Pontificia Universidad Católica de Valparaíso
Perspectiva Educacional. Formación de Profesores
Noviembre 2025, Vol 64(3), pp. 246-274

SOCIOECONOMIC FACTORS AND SCHOOL CONTEXT IN MATH ADMISSION TEST SCORES IN CHILE: A MULTILEVEL ANALYSIS

FACTORES SOCIOECONÓMICOS Y CONTEXTO ESCOLAR EN LOS PUNTAJES DE LA PRUEBA DE ADMISIÓN DE MATEMÁTICAS EN CHILE: UN ANÁLISIS MULTINIVEL

Bladimir Padilla (*)
University of Iowa
Lesa Hoffman
University of Iowa
Claudia Suazo
Universidad del Bio-Bio

Resumen

Los puntajes de la prueba de admisión de matemáticas en Chile han recibido constante atención debido a su fuerte asociación con el contexto educativo y los antecedentes socioeconómicos de los estudiantes. Este estudio analiza las relaciones entre los atributos socioeconómicos individuales de los estudiantes, las características de las escuelas y los puntajes obtenidos en la prueba de matemáticas del proceso de admisión a la educación superior chilena en 2022, empleando técnicas de modelamiento multinivel que cuantificaron diferencias intra- e inter-escolares en dichos puntajes. Los resultados revelaron efectos significativos del nivel socioeconómico del estudiante, la educación de los padres y el sexo biológico, además de efectos contextuales relevantes derivados de la composición socioeconómica de las escuelas. Particularmente, una mayor proporción de pares con ventajas económicas, padres con educación superior y estudiantes mujeres dentro de las escuelas estuvo asociado con puntajes más altos en matemáticas. Estos hallazgos resaltan la importancia del contexto socioeconómico en las evaluaciones estandarizadas y plantean preocupaciones sobre la equidad y justicia en el proceso de admisión universitaria en Chile.

Palabras clave: Educación matemática; factores socioeconómicos; desigualdad educativa; admisión a la educación superior; análisis multinivel.

Abstract

Mathematics admission test scores play a critical role in Chilean higher education access, yet their strong association with students' socioeconomic backgrounds and school contexts has raised persistent concerns about equity and fairness in university admissions. In the Chilean admission system for selective universities, standardized test performance carries substantial weight in determining student placement, particularly in mathematics, making it essential to understand how individual and contextual factors shape these outcomes. Despite the test's intended purpose to measure mathematical knowledge and skills aligned with the national curriculum, research evidence suggests that scores systematically reflect socioeconomic stratification rather than purely academic merit. In this context, this study examines the complex relationships between student-level socioeconomic attributes, school-level characteristics, and mathematics test performance in the 2022 Chilean higher education admission process, employing multilevel analysis techniques to disentangle individual and contextual influences.

Drawing on a comprehensive dataset of 253,380 students nested within 2,832 schools, this research quantifies both intra-school and inter-school variation in mathematics scores while examining how individual characteristics and school composition jointly predict math performance. The analytical approach utilized multilevel models that account for the nested structure of students within schools, allowing for simultaneous estimation of student-level effects (Level 1) and school-level contextual effects (Level 2). Key predictors included individual socioeconomic

(*) Autor para correspondencia:
Bladimir Padilla
University of Iowa
240 S Madison St, Iowa City, IA 52242, USA
Correo de contacto: Geraldo-Padilla@UIowa.edu

RECIBIDO: 14.08.2024
ACEPTADO: 27.04.2025
DOI: 10.4151/07189729-Vol.64-Iss.3-Art.1617

status, parental educational attainment, biological sex, school type, and school-level aggregates of these characteristics representing the socioeconomic and demographic composition of the student body.

Results revealed substantial and statistically significant effects across both levels of analysis. At the individual level, students from higher socioeconomic backgrounds, those with parents who attained higher education, and male students demonstrated significantly higher mathematics scores. These individual-level effects remained robust even after accounting for school-level variation, confirming their independent contribution to test performance. Critically, school-level analyses uncovered powerful contextual effects: schools with higher proportions of economically advantaged students, greater concentrations of parents with tertiary education, and increased representation of female students were associated with elevated mathematics scores for all students attending those institutions, independent of individual characteristics. These findings suggest that peer composition and the broader school environment exert considerable influence on math performance beyond individual student attributes.

