). Socioeconomic status of parents and its effects on the academic performance of students of Yamfo Anglican senior high school in the BrongAhafo region of Ghana. *Adv. Soc. Sci. Res. J.* 5, 48–45.

Osonwa, O. K., Adejobi, A. O., Iyam, M. A., and Osonwa, R. H. (2013). Economic rank of parents: A Determinant on academic performance of senior secondary schools students in Ibadan Nigeria. *J. Educ. Soc. Res.* 3:115.

Ovansa, J. U. (2017). Effect of socio-economic status on the academic performance of senior secondary school students (A case study of public senior secondary schools in Adavi L.G.A of Kogi State). *Int. J. Educ. Eval.* 3, 7–17.

Pandit, R. (2011). *Problem-solving style questionnaire.* Available online at: https://www.scribd.com/doc/67905634/Problem-Solving-Style-Questionnaire Legend (accessed May 1, 2022).

Pant, K. R. (2020). Influences of parental socio-economic status on academic achievement: A case study of rural communities in Kailali, Nepal. *Contemp. Res. Interdiscip. Acad. J.* 4, 95–109.

Pinquart, M. (2016). Associations of parenting styles and dimensions with academic achievement in children and adolescents: A meta-analysis. *Educ. Psychol. Rev.* 28, 475–493.

Pinquart, M. (2017). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: An updated meta-analysis. *Dev. Psychol.* 53, 873–932. doi: 10.1037/dev0000295

Poulain, T., Mandy, V., Carolin, S., Anja, H., Antje, K., and Wieland, K. (2019). Associations between socioeconomic rank and child health: Findings of a large German cohort study. *Int. J. Environ. Res. Public Health* 16:677.

Pungello, E. P., Iruka, I. U., Dotterer, A. M., Mills-Koonce, R., and Reznick, J. S. (2009). The effects of socioeconomic status, race, and parenting on language development in early childhood. *Dev. Psychol.* 45:544.

Qasem, L. (2018). Active parental involvement relationship with socioeconomic status and impact on students' academic performance. *Int. J. Pedagogy Soc. Stud.* 3, 123–132.

Reay, D. (2019). *Poverty and education.* Leeds: Leeds Trinity University.

Repetti, R. L., Taylor, S. E., and Seeman, T. E. (2002). Risky families: Family social environments and the mental and physical health of offspring. *Psychol. Bull.* 128, 330–366.

Reynolds, A. J., and Temple, J. A. (1998). Extended early childhood intervention and school achievement: Age thirteen findings from the Chicago longitudinal study. *Child Dev.* 69, 231–246.

Reynolds, J., Houlston, C., Coleman, L., and Harold, G. (2014). *Parental conflict: Outcomes and interventions for children and families (Understanding and Strengthening Relationships), 1st ed*. Chicago: University of Chicago Press.

Roubinov, D. S., and Boyce, W. T. (2017). Parenting and SER: Relative values or enduring principles? *Curr. Opin. Psychol.* 15, 162–167.

Sindiani, A. M., Obeidat, N., Alshdaifat, E., Elsalem, L., Alwani, M. M., Rawashdeh, H., et al. (2020). Distance education during the COVID-19 outbreak: A cross-sectional study among medical students in North of Jordan. *Ann. Med. Surg. (Lond)* 59, 186–194. doi: 10.1016/j.amsu.2020.09.036

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Rev. Educ. Res.* 75, 417–453.

Torr, B. M. (2011). The changing relationship between education and marriage in the United States, 1940–2000. *J. Fam. Hist.* 36, 483–503.

Usman, M. T., Mukhtar, U., and Auwa, S. (2016). Parents socio-economic rank and students academic performance in Nigeria educational system. *Int. J. Educ. Inform. Technol.* 2, 14–18.

Wadsworth, M., and Raviv, T. (2008). An indirect effects model of the association between poverty and child functioning: The role of children's poverty-related stress. *J. Loss Trauma* 13, 156–185. doi: 10/1532502070174 2185

Wen, M., and Lin, D. (2012). Child development in rural China: Children left behind by their migrant parents and children of non-migrant families. *Child Dev.* 83, 120–136. doi: 10.1016/j.socscimed.2014.07.051

Wen, M. (2017). Social capital and adolescent substance use: The role of family, school, and neighborhood contexts. *J. Res. Adolesc.* 27, 362–378.

Wilkinson, I. A. G., Hattie, J. A., Parr, J. M., Michael, A. R., Townsend, I. F., Charlotte, U., et al. (2000). *Influence of peer effects on learning outcomes: A review of the literature; Wellington: Ministry of education.* Available online at: https://eric.ed.gov/?id=ED478708 (accessed Oct 20, 2020).

Zhou, Q., Jie, Z., and Hennessy, D. A. (2019). The role of family absolute and relative income in suicide among Chinese rural young adults: Mediation effects of social support and coping strain. *J. Public Health* 41, 609–617. doi: 10.1093/pubmed/fdy123

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 29 September 2022
DOI 10.3389/fpsyg.2022.846466

OPEN ACCESS

EDITED BY
Ana E. Azpilicueta,
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

REVIEWED BY
Sergio Longobardi,
University of Naples Parthenope, Italy
Fabio Alivernini,
Sapienza University of Rome, Italy
Sara Manganelli,
Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione (INVALSI), Italy

*CORRESPONDENCE
Guangbao Fang
fgb20071507@163.com

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

RECEIVED 31 December 2021
ACCEPTED 11 July 2022
PUBLISHED 29 September 2022

CITATION
Jin S, Fang G, Cheung KC and Sit PS (2022) Factors associated with academic resilience in disadvantaged students: An analysis based on the PISA 2015 B-S-J-G (China) sample. *Front. Psychol.* 13:846466. doi: 10.3389/fpsyg.2022.846466

Factors associated with academic resilience in disadvantaged students: An analysis based on the PISA 2015 B-S-J-G (China) sample

Songli Jin[1], Guangbao Fang[2]*, Kwok Cheung Cheung[3] and Pou Seong Sit[3]

[1]Institute of Educational Innovation, Chongqing University of Arts and Sciences, Chongqing, China, [2]School of Education, Fujian Normal University, Fuzhou, China, [3]Faculty of Education, University of Macau, Macau, Macau SAR, China

Academic resilience is evident in students who are living in vulnerable environments, yet achieve success in academic outcomes. As a result, substantial attention has been devoted to identifying the factors associated with academic resilience and supporting students to be resilient. This study used the Classification and Regression Tree and Multilevel Logistic Regression modeling to identify the potential factors related to students' academic resilience. Using these tools, the study analyzed the B-S-J-G (China) sample in PISA 2015. The variables that significantly predicted whether a student is disadvantaged and resilient (DRS) or not resilient (DNRS) were shown to be: *Proportion of teachers in school with master's degrees, Proportion of teachers in school with bachelor's degrees, Environmental awareness, Science learning time per week, Number of learning domains with additional instruction, and Students' expected occupational status*. These findings may enlighten governments, teachers, and parents on ways to assist students to be resilient.

KEYWORDS

academic resilience, classification and regression tree, disadvantaged students, program for International Student Assessment, scientific literacy

Introduction

Socioeconomic status (SES) is highly associated with students' academic achievement (White, 1982; Sirin, 2005; Wang, 2009; Ren and Xin, 2013), suggesting that students with higher SES are more likely to outperform their classmates. However, some students from low-SES households attain high levels of academic success. These

children are designated as Disadvantaged Resilient Students (DRS) because they are able to overcome the negative effects of their adverse circumstances and achieve educational success beyond the predicted SES-based outcomes (Cheung, 2017). In addition, there is a subset of pupils known as Disadvantaged Non-resilient Students (DNRS) who are from households with low socioeconomic status and have low academic achievement. Moreover, as digital natives (Prensky, 2001) who grew up with technology, the millennial generation has been compelled to increase its scientific literacy in order to adapt to the current society. In recent decades, this demand for enhancing pupils' scientific literacy has received considerable attention (Chang, 2015). How teachers, parents, and educational policymakers can assist children from low socioeconomic backgrounds to overcome their adverse situations and develop resilience in science learning is a crucial challenge for educators. The first step in answering this question is to identify the potential factors that are strongly related to students' resilience in scientific literacy performance.

PISA 2015 provides an opportunity to address this problem in the domain of scientific literacy. PISA was developed by the Organization for Economic Co-operation and Development (OECD) to assess 15-year-old students' literacy in the fields of science, mathematics and reading needed for full participation in modern societies. Assessments occur every 3 years in many regions of the world. 72 countries and economies participated in the 2015 PISA. The main domain of PISA 2015 was science, thereby providing a comprehensive measure of student performance in this domain.