The variance decomposition revealed that approximately 24% of the total variation in mathematics scores occurred between schools rather than within them, underscoring the substantial role of educational context in shaping achievement. The multilevel models explained considerable proportions of this variation, with school-level predictors accounting for approximately 77% of between-school variance and individual-level predictors explaining about 21% of within-school variance. These findings illuminate how Chilean schools function as socially stratified spaces where educational opportunities and outcomes are fundamentally shaped by institutional composition.

The study's findings carry significant implications for educational policy and university admissions in Chile. The documented associations between socioeconomic characteristics and mathematics test performance challenge assumptions of objectivity and meritocracy in the current admission system. Results suggest that the mathematics test may inadvertently perpetuate existing social inequalities by systematically advantaging students from privileged backgrounds and well-resourced schools while creating barriers for equally capable students from disadvantaged contexts.

Keywords: Mathematics education; socioeconomic factors; educational inequality; higher education admission; multilevel analysis.

1. Introduction

1.1. Admission Process to Selective Universities in Chile

Among Chilean higher education institutions, selective universities are highly attractive to students due to their academic quality, diverse study options. exclusive programs, recognized degrees, and prestige amongst employers (Améstica et al., 2014). Consequently, thousands of students apply annually to these institutions, which have a unified admission process. Regular admission depends on a composite score derived from standardized admission tests in specific subjects and high school grades. Each university program assigns particular weights to these selection factors, and students are selected according to their position among applicants (Larroucau et al., 2015).

Historically, programs have assigned great weight to the admission test scores (Rodríguez et al., 2022a), especially in mathematics. However, this test (referred to as math throughout) has received substantial criticism due to its relationship with students' individual and school characteristics (Catalán & Santelices, 2014), which some researchers consider an unfair barrier for students lacking good mathematics education (Jarpa & Rodríguez, 2017).

1.2. Relationship Between Math Scores and Socioeconomic Attributes

The math test is intended to measure students' knowledge and skills in mathematics (Department of Educational Evaluation, Measurement, and Registration [DEMRE], 2022), aligning its questions with the national curriculum for primary and secondary education (DEMRE, 2023). Despite its psychometric properties (details are described in DEMRE, 2022), math scores have frequently shown a strong correlation with the socioeconomic attributes of the students and their schools.

Local researchers such as Rodríguez and Jarpa (2015), and Vergara and Peredo (2017) have found a strong association between math scores and variables like gender, high school grades, family income, and parental education. One finding is repeated across admissions processes: Students from lower-income families whose parents did not attain higher education typically achieve lower math scores (Catalán & Santelices, 2014; Rodríguez et al., 2021). Additionally, a gender gap in favor of male students has been documented repeatedly (Arias, 2016; Díaz et al., 2019).

Criticism arises because such associations challenge the fairness of the math test, suggesting that it reproduces social inequalities from earlier educational stages into higher education (Jarpa & Rodríguez, 2017; Muñoz & Blanco, 2013). Similar gaps have been observed in national and international standardized assessments, such as SIMCE, PISA, and TIMSS. International research underscores how female and socioeconomically disadvantaged students consistently achieve lower mathematics performance, revealing a broader systemic issue rather than one specific to Chile (León & Salazar, 2014; Schleicher, 2019; Vargas & Matus, 2022).

School characteristics further amplify these differences among students. Studies by Alessandri and Peñafiel (2022), Antivilo et al. (2017), and Muñoz and Redondo (2013) show that students attending private and humanistic schools with a high proportion of socioeconomically well-off peers, score higher than the rest of the students in math. For this reason, some authors have argued that in Chile there are schools for the poor and others for the rich (Bellei, 2015). Chilean schools differ greatly in economic, pedagogical, and human resources, directly affecting their ability to prepare students adequately for standardized testing (Gayo et al., 2019; González, 2017).

International studies have documented educational segmentation as a global phenomenon affecting mathematics achievement. Research from the Organisation for Economic Co-operation and Development (OECD, 2019) and the United Nations Educational, Scientific and Cultural Organization (UNESCO, 2020) points out that students in socioeconomically vulnerable schools often have unequal access to experienced teachers, advanced curricula, and supportive learning environments, which maintain or extend inequalities (Lamb & Fullarton, 2002; Reardon & Owens, 2014).