Literature review

Conceptual framework

Walberg (1981) proposed the education productivity theory, which asserts that students' learning is inextricably linked to their social settings. The social context was further defined in a series of studies (Walberg, 1984), as nine elements classified into three groups. The first is about student aptitude, which encompasses ability, development, and motivation. The second category is concerned with instruction and is comprised of two components: instructional quality and quantity. Finally, there is the category of environment, which includes the home, classroom, peer group, and mass media-environments (Walberg, 1986; Fraser et al., 1987). Furthermore, the education productivity theory considered that the influences of all these elements on students' learning should be studied holistically, rather than individually, because their effects are more apparent when combined (Chen et al., 2021). Therefore, consistent with the education productivity theory, this study investigated factors associated with students, family, and schools overall, as well as prospective factors associated with students' academic resilience and scientific literacy performance.

Scientific literacy was defined in the PISA 2015 as the capacity to engage as a reflective citizen in issues and concepts related to science, and included three specific competencies involving being able to: scientifically explain phenomena, evaluate and design scientific enquiry, and scientifically interpret data and evidence (Organization for Economic Cooperation and Development [OECD], 2016a). PISA 2015 provides a framework for identifying probable factors affecting pupils' scientific literacy. According to the PISA 2015 framework for assessing scientific literacy, students' scientific literacy is related to three types of knowledge (content, procedural, and epistemic knowledge), students' attitudes toward science (such as students' interest in science and environmental awareness), and context variables (personal, local/national, and global contexts) (Organization for Economic Cooperation and Development [OECD], 2016a). Accordingly, PISA 2015 took a holistic and comprehensive approach to explaining how students' scientific literacy developed. Based on scrutiny of the educational productivity theory and the PISA 2015 framework on scientific literacy, this study concluded that these two frameworks emphasized the importance of personal and local elements, which are students-related, parent-related, teacher-related, and school-related factors in the educational context. Given the overlap of explanatory factors associated with students' learning (in their scientific literacy performance), this study created a new framework (see **Figure 1**) by combining the education productivity theory and the PISA 2015 scientific literacy assessment framework to guide the research design and selection of variables that could affect the performance of PISA index of economic, social and cultural status (ESCS) disadvantaged students in science literacy.

Factors associated with students' academic resilience in scientific literacy

Academic resilience is regarded as a characteristic of the students with low social-economic status who have achieved outstanding academic outcomes (Organization for Economic Cooperation and Development [OECD], 2011). Although PISA alternates its main domain between reading, mathematics, and science every 3 years, the discriminating factors affecting students' academic resilience may be similar across these three domains, as some students who exhibit resilience in one domain are likely to demonstrate resilience in others (Organization for Economic Cooperation and Development [OECD], 2011). Moreover, the majority of existing research ignores the specific domains of students' resilience when examining the elements that contribute to students' academic resilience. These two findings may indicate that the research

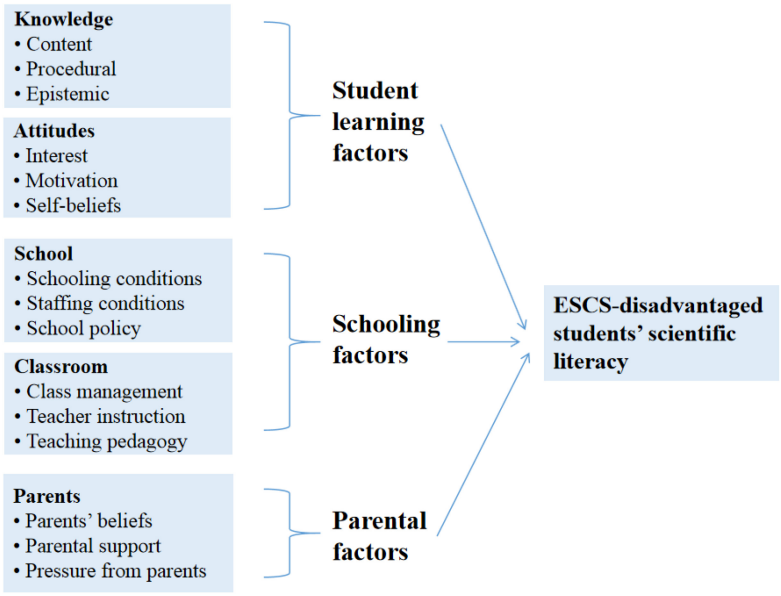


FIGURE 1
Conceptual framework of factors associated with the scientific literacy performance of economic, social and cultural status (ESCS)-disadvantaged students.

in other domains may have relevance for our work. Therefore, this study included research in the domains of mathematics and reading to indicate the possible factors associated with students' academic resilience.

On the students' level, their enjoyment of learning a subject, metacognitive awareness of learning strategies, and participation in a variety of learning activities all correlated positively with students' academic resilience, such as in reading (Shen, 2012; Cheung et al., 2014), and in mathematics (Alivernini et al., 2016; Cheung, 2017). Additionally, Clavel et al. (2021) found a strong association between students' enjoyment and interest in science and their academic resilience in science. Furthermore, the analysis of PISA 2015 revealed that students' epistemic beliefs about science, learning time, and science self-efficacy are all positively associated with students' resilience in science (Alivernini and Manganelli, 2015; She et al., 2019). Agasisti et al. (2016) indicated that students in a class whose peers have higher academic achievements are more likely to be resilient. Similarly, Cordero and Mateos-Romero (2021) suggested that students' learning skills prior to entering school, as well as their primary school classmates' socioeconomic status, are strongly associated with their academic resilience, based on an analysis of TIMSS (2015) and PIRLS (2016) data. In addition, Agasisti et al. (2021) indicated that students who attend schools with a supportive disciplinary climate, and receive additional time for instruction in critical areas are more likely to develop resilient capabilities. Agasisti and Longobardi (2017) argued in another study that if schools could provide more extracurricular activities for students, they would be more resilient. A comparable study discovered that for the African American women they studied, experiences outside of school were more critical than experiences within their schools for building up resilience (Ferguson and Martin-Dunlop, 2021). Chirkina et al. (2020) found that students' attitudes toward mathematics, their general test scores, and the average school social economic status and school type, are significantly correlated with their academic resilience. Alivernini and Manganelli (2015) asserted that teachers' salaries, parental pressure on schools, and school size are all associated with students' resilience in science.

Factors associated with students' science performance

This study categorizes the factors associated with students' scientific literacy performance into three areas, namely student-related, school-related, and family-related components. On the student level, the enjoyment of science learning was identified as the strongest factor in students' scientific performance (Altun and Kalkan, 2021; Lau and Ho, 2022). Kalkan et al. (2020) revealed that male students' scientific performance was much higher than female students in Turkey, Singapore, the United States, Italy, and Brazil. This finding is consistent with the results of various other studies, which showed that male students outperform female students (Sun et al., 2012; Lam and Lau, 2014; Chi et al., 2018). Moreover, Lau and Ho (2022) argued that, compared with male students, female students have

lower science performance and less positive attitudes toward science with an international sample. In addition, studies have also reported that students' lack of motivation when learning science increases their possibility of low achievement in science (Glynn et al., 2007; Areepattamannil et al., 2011). Chen et al. (2021) also stated that students' science self-efficacy is a positive significant factor in students' science performance. This finding is also echoed in the research of Alatli (2020).

On the teacher level, it has been demonstrated that teacher shortage is negatively connected with students' scientific literacy performance in many countries, such as in Brazil (Kalkan et al., 2020), in Turkey and Singapore (Alatli, 2020), and in Finland (Nissinen et al., 2018). By studying PISA 2015 data, Ilgaz et al. (2019) extended this negative relation to 70 countries. Furthermore, teachers' teaching methods have a strong correlation with students' science performance. For example, Lau and Ho (2022) revealed that teachers' teaching practices, direct teaching, and adapted instruction are positively associated with students' enjoyment of scientific learning and performance (Alatli, 2020; Chen et al., 2021). In addition, it has been contended that teachers' experiences and engagement in professional development activities are positively associated with students' science performance (Wenglinsky, 2002; Blank and De Las Alas, 2009). By contrast, You et al. (2020) argued that teachers' teaching experience, and engagement in professional development have no discernible relationships with students' scientific literacy.

Regarding the school level, You et al. (2020) stated that school-level factors might account for 21% of the variance on students' scientific literacy. Previous research revealed a variety of factors related to students' scientific literacy performance, including the school disciplinary climate (Altun and Kalkan, 2021), school leadership, and instructional resources (Areepattamannil et al., 2015; Topcu et al., 2015; Chi et al., 2018; Chen and Cui, 2019; Chen et al., 2021).

The relationship between school resources and students' achievements in science is still inconclusive. On one hand, some studies have indicated that school resources, such as those devoted to enhancing classroom conditions or teacher quality, show no substantial association with students' science performance (Hanushek, 1996), or any direct positive effect on their achievements (Hanushek, 1997; Picus et al., 2005). On the other hand, a review of the research insisted that school resources positively related to students' achievements in various subjects (Greenwald et al., 1996). In terms of the school's mean socioeconomic status, research has indicated that this correlates positively with students' achievements, including their science performance (Perry and McConney, 2010; You and Delgado, 2015).

On the basis of the literature reviewed above, this study concludes that few studies have incorporated all of these viewpoints; nor have they found which variables have a greater impact on students' science literacy.

Research questions

Therefore, this study used the sample of B-S-J-G (China) to address these gaps, the research questions is: What are the potential variables underlying the distinctions between the DRS and DNRS in the sample of B-S-J-G (China) in PISA 2015?

Materials and methods

Sample

Using the B-S-J-G (China) data from the PISA 2015 database, which is publicly accessible *via* the official website of the OECD, this study seeks to address the research objective outlined above. There were 9841 students from B-S-J-G (China) in the initial sample, including 4682 females and 5159 males. This study's sample consisted of 2,450 DRS and DNRS aged 15 from B-S-J-G (China), with 1168 female and 1282 male disadvantaged pupils. These students are positioned in the lowest quarter of the socioeconomic status distribution, as defined by the PISA index of ESCS; they are referred to as home-disadvantaged pupils in the current study. Technically, the DRS and DNRS are identified through three phases that are consistent with the method used in the PISA report to identify resilient students (Organization for Economic Cooperation and Development [OECD], 2016b).

The first stage is to identify disadvantaged students whose ESCS falls inside the bottom quarter of the B-S-J-G (China) ESCS distribution. Among the 9841 students of B-S-J-G (China), 2450 have been identified as disadvantaged students.

Next, the observed student scores on scientific literacy were regressed on the student ESCS across all participating countries/economies to establish the international performance-ESCS regression line. This regression line calculates anticipated student scores on the scientific literacy test. It is noteworthy that this study selected the first plausible value from the ten plausible values of science literacy in the PISA database as the observed scientific literacy score, because using one plausible value or all plausible values does not make a significant difference on large sample sizes (Organization for Economic Cooperation and Development [OECD], 2009).

Finally, the residual scores of pupils are calculated by subtracting their observed scientific literacy scores from their expected scientific literacy ratings. If students' residual scores exceed the international top quarter residual, they will be classed as DRS, and if their residual scores fall below the international top quarter residual, they will be categorized as DNRS (see **Figure 2**). According to this classification, there are 1186

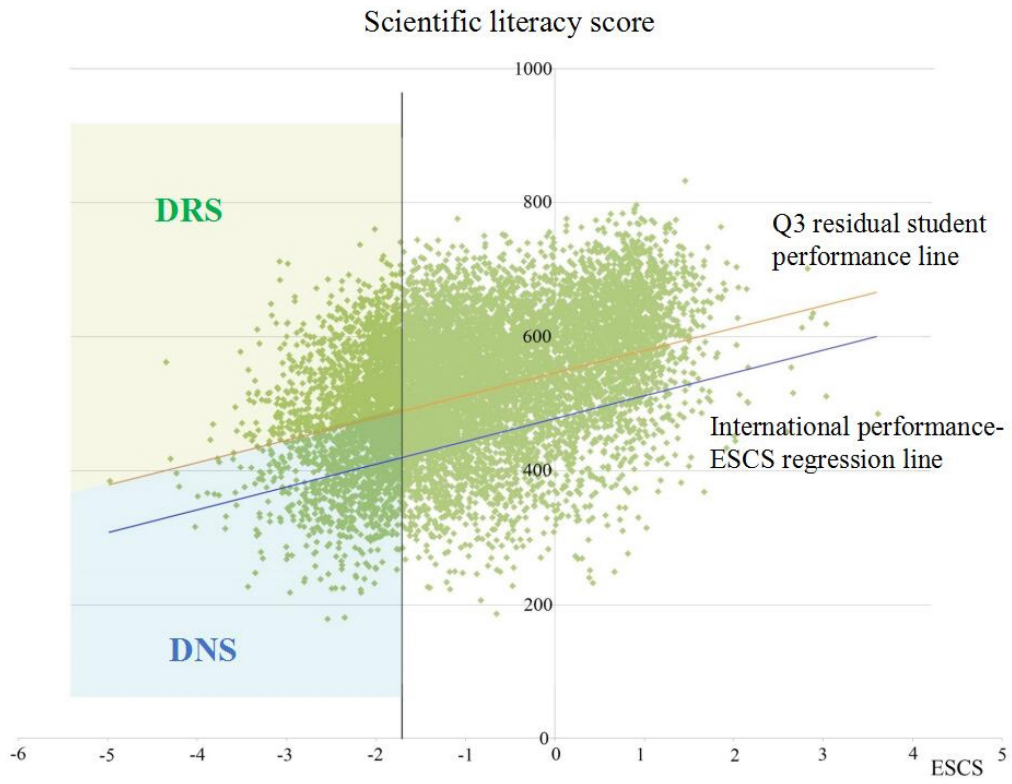


FIGURE 2
Identification of disadvantaged and resilient (DRS) and disadvantaged non-resilient students (DNRS) in the PISA 2015 B-S-J-G (China) sample.

(48.4%) DRS and 1264 (51.6%) DNRS in the B-S-J-G (China) sample for further research in this study. In particular, the DRS should meet the two following criteria: (1) Their ESCS is in the lowest quartile of B-S-J-G (China) ESCS's distribution, and (2) Their residual performance exceeds the international top quartile residual performance.

Variables

The dependent variable is the classification of resilient and non-resilient ESCS-disadvantaged students. According to the PISA, academically resilient children are those who come from families with a low socioeconomic position but yet obtain better results than expected (Organization for Economic Cooperation and Development [OECD], 2016b).

PISA collects a multitude of variables. In accordance with the conceptual model (**Figure 1**). **Appendix 1** presents the independent variables evaluated in this study. These variables are obtained from the PISA 2015 tests as well as student, school, and parent questionnaires.

The modular structure of the PISA 2015 questionnaires contains two rows of topics separated into two portions: science-related topics and general topics. Science-related topics consist of the learning environments at the school level that explicitly promote science education, such as laboratories, science-related education curriculum, collaboration among science professionals, and the values ascribed to science by the school community. The modular structure of PISA 2015 summarizes student background characteristics and science learning processes, respectively. In particular, the student background variables are associated with family and family members' education, whereas the processes are associated with three themes for in-depth examination (i.e., teaching and learning, school policies, and governance). The modular structure of PISA 2015 also discusses the non-cognitive outcomes of education (e.g., motivation, interest, beliefs, and career aspirations). The present study selected factors in accordance with the conceptual model (**Figure 1**) and the modular structure of the PISA 2015 questionnaires.

Data analysis methodologies

This study utilized the Classification and Regression Trees (CART) and Multilevel Logistic Regression (MLR) to analyze data. CART provides various benefits over other classification and regression techniques. First, it can analyze tens of thousands

of nominal, ordinal, and continuous independent variables with varying degrees of measurement. In addition, no assumptions are made about the distribution of the independent variables. Second, multicollinearity between independent variables has no effect on CART. It is a data mining method that evaluates a vast variety of predictor variables and is unaffected by the multitude of complex interactions between them. Therefore, some studies involving a large number of predictor variables have applied CART for data analysis of international educational data sets. For example, Alivernini (2013) used CART to identify variables (at country, school and student levels) associated with the differences of highest and lowest ability readers based on the Progress in International Reading Literacy Study (PIRLS) 2006 data. Sanzana et al. (2015) study identified groups of eighth-grade elementary students according to their performance in the mathematics test, using features related to individual and family behavior through random forest (RF) and CART with a database provided by the Education Quality Measurement System of Chile. Moreover, Liu and Ruiz (2008) predicted K-12 students' competence levels on test items related to energy using data mining algorithms similar to CART, based on data sets of the Third International Mathematics and Science Study (TIMSS; 1995, 1999, and 2003) and the National Assessment of Educational Progress (NAEP). Third, the advanced algorithms used in CART can effectively manage missing data. Fourth, the results are reasonably straightforward to interpret (Breiman et al., 1998; Allore et al., 2005; Strobl et al., 2009).

However, the main problem with CART is that the sample in a subset of the analysis are the students in this subset instead of the whole sample. If the study needs to check the factors associated with the differences between the DRS and the DNRS in relation to the whole sample, MLR would be an appropriate choice. MLR is well-suited for describing and testing hypotheses about relationships between a categorical outcome variable and several predictor variables with the whole sample (Peng et al., 2002). Therefore, this study employs MLR to expand upon the findings of CART analysis. In the present study, WesVar 5.1 software was employed to conduct MLR using replicated weights and complex design weights for an unbiased estimation of the parameters (WESTAT, 2007).