Favorable educational contexts facilitate effective teaching, enhance study habits, and build positive expectations about higher education (Canales et al., 2016; Espinoza, 2017). Conversely, adverse conditions, such as limited pedagogical resources, poor school climate, and low student motivation, negatively influence student performance (Donoso et al., 2016). The school context has an effect on students that can reinforce or mitigate the effect of certain individual attributes, so it is important to separate both dimensions when analyzing math performance.

To summarize, extensive evidence highlights the strong link between math scores, student socioeconomic attributes, and school contexts. This violates the assumptions of objectivity in the admissions process (DEMRE, 2022) and is a potential bias for student selection. This study aims to analyze the relationship between socioeconomic attributes and scores on the 2022 math test. The leading research question was: Is there a relevant relationship between students' socioeconomic and educational characteristics and scores on the 2022 math test? Furthermore, if so, what is the predictive capacity of these variables?

2. Method

2.1. Data

The data corresponded to the 2022 admission process to selective higher education institutions in Chile. Students included in the dataset belonged to the Chilean educational system who were transitioning from secondary to higher education (that is why we refer to this group as students throughout). The data were collected and processed by the Department of Educational Evaluation, Measurement, and Registration (DEMRE) of the Universidad de Chile. The original database contained information about the students, their families, and their respective schools. Furthermore, additional contextual variables from the national repository of schools, managed by the Chilean Ministry of Education, were incorporated to enrich the dataset and provide comprehensive information about the students' educational context.

2.2. Sample

In 2021, 281,457 students registered to participate in the 2022 admission process. To keep the sample size constant across regression models and be able to assess their relative model fit, students with incomplete data were removed (different sample sizes between models affect their comparison), which decreased the sample size to 150,317. This means that missing data were treated as missing completely at random, which could be modified in future studies using multiple imputation techniques under the assumption of data missing at random (Enders, 2022).

After including the additional school variables in the DEMRE database, incomplete data were removed again, and schools with less than ten students were excluded from the sample to prevent school variables derived from student information from introducing some atomistic fallacy in the analyses. Thus, the final sample consisted of 147,562 students in 2,810 schools, which in the following sections are also referred to as units/observations and clusters, respectively.

2.3. Variables

In our study, level 1 (L1) of the multilevel models refers to the relationship between student characteristics and predicted math scores within a particular school. In contrast, level 2 (L2) refers to the relationship between school characteristics and predicted mean math scores across schools. Although our focus was the socioeconomic differences in math scores, biological student sex was also included, given its relevance in the analysis of math scores in Chile (e. g., in Arias [2016] and Arias et al. [2017]).

For students at L1, binary biological sex, quintile of family income, father's education, mother's education, ranking, and math scores were included. For schools at L2, type, sector, costs, school subsidy agreement (SEP), and participation in the program for access to higher education (PACE) were included. Table 1 shows the variables considered in the analysis.

Table 1

Description of the measures included

Variable	Description	Descriptives
Student attributes		
Graduation	Year of student's graduation.	< 2020 = 12.76% 2020 = 18.63% 2021 = 68.61%
Biological sex	Biological sex declared by students	Male = 42.94% Female = 57.06%
Family Income Quintile	Quintile of family income, created by DEMRE	Quintile 1 = 33.30% Quintile 2 = 18.00% Quintile 3 = 12.80% Quintile 4 = 18.53% Quintile 5 = 17.37%
Father Education	The educational level of the father	Primary = 16.36% Secondary = 53.21% Higher education = 30.34%
Mother Education	The educational level of the mother	Primary = 13.55% Secondary = 56.13% Higher education = 30.32%

Ranking score	Score of the student's GPA in their educational context	Range = 187.00 – 850.00 Mean = 635.34 SD = 125.17
Math score	Student scores on the math test	Range = 150.00 – 850.00 Mean = 505.16 SD = 109.47
School attributes		
Type	Type of education offered by the school	Technical = 22.52% Humanist = 77.48%
Sector	Administration of the school	Public = 32.71% Private Sub. = 55.65% Private = 11.64%
Costs	Associated tuition or fees	No = 61.58% Yes = 38.42%
SEP	Public financing by SEP law	No = 29.20% Yes = 70.80%
PACE	Participant in the Access to Higher Education Program	No = 83.98% Yes = 16.02%

Note. Sub. = Subsidized; SEP = Preferential School Subsidy; PACE = Access to Higher Education Program; n students = 147,562.