This study employed the Gini index as the statistical criterion for terminating successive CART iterations and was conducted with SPSS 26.0. Based on the idea of producing the most homogeneous groupings, CART automatically selects the most influential partitioning variable from the independent variables. The target sample (the parent node) is separated into two homogenous subgroups (the child nodes) depending on a particular independent variable. Afterward, each of the child nodes are separated into two subgroups using the same technique. This procedure is repeated until the impurity reduction satisfies a predefined criterion (Gini index.001) or the number of students in a subgroup falls below a predetermined threshold, which in this study was set at 50 (Strobl et al., 2009). In this study, a maximally homogenous node includes students who are either DRS or DNRS. Cross-validation procedures are then utilized to confirm the results.

Results

Factors related with the classification of academically resilient students

Figure 3 demonstrates the classification tree generated using CART. It consists of seven terminal nodes (subgroups): node 3, node 6, node 8, node 9, node 10, node 11, and node 12. Students inside nodes 6, 8, 10, and 11 are expected to be DRS, whereas students within nodes 3, node 9, and 12 are predicted to be DNRS. The accuracy of the model is estimated using 10-fold cross-validation, which is superior to other cross-validation techniques with fewer iterations (Breiman et al., 1998).

The most important variable among the 81 variables is the *Proportion of teachers in school with master's degrees*. This split improves the Gini index by 0.039. This means the impurity of the target sample (Node 0) is reduced by 0.039. That is to say, the subgroups (Node 1 and Node 2) become more homogeneous. If a school has more than.8% of teachers with a master's degree qualification (at the 55th percentile of the classifying variable), the probability of ESCS-disadvantaged students at the school becoming resilient rises from 48.4% to 63.9% (node 2). However, if a school has equal to or less than 0.8% of teachers with a master's degree qualification, there is a possibility that the level of DRS students will go down from 48.4% to 35.8% (node 1).

The classification tree is interpreted from the right to the left. The next variable is *Science learning time per week*. This split improves the Gini index further by 0.018 based on 0.039. It means the impurity of the parent node is reduced by 0.057, and the subgroups (Node 5 and Node 6) become more homogeneous. For students who study at a school with more than 0.8% of teachers holding a master's degree (on the right-hand side of the tree), and if they learn science for more than 232.5 min per week (at the 35th percentile), the percentage of DRS students rises from 63.9% to 73.7% (Node 6). Also, if students learn science for equal to or less than 232.5 min per week, the percentage of DRS students goes down from 63.9% to 45.5% (node 5).

Furthermore, when students learn science for equal to or less than 232.5 min per week, the variable that determines whether they are DRS or DNRS relates to the *Proportion of teachers in school with bachelor's degrees*. This split improves the Gini index further by 0.005 based on 0.057, that is 0.062, and the subgroups (Node 9 and Node 10) become more homogeneous. If a school has more than 93.4% of teachers with a bachelor's degree (at the 77th percentile), the percentage of DRS students in this school rises from 45.5% to 67.7% (Node 10). However, if this school

10.3389/fpsyg.2022.846466

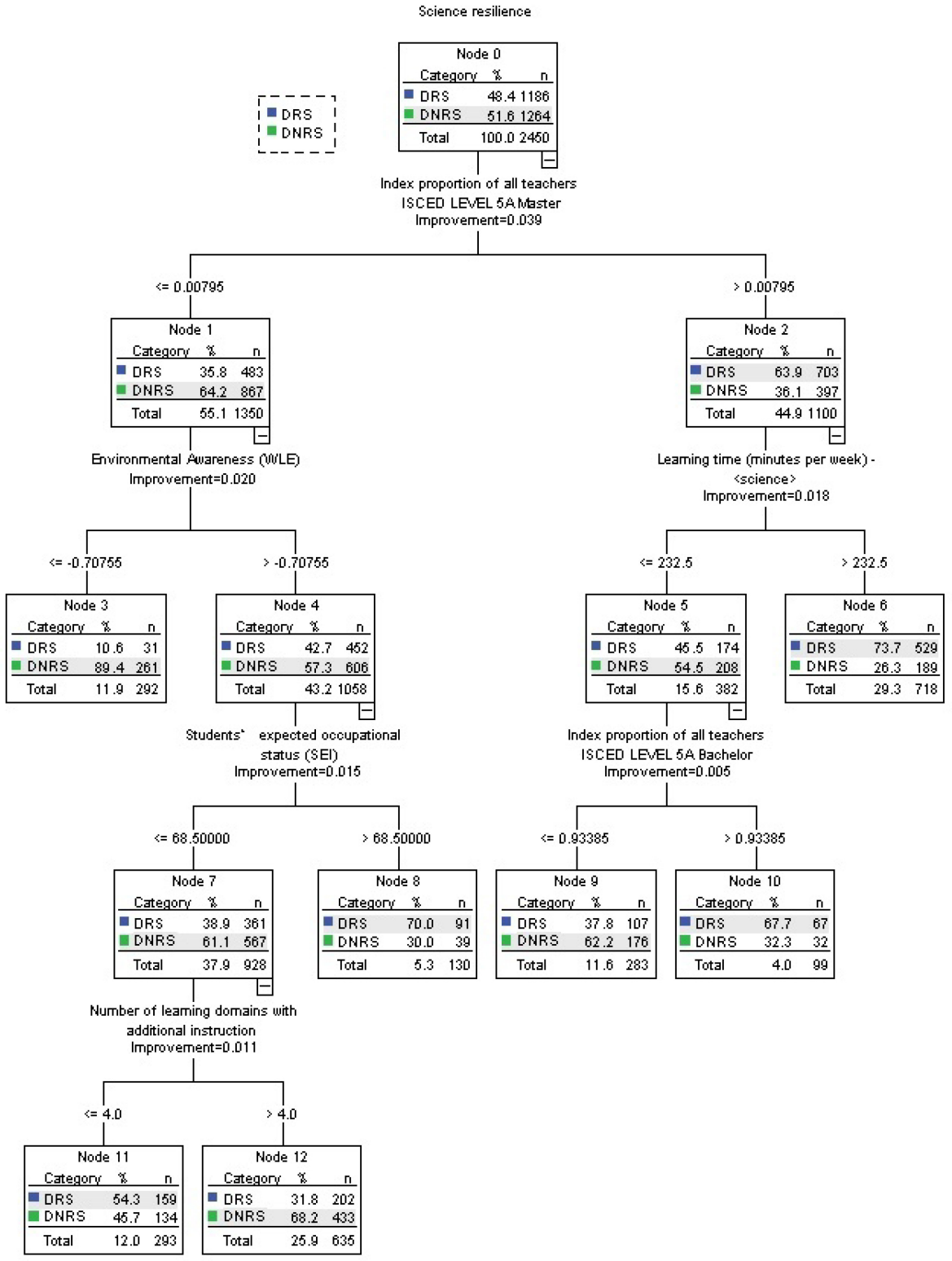


FIGURE 3
Results of the Classification and Regression Trees (CART) analysis for the B-S-J-G (China) data.

has equal to or less than 93.4%, the percentage of DRS students in this school goes down slightly from 45.5% to 37.8% (node 9).

On the right branch of the tree relating to students who study in a school with equal to or less than.8% of teachers with a master's degree (in Node 1). This split improves the Gini index further by 0.020 based on 0.039. This means the impurity of the target sample (Node 0) is reduced by 0.059, and the subgroups (Node 3 and Node 4) become more homogeneous. If a student in Node 1 has a high *Environmental awareness* above −0.708 (at the 20th percentile), the probability of him/her being a DRS increases from 35.8% to 42.7% (Node 4); but if this is equal to or lower than −0.708, the probability decreases remarkably from 35.8% to 10.6% (Node 3). Students in Nodes 3 and 4 will be predicted to be DNRS.

However, the percentage of DRS in Node 4 can be improved by the variable of *Students' expected occupational status*. This split improves the Gini index further by 0.015 based on 0.059, that is 0.074, and the subgroups (Node 7 and Node 8) become more homogeneous. If a student in Node 4 has a high expected occupational status above 68.5 (at the 80th percentile), the probability of him/her rising to a DRS increases dramatically from 42.7% to 70.0% (Node 8); but if it is equal to or lower than 68.5, the probability of him/her becoming a DRS decreases a little from 42.7% to 38.9% (Node 7).

Finally, the variable relating to the *Number of learning domains with additional instruction* affects whether a student is considered as DRS or DNRS in Node 7. This split improves the Gini index further by 0.011 based on 0.074. If the student in Node 7 learns equal to or less than four learning domains of additional instruction (at the 40th percentile), the probability of him/her becoming a DRS increases from 38.9% to 54.3%. However, if the student receives more than four learning domains of additional instruction, the probability of him/her becoming a DRS decreases slightly to 31.8%.

Factors associated with the differences of disadvantaged and resilient and disadvantaged non-resilient students

This study employed MLR to duplicate and expand upon the findings of CART analysis (Alivernini and Manganelli, 2015). The samples in the child nodes of CART analysis consist only of the students in its parent nodes, not the entire sample (for example, the sample of the analysis to divide Node1 into Node 3 and Node 4 is just 1350 students in Node 1, rather than the entire sample of 2450), so it is necessary to use MLR to test the factor associated with the differences between the DRS and the DNRS in the entire sample.

The dependent variable in MLR is dichotomous (DRS vs. DNRS), and the independent variables are those differentiating variables found in CART, namely, *Proportion of teachers in school with master's degrees, Proportion of teachers in school with bachelor's degrees, Environmental awareness, Science learning time per week, Students' expected occupational status,* and *Number of learning domains with additional instruction*. In the MLR analysis, the average of the school ESCS serves as a control variable. Both the student and school levels employ weights. This study employed the final student weight variable (W_FSTUWT) from the PISA raw data to weight the student level variables and the total of W_FSTUWT within each school as between-school weights for school level analysis in MLR (Organization for Economic Cooperation and Development [OECD], 2016b). The outcome is shown in **Table 1**.

The results indicate that the six essential factors show adequate effect in predicting whether a student is a DRS or a DNRS, with Negative log-likelihood, Cox-Snell, and Estrella indices ranging from 0.209 to 0.277 (WESTAT, 2007). According to **Table 1**, teacher qualifications are of utmost importance in B-S-J-G (China), since an increase of one standard deviation in the *Proportion of teachers in school with master's degrees* results in a 9455-fold increase in the likelihood that a pupil is academically resilient. There is a 4-fold rise for teachers with bachelor's degrees. An increase of one standard deviation in *Environmental awareness* improves the likelihood that an ESCS-disadvantaged youngster will be academically resilient by 56.9 percent. An increase of one standard deviation in the *Number of learning domains with additional instruction* reduces the likelihood that an ESCS-disadvantaged student will be academically resilient by 16.6 percent. For a one-unit increase in the standard deviation of the indices, the influence of the *Science*

TABLE 1 Results of Multilevel Logistic Regression (MLR) of the disadvantaged and resilient (DRS) vs. disadvantaged non-resilient students (DNRS) classification.

	Variable	*B*	S.E.	Exp (*B*)
	Intercept	1.104	1.594	
School level	Proportion of teachers in school with master's degrees	9.154*	3.818	9454.813
	Proportion of teachers in school with bachelor's degrees	1.381**	0.427	3.978
	School ESCS (control variable)	1.596*	0.644	4.934
Student level	Environmental awareness	0.451**	0.102	1.569
	Science learning time per week	0.002**	0.000	1.002
	Students' expected occupational status	0.028**	0.005	1.029
	Number of learning domains with additional instruction	−0.182**	0.022	0.834

(1) ** <0.01, * <0.05. (2) Negative log-likelihood = 0.209; Likelihood ratio (Cox-Snell) = 0.251; Likelihood ratio (Estrella) = 0.277. (3) Variable is at school level or at student level is according to PISA 2015 technical report (Organization for Economic Cooperation and Development [OECD], 2017).

learning time per week and the *Students' expected occupational status* is 0.2% and 2.9%, respectively.

To summarize, the six variables found to significantly predict whether a student is a DRS or a DNRS are: *Proportion of teachers in school with master's degrees, Proportion of teachers in school with bachelor's degrees, Environmental awareness, Science learning time per week, Number of learning domains with additional instruction, and Students' expected occupational status. Environmental awareness, Science learning time per week, and Students' expected occupational status* are literacy learning factors. *Proportion of teachers in school with master's degrees* and *Proportion of teachers in school with bachelor's degrees* are schooling factors; and, *Number of learning domains with additional instruction* is considered in this present study as both learning and parental factors.

Discussion and conclusion

Discussion

This study has sought to elucidate the various characteristics in their family, personal and educational backgrounds that differentiate DRS from DNRS adolescents in scientific literacy. This study suggests that the number of teachers with master's or bachelor's degrees in schools could significantly increase the probability of academic resilience on scientific literacy of ESCS-disadvantaged children, a finding that is supported by numerous other studies (Goldhaber and Brewer, 1997, 2000; Clotfelter et al., 2007; Chu et al., 2015). In a research review, Wayne and Youngs (2003) offered a possible reason for this link, as teachers with advanced degrees are more likely to teach more effectively and offer their students more learning supports.

This study also indicates that students' environmental awareness could add their possibility of been resilience in scientific literacy. This study's findings are congruent with other studies (Bybee, 2008; Organization for Economic Cooperation and Development [OECD], 2016a). In the PISA 2015, awareness of environmental issues was one important aspect of the construct attitudes toward science (Organization for Economic Cooperation and Development [OECD], 2016a). According to Bybee (2008), the scientific literacy of students should be developed in parallel with their attitudes and beliefs regarding natural resources and the quality of the environment. Natural resources and environmental quality are two major areas in which scientific literacy has significant importance for promoting and preserving the quality of life and formulating public policy for individuals and communities (Organization for Economic Cooperation and Development [OECD], 2016a). Therefore, environmental awareness is an essential component of ESCS-disadvantaged students' scientific literacy.

This study presents that spending more than 4 h per week studying science is significantly related with the ESCS-disadvantaged students' resilience in scientific literacy, while the regression coefficient is modest. Agasisti et al. (2021) have also demonstrated that students are more likely to develop resilience when students were given additional time for instruction in essential courses. Because of the exam-based feature of high school courses in China, many Chinese high school students devote the majority of their time to study (Leung, 2021). This study reveals that 65 percent of B-S-J-G (China) ESCS underprivileged students who study science for more than 4 h per week are likely to be marginally more academically resilient in scientific literacy.

To our knowledge, there is no direct evidence to confirm the positive relationship between students' career expectations and their resilience in scientific literacy. This research filles this gap by indicating that the possibility of students be resilient in science literacy would increase when students have a higher expected occupational status. Moreover, as some research suggested, students' career expectation is positively related with their learning motivations (Domene et al., 2011), and students' science learning motivation is an essential component affecting their science accomplishment (Glynn et al., 2007; Areepattamannil et al., 2011). Therefore, students' learning motivation may mediate the relationship between students' career expectation and their academic resilience in scientific literacy, which requires to be assessed in the future study.

The number of learning domains in which ESCS-disadvantaged pupils receive additional instruction has a negative relationship with their probability of be resilient in scientific literacy. The potential reason of this relationship may be that students spent so much time on learning other subjects, such as mathematics, and English, which are critical in College Entrance Examination (高考) (Zhang, 2011; Zhang and Bray, 2018) rather than in science. For example, Zhang and Bray (2018) indicated that 58.7 percent of sampled students in Grades 3–9 in Shanghai had received various additional instruction, while 81.5 percent and 76.8 percent of those students received tutoring in Mathematics and English, respectively. Similarly, this study finds that 60 percent of ESCS-disadvantaged students in the B-S-J-G (China) sample participated in more than four learning domains requiring additional instruction. There is a possibility for these students that they may not have sufficient time to learning science.

In addition, receiving excessive amount of additional instruction may increase students' academic pressure, resulting in a decline in instructional efficacy (Šťastný et al., 2021). As indicated by Huang and Chen (2008), the link between the quantity of additional instruction and academic performance is not a positive and linear relationship; rather, it follows a "climb first and then fall" pattern. To some extent, additional instruction may improve students' academic performance, whereas excessive additional instruction definitely have

negative influences on students' learning. Thus, when ESCS-disadvantaged pupils participated in excessive additional instruction, their scientific literacy performance would decline and be non-resilient.

Implications

This exploratory study aimed to identify the influential predictors of whether a student is a DRS or DNRS in four prosperous Chinese cities/provinces. The PISA 2015 data offer a wealth of information and can be investigated in depth through educational data mining to reveal important knowledge to bolster informed policy-making. This research illustrates that it is possible to obtain information with crucial significance. Important characteristics that predict the classification of pupils as DRS or DNRS in the other PISA 2015 participating economies may or may not be identical to those discovered for B-S-J-G (China). However, the most significant features discovered may indicate to researchers and educational practitioners in various educational systems the way forward in terms of enhancing the scientific literacy of ESCS-disadvantaged pupils.

This study has suggested that the proportion of teachers with master's or bachelor's degrees in schools is related to the students' resilience in scientific literacy. In order to effectively improve the students' resilience in scientific literacy, it is necessary to increase the academic qualifications of teachers. Understanding this is crucial for the B-S-J-G (China) Governments and school administrators, since on the basis of this knowledge they can provide instructors with more opportunity to advance their education and achieve higher academic levels.