2.4. Procedure

All model parameters were estimated using restricted maximum likelihood (REML) in R v.4.2.3 (Posit team, 2023) with the functions in the lme4 package (Bates et al., 2015). The statistical significance of the fixed effects was evaluated at chosen $\alpha < .01$. Relative model fit involving fixed effects was assessed via univariate or multivariate Wald tests using Satterthwaite denominator degrees of freedom. The statistical significance of random effects was assessed through *deviance tests* (−2ΔLL), which compares the sample Log-Likelihood of two nested models using a χ^2 distribution with degrees of freedom equal to the difference in the number of model parameters between competing models (Snijders & Bosker, 2012).

Effect sizes for the fixed effects were computed as *r* or Cohen's *d* indices (i.e., from the t-test statistic and estimated denominator degrees of freedom for each fixed effect). Contrasts between fixed effects were calculated as linear combinations using the lmertTest package (Kuznetsova et al., 2017).

A validation set approach was used to explore the accuracy and replicability of our results (James et al., 2013). This means that the sample was randomly halved at the schools' level. The *training* sample consisted of 72,968 students distributed in 1,405 schools, and the *validation* sample consisted of 74,594 students nested in 1,405 schools. In both datasets, students had similar characteristics.

Finally, as suggested by Rights and Sterba (2019), *total*, *within*-cluster, and *between*-cluster multilevel R^2 measures were estimated, which have similar interpretations to the traditional R^2 measure of single-level regression models: proportion of the respective variance component accounted for by the fixed effects. These estimates were computed using the r2mlm package (Shaw et al., 2023).

3. Results

As a first step, fitting an empty multilevel model for the dependent variable is optional but highly advisable (Hoffman & Walters, 2022). This model predicted the overall math score in the sample. The $L1$ residuals and the $L2$ random intercepts are random variables assumed to be independent and normally distributed with mean 0 and estimated constant variances, which served as the reference to compute R^2 measures.

In the training dataset, adding a random intercept significantly improved the fit of the empty model for math, −2ΔLL (1) = 21,94, $p < .01$. This indicated that the educational context was relevant (i.e., there were substantial differences in average math scores between schools). In fact, the intraclass correlation index (Hox et al., 2017) was .31, which means that approximately 31% of the total variance of math scores in 2022 was attributable to mean differences between schools (i.e., school attributes can account for up to 30% of the variability in math scores).

Other empty models were fitted for $L1$ predictors, given that individual variables can also have significant variance between clusters (Raudenbush & Bryk, 2002). Individual predictors, like the outcome, can exhibit systematic variance between clusters that must be controlled for in the model to avoid conflated or smushed fixed and random effects, which lead to conceptual and statistical errors (Curran et al., 2012; McNeish & Kelley, 2019; Yaremych et al., 2023).

As shown in Table 2, all *L*1 predictors had statistically significant variance at *L*2 (as reported by the −2ΔLL tests), ranging from 15–38% of their total variance. This means that each of these predictors had within- and between-school differences. Consequently, we added their school means as predictors at *L*2 to separate both sources of information (Hoffman, 2019). Otherwise, we would estimate *smushed* fixed effects at *L*1, i.e., weighted combinations of the within- and between-school fixed effects for each predictor (Hoffman & Walters, 2022; Rights & Sterba, 2023).

Table 2

Summary of empty multilevel models for level 1 variables

Predictor	Unconditional ICC	−2ΔLL $p <$
Sex	.21	.01
Family Income Quintile	.24	.01
Education Father	.38	.01
Education Mother	.32	.01
Ranking	.15	.01

Note. ICC = intraclass correlation; n students = 72,968; n schools = 1,405.

The student predictors were *school-mean-centered* to estimate within-school fixed effects at *L*1, whereas the *L*2 predictors were *constant-centered* to estimate between-school fixed effects (McNeish & Kelley, 2019). This was useful for holding constant the marginal frequencies of the non-compared categories in the figures showing within- and contextual fixed effects.

Table 3 shows the results of the multilevel model with the best relative fit. Importantly, although the model was fitted using between-school fixed effects at *L*2, some of them were then