In addition, this study indicated that environmental awareness is a significant element related to the DRS's scientific literacy. Considering the significance of environmental challenges to the sustainability of life on Earth and the survival of humanity, the OECD has suggested that young people must learn to plan their lives in accordance with ecological principles (Organization for Economic Cooperation and Development [OECD], 2016a). Thus, fostering awareness of environmental issues and a responsible attitude toward the sustainability of the environment are essential for modern science education.

In recent years, a growing number of Chinese policymakers and academics have focused on natural resources and environmental quality (e.g., Stalley and Yang, 2006; He et al., 2011; Lumkes et al., 2012; Wu, 2013). However, schools in different districts place varying amounts of emphasis on environmental education and implement it in various ways. Consequently, children in B-S-J-G (China) acquire various types of environmental education and achieve varying academic levels in this subject at their respective schools. The B-S-J-G (China) Governments must review the implementation of environmental education in schools so that each student receives a high-quality environmental education. Students must also share responsibilities for fostering environmental awareness.

This study found that spending more than 4 h per week studying science is connected with students' resilience in scientific literacy. However, time is only one aspect of the equation; learning efficacy is the other. Throughout their studies, ESCS-disadvantaged students must guarantee that their learning is effective. Educatively relevant teacher scaffolding is required in this regard.

If a student's expected occupational status is above the 80th percentile of his or her peers in B-S-J-G (China), he or she will have a greater chance of being categorized as a DRS rather than a DNRS. The predicted occupational position of Chinese pupils may be affected by their self-evaluation, parental background, expectations, as well as social appraisal and support from instructors, peers, parents, and relatives (Hsieh, 2005; Wang, 2007). Therefore, instructors, parents, classmates, and relatives are encouraged to provide children and adolescents with accurate evaluations and to encourage them to reach a higher occupational standing.

This study indicates that excessive additional education may negatively impact students' resilience in scientific literacy performance. Due to cost constraints, parents in China determine the extent to which ESCS-disadvantaged students receive additional teaching in domain courses learned at school. According to the findings of the present study conducted in B-S-J-G (China), parents should not force their children to take on more than four learning domains at any one time. China has introduced a "double reduction" strategy in 2021 to alleviate the strain of excessive homework and off-campus tutoring for compulsory education students (Xu and Jianli, 2021).

Limitation

In this study, CART and MLR were used to identify the significant factors of the DRS to DNRS classification of teenagers in B-S-J-G (China) based on the assessment of scientific literacy. However, this cannot depict the intricate interactions between these significant components. Consequently, future research could investigate the links between these variables using structural equation modeling or other causal modeling techniques.

Conclusion

Based on the education productivity theory, this study utilized the B-S-J-G (China) data from the PISA 2015 database to investigate the relationship between student-related, school-related, and parental factors by differentiating between DRS and DNRS pupils. The CART and MLR were used to determine the

distinguishing factors. According to the findings of this study, the following variables significantly predicted whether a student is a DRS or DNRS are: *Proportion of teachers in school with master's degrees, Proportion of teachers in school with bachelor's degrees, Environmental awareness, Science learning time per week, Number of learning domains with additional instruction,* and *Students' expected occupational status*. These findings also enlighten governments, educational practitioners, and parents about ways to assist DNRS youth in attaining a greater level of scientific literacy.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: OECD-PISA2015 China dataset, https://www.oecd.org/pisa/data/2015database/.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

SJ: writes the raw manuscript. GF: revises and proofreads the manuscript. KC and PS: supervise the writing of the raw manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Humanities and Social Science Foundation of Ministry of Education of China (20XJC880003).

Acknowledgments

The author to Alivernini, Manganelli, and Lucidi's research in 2016 for giving this study great inspiration.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Agasisti, T., Avvisati, F., Borgonovi, F., and Longobardi, S. (2021). What school factors are associated with the success of socio-economically disadvantaged students? An empirical investigation using PISA data. *Soc. Indic. Res.* 157, 749–781. doi: 10.1007/s11205-021-02668-w

Agasisti, T., and Longobardi, S. (2017). Equality of educational opportunities, schools' characteristics and resilient students: an empirical study of EU-15 countries using OECD-PISA 2009 data. *Soc. Indic. Res.* 134, 917–953. doi: 10.1007/s11205-016-1464-5

Agasisti, T., Soncin, M., and Valenti, R. (2016). School factors helping disadvantaged students to succeed: empirical evidence from four Italian cities. *Policy Stud.* 37, 147–177. doi: 10.1080/01442872.2015.1127341

Alatli, B. K. (2020). Investigation of factors associated with science literacy performance of students by hierarchical linear modeling: PISA 2015 comparison of Turkey and Singapore. *Educ. Sci.* 45, 17–49.

Alivernini, F. (2013). An exploration of the gap between highest and lowest ability readers across 20 countries. *Educ. Stud.* 39, 399–417. doi: 10.1080/03055698.2013.767187

Alivernini, F., and Manganelli, S. (2015). Country, school and students factors associated with extreme levels of science literacy across 25 countries. *Int. J. Sci. Educ.* 37, 1992–2012. doi: 10.1080/09500693.2015.1060648

Alivernini, F., Manganelli, S., and Lucidi, F. (2016). The last shall be the first: competencies, equity and the power of resilience in the Italian school system. *Learn. Individ. Diff.* 51, 19–28. doi: 10.1016/j.lindif.2016.08.010

Allore, H., Tinetti, M. E., Araujo, K. L., Hardy, S., and Peduzzi, P. (2005). A case study found that a regression tree outperformed multiple linear regression in predicting the relationship between impairments and social and productive activities scores. *J. Clin. Epidemiol.* 58, 154–161. doi: 10.1016/j.jclinepi.2004.09.001

Altun, A., and Kalkan, O. K. (2021). Cross-national study on students and school factors affecting science literacy. *Educ. Stud.* 47, 403–421. doi: 10.1080/03055698.2019.1702511

Areepattamannil, S., Chiam, C. L., Lee, D. H., and Hong, H. (2015). "Correlates of science achievement in Singapore: a multilevel exploration," in *Science Education in East Asia*, ed. M. S. Khine (New York, NY: Springer), 607–629. doi: 10.1007/978-3-319-16390-1_25

Areepattamannil, S., Freeman, J. G., and Klinger, D. A. (2011). Influence of motivation, self-beliefs, and instructional practices on science achievement of adolescents in Canada. *Soc. Psychol. Educ.* 14, 233–259. doi: 10.1007/s11218-010-9144-9

Blank, R. K., and De Las Alas, N. (2009). The Effects of Teacher Professional Development on Gains in Student Achievement: How Meta Analysis Provides Scientific Evidence Useful to Education Leaders. Avalable at: https://files.eric.ed.gov/fulltext/ED514190.pdf (accessed June 12, 2020).

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1998). *Classification and Regression Trees*. Boca Raton, FL: CRC Press.

Bybee, R. W. (2008). Scientific literacy, environmental issues, and PISA 2006: the 2008 Paul F-Brandwein lecture. *J. Sci. Educ. Technol.* 17, 566–585. doi: 10.1007/s10956-008-9124-4

Chang, Y. (2015). Science motivation across Asian countries: links among future-oriented motivation, self-efficacy, task values, and achievement outcomes. *Asia Pac. Educ. Res.* 24, 247–258. doi: 10.1007/s40299-014-0176-3

Chen, F., and Cui, Y. (2019). Investigating the relation of perceived teacher unfairness to science achievement by hierarchical linear modeling in 52 countries and economies. *Educ. Psychol.* 40, 273–295. doi: 10.1080/01443410.2019.1652248

Chen, J. P., Zhang, Y., Wei, Y., and Hu, J. (2021). Discrimination of the contextual features of top performers in scientific literacy using a machine learning approach. *Res. Sci. Educ.* 51(Suppl. 1), 129–158. doi: 10.1007/s11165-019-9835-y

Cheung, K. C. (2017). The effects of resilience in learning variables on mathematical literacy performance: a study of learning characteristics of the academic resilient and advantaged low achievers in Shanghai, Singapore, Hong Kong, Taiwan and Korea. *Educ. Psychol.* 37, 965–982. doi: 10.1080/01443410.2016.1194372

Cheung, K. C., Sit, P. S., Soh, K. C., Ieong, M. K., and Mak, S. K. (2014). Predicting academic resilience with reading engagement and demographic variables: comparing Shanghai, Hong Kong, Korea, and Singapore from the PISA perspective. *Asia Pac. Educ. Res.* 23, 895–909. doi: 10.1007/s40299-013-0143-4

Chi, S., Liu, X., Wang, Z., and Won Han, S. (2018). Moderation of the effects of scientific inquiry activities on low SES students' PISA 2015 science achievement by school teacher support and disciplinary climate in science classroom across gender. *Int. J. Sci. Educ.* 40, 1284–1304. doi: 10.1080/09500693.2018.1476742

Chirkina, T., Khavenson, T., Pinskaya, M., and Zvyagintsev, R. (2020). Factors of student resilience obtained from TIMSS and PISA longitudinal studies. *Issues Educ. Res.* 30, 1245–1263.

Chu, J. H., Loyalka, P., Chu, J., Qu, Q., Shi, Y., and Li, G. (2015). The impact of teacher credentials on student achievement in China. *China Econ. Rev.* 36, 14–24. doi: 10.1016/j.chieco.2015.08.006

Clavel, J. G., Garcia Crespo, F. J., and Sanz San Miguel, L. (2021). Rising above their circumstances: what makes some disadvantaged east and south-east Asian students perform far better in science than their background predicts? *Asia Pac. J. Educ.* 1, 1–16. doi: 10.1080/02188791.2021.1886905

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2007). Teacher credentials and student achievement: longitudinal analysis with student fixed effects. *Econ. Educ. Rev.* 26, 673–682. doi: 10.1016/j.econedurev.2007.10.002

Cordero, J. M., and Mateos-Romero, L. (2021). Exploring factors related with resilience in primary education: evidence from European countries. *Stud. Educ. Eval.* 70:101045. doi: 10.1016/j.stueduc.2021.101045

Domene, J. F., Socholotiuk, K. D., and Woitowicz, L. A. (2011). Academic motivation in post-secondary students: effects of career outcome expectations and type of aspiration. *Can. J. Educ.* 34, 99–127.

Ferguson, D., and Martin-Dunlop, C. (2021). Uncovering stories of resilience among successful African American women in STEM. *Cultural Stud. Sci. Educ.* 16, 461–484. doi: 10.1007/s11422-020-10006-8

Fraser, B. J., Walberg, H. J., Welch, W. W., and Hattie, J. A. (1987). Syntheses of educational productivity research. *Int. J. Educ. Res.* 11, 147–252.

Glynn, S. M., Taasoobshirazi, G., and Brickman, P. (2007). Nonscience majors learning science: a theoretical model of motivation. *J. Res. Sci. Teach.* 44, 1088–1107. doi: 10.1002/tea.20181

Goldhaber, D. D., and Brewer, D. J. (1997). "Evaluating the effect of teacher degree level on educational performance," in *Developments in School Finance, 1996*, ed. W. J. Fowler (Washington, DC: National Center for Education Statistics), 197–210.

Goldhaber, D. D., and Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educ. Eval. Policy Anal.* 22, 129–145. doi: 10.3102/01623737022002129

Greenwald, R., Hedges, L. V., and Laine, R. D. (1996). The effect of school resources on student achievement. *Rev. Educ. Res.* 66, 361–396. doi: 10.3102/00346543066003361

Hanushek, E. A. (1996). Measuring investment in education. *J. Econ. Perspect.* 10, 9–30. doi: 10.1257/jep.10.4.9

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: an update. *Educ. Eval. Policy Anal.* 19, 141–164. doi: 10.3102/01623737019002141

He, X., Hong, T., Liu, L., and Tiefenbacher, J. (2011). A comparative study of environmental knowledge, attitudes, and behaviors among university students in China. *Int. Res. Geogr. Environ. Educ.* 20, 91–104.

Hsieh, M. F. (2005). 性別、家庭背景及職業興趣對國中資優班學生教育及職業選 [The effects of education and job selection of different gender, background and interests on gifted students in junior high school]. 擇影響之研究 *[Corpus Spec. Educ.]* 7, 1–32.

Huang, Y. J., and Chen, C. W. (2008). **國立屏東教育大學特殊教育文集**, 學科補習, [Academic Cram Schooling, Academic Performance, and Opportunity of Entering Public Universities]. 成績表現與升學結果 *[Bull. Educ. Res.]* 54, 117–149.

Ilgaz, G., Eskici, M., and Vural, L. (2019). Which country is more effective in science teaching? Evidence from PISA 2015 as a secondary school assessment tool. *Asian J. Educ. Train.* 5, 349–361. doi: 10.20448/journal.522.2019.52.349.361

Kalkan, O. K., Altun, A., and Atar, B. (2020). Role of teacher-related factors and educational resources in science literacy: an international perspective. *Stud. Educ. Eval.* 67:100935.

Lam, T. Y. P., and Lau, K. C. (2014). Examining factors affecting science achievement of Hong Kong in PISA 2006 using hierarchical linear modeling. *Int. J. Sci. Educ.* 36, 2463–2480. doi: 10.1080/09500693.2013.879223

Lau, K. C., and Ho, S. C. (2022). Attitudes towards science, teaching practices, and science performance in PISA 2015: multilevel analysis of the Chinese and western top performers. *Res. Sci. Educ.* 52, 415–426. doi: 10.1007/s11165-020-09954-6

Leung, F. K. S. (2021). Jewish culture, Chinese culture, and mathematics education. *Educ. Stud. Math.* 107, 405–423. doi: 10.1007/s10649-021-10034-3

Liu, X., and Ruiz, M. E. (2008). Using data mining to predict K-12 students' performance on large-scale assessment items related to energy. *J. Res. Sci. Teach.* 45, 554–573. doi: 10.1002/tea.20232

Lumkes, J. H., Hallett, S., and Vallade, L. (2012). Hearing versus experiencing: the impact of a short-term study abroad experience in China on students' perceptions regarding globalization and cultural awareness. *Int. J. Intercult. Relat.* 36, 151–159. doi: 10.1016/j.ijintrel.2011.12.004

Nissinen, K., Olafsson, R. G., Rautopuro, J., Halldorsson, A. M., and Vettenranta, J. (2018). *The Urban Advantage in Education? Science Achievement Differences Between Metropolitan and Other Areas in Finland and Iceland in PISA 2015*. Copenhagen: Nordic Council of Ministers.

Organization for Economic Cooperation and Development [OECD] (2009). *PISA Data Analysis Manual: SPSS*. Paris: OECD Publishing.

Organization for Economic Cooperation and Development [OECD] (2011). *Against the Odds: Disadvantaged Students Who Succeed in School*. Paris: OECD Publishing.

Organization for Economic Cooperation and Development [OECD] (2016a). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematics and Financial Literacy*. Paris: OECD Publishing.

Organization for Economic Cooperation and Development [OECD] (2016b). *PISA 2015 Results: Excellence and Equity in Education (Volume 1)*. Paris: OECD Publishing.

Organization for Economic Cooperation and Development [OECD] (2017). *PISA 2015 Technical Report*. Paris: OECD Publishing.

Peng, C. Y. J., Lee, K. L., and Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *J. Educ. Res.* 96, 3–14. doi: 10.1080/00220670209598786

Perry, L. B., and McConney, A. (2010). Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. *Teach. College Rec.* 112, 1137–1162. doi: 10.1177/016146811011200401

Picus, L. O., Marion, S. F., Calvo, N., and Glenn, W. J. (2005). Understanding the relationship between student achievement and the quality of educational facilities: evidence from Wyoming. *Peabody J. Educ.* 80, 71–95. doi: 10.1207/s15327930pje8003_5

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon* 9, 1–6. doi: 10.1108/10748120110424816

Ren, C. R., and Xin, T. (2013). **教育研究集刊** [Longitudinal study on predicting effect of social economic status on students' performance]. 家庭社会经济地位对小学生成绩预测效应的追踪研 *[Educ. Res.]* 3, 79–87.

Sanzana, M. B., Garrido, S. S., and Poblete, C. M. (2015). Profiles of Chilean students according to academic performance in mathematics: an exploratory study using classification trees and random forests. *Stud. Educ. Eval.* 44, 50–59. doi: 10.1016/j.stueduc.2015.01.002

She, H. C., Lin, H. S., and Huang, L. Y. (2019). Reflections on and implications of the Programme for International Student Assessment 2015 (PISA 2015) performance of students in Taiwan: the role of epistemic beliefs about science in scientific literacy. *J. Res. Sci. Teach.* 56, 1309–1340. doi: 10.1002/tea.21553

Shen, X. (2012). Shanghai resilient students' learning characteristics: evidence-based research with PISA 2009 data. *Exploring Educ. Dev.* 18, 25–36.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: a meta-analytic review of research. *Rev. Educ. Res.* 75, 417–453. doi: 10.3102/00346543075003417

Stalley, P., and Yang, D. (2006). An emerging environmental movement in China? *China Q.* 186, 333–356. doi: 10.1017/S030574100600018X

Šťastný, V., Greger, D., and Soukup, P. (2021). Does the quality of school instruction relate to the use of additional tutoring in science? Comparative analysis of five post-socialist countries. *Sch. Effect. Sch. Improv.* 32, 24–46. doi: 10.1080/09243453.2020.1770809

Strobl, C., Malley, J., and Tutz, G. (2009). An introduction to recursive partitioning: rationale, application, and characteristics of classi?cation and regression trees, bagging, and random forests. *Psychol. Methods* 14, 323–348. doi: 10.1037/a0016973

Sun, L., Bradley, K. D., and Akers, K. (2012). A multilevel modelling approach to investigating factors impacting science achievement for secondary school students: PISA Hong Kong sample. *Int. J. Sci. Educ.* 34, 2107–2125. doi: 10.1080/09500693.2012.708063

Topcu, M. S., Arikan, S., and Erbilgin, E. (2015). Turkish students' science performance and related factors in PISA 2006 and 2009. *Austral. Educ. Res.* 42, 117–132. doi: 10.1007/s13384-014-0157-9

Walberg, H. J. (1981). "A psychological theory of educational productivity," in *Psychology and Education*, eds F. H. Farley and N. Gordon (Berkeley, CA: McCutchan), 81–110.

Walberg, H. J. (1984). Improving the productivity of America's schools. *Educ. Leadersh.* 41, 19–27.

Walberg, H. J. (1986). "Synthesis of research on teaching," in *Handbook of Research on Teaching*, 3rd Edn, ed. M. C. Wittrock (New York, NY: Macmillan), 214–229.

Wang, C. H. (2009). 教育研究 [Study on the relationship between family background and students' academic achievement in compulsory education]. 义务教育阶段学生学业成就与家庭背景的关系研究 *[Modern Prim. Second. Educ.]* 12, 59–61.

Wang, R. (2007). *The Influence of Self-Appraisal and Social Supports on Self-Occupation Expectation Among Graduate Students*. Doctoral dissertation. Changchun: Northeast Normal University.

Wayne, A. J., and Youngs, P. (2003). Teacher characteristics and student achievement gains: a review. *Rev. Educ. Res.* 73, 89–122. doi: 10.3102/00346543073001089

Wenglinsky, H. (2002). The link between teacher classroom practices and student academic performance. *Educ. Policy Anal. Arch.* 10, 1–29. doi: 10.1016/j.jsp.2021.09.004 doi: 10.14507/epaa.v10n12.2002

WESTAT (2007). *WesVar User's Guide*. Rockville: WESTAT.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychol. Bull.* 91, 461–481. doi: 10.1037/0033-2909.91.3.461

Wu, F. (2013). Environmental activism in provincial China. *J. Environ. Policy Plann.* 15, 89–108. doi: 10.1080/1523908X.2013.763634

Xu, Z., and Jianli, F. (2021). "双减"背景下如何实现课堂教学的应教尽教 [How to realize classroom teaching under the background of "Double Reduction" policy]. 中国教育学刊 *[J. Chinese Soc. Educ.]* 12, 1–5.

You, H. S., and Delgado, C. (2015). "Revisiting the Coleman Report: exploring school effects on scientific literacy in PISA 2012 using hierarchical linear modeling," in *Paper Presented at the Annual meeting of the National Association for Research in Science Teaching (NARST)*, (Boston, MA: NARST).

You, H. S., Park, S., and Delgado, C. (2020). A closer look at US schools: what characteristics are associated with scientific literacy? A multivariate multilevel analysis using PISA 2015. *Sci. Educ.* 105, 406–437. doi: 10.1002/sce.b21609

Zhang, W., and Bray, M. (2018). Equalising schooling, unequalising private supplementary tutoring: access and tracking through shadow education in China. *Oxf. Rev. Educ.* 44, 221–238. doi: 10.1080/03054985.2017.1389710

Zhang, Y. (2011). *The Determinants of National College Entrance Exam Performance in China With an Analysis of Private Tutoring*. Doctoral dissertation. New York, NY: Columbia University.

Appendix

TABLE A1 Independent variables drawn from the PISA 2015 database.

Variable type	Classification	Variable label
Personal factors	Student background	Gender
		Grade repetition
		Duration in early childhood education and care
		Number of school changes
		Number of changes in educational biography
Learning factors	Knowledge and experience	Index of science activities
		Perceived feedback
		Adaption of instruction
		Number of learning domains with additional instruction
		Total hours of additional instruction
		Number of science disciplines and subjects with additional instruction
		Out-of-school study time per week
		Science learning time per week
		Learning time per week in total
		Child's past science activities
		Student behavior hindering learning
		ICT use outside of school for schoolwork
		ICT use outside of school leisure
		Students' perceived ICT competence
		Students' ICT as a topic in social interaction
		Students' perceived autonomy related to ICT use
	Attitudes and beliefs	Student attitudes, preferences and self-related beliefs: Achieving motivation
		Collaboration and teamwork dispositions: Enjoy cooperation
		Collaboration and teamwork dispositions: Value cooperation
		Students' ICT interest
		Students' expected occupational status
		Personality: Test anxiety
		Subjective well-being: Sense of belonging to school
		Environmental awareness
		Environmental optimism
		Enjoyment of science
		Interest in broad science topics
		Instrumental motivation
		Science self-efficacy
		Epistemological beliefs
Schooling factors	School-related variables	School size
		Class size
		School ownership
		Shortage of educational material
		Creative extra-curricular activities
		Index of science specific resources

(Continued)

TABLE A1 (Continued)

Variable type	Classification	Variable label
		Professional development of teachers
		Teachers' participation
		Shortage of educational staff
		Proportion of teachers in school with bachelor's degree
		Proportion of teachers in school with master's degree
		Proportion of teachers in school with doctoral degree
		Proportion of all teachers fully certified
		Total number of all teachers at school
		Proportion of science teachers by all teachers
		Proportion of science teachers fully certified
		Proportion of science teachers with bachelor's/master's degree and a major in science
		Total number of science teachers at school
		Student-teacher ratio
		School policies for parental involvement
		Educational leadership
		Curricular development
		Instructional leadership
		Responsibility for curriculum
		Responsibility for resources
		School autonomy
		ICT resources
		Use of ICT at school in general
		Number of available computers per student at modal grade
		Proportion of available computers that are connected to the Internet
	Classroom-related variables	Disciplinary climate in science classes
		Teacher support in science classes of students' choice
		Inquiry-based science teaching and learning practices
		Teacher-directed science instruction
		Comparison of science school lessons and additional instruction: Support
		Comparison of science school lessons and additional instruction: Structuredness of lessons
		Comparison of science school lessons and additional instruction: Structuredness of content
		Comparison of science school lessons and additional instruction: Teacher-student relation
		Teacher fairness
		Teacher behavior hindering learning
Parental factors	Parental beliefs and support	Parents' perceived school quality
		Parents' view on science
		Parents' concerns regarding environmental topics
		Parents' view on future environmental topics
		Parents' emotional support
		Parents' current support for learning at home

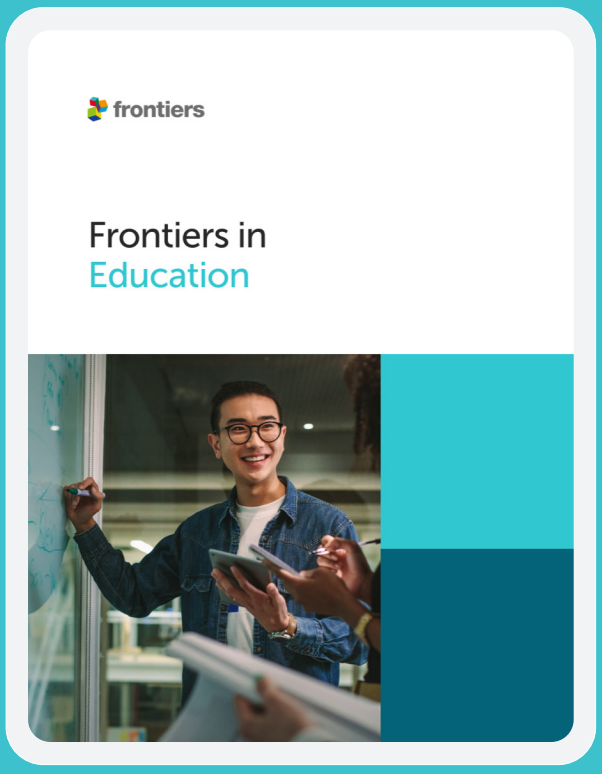
Frontiers in Education

Explores education and its importance for individuals and society

A multidisciplinary journal that explores research-based approaches to education for human development. It focuses on the global challenges and opportunities education faces, ultimately aiming to improve educational outcomes.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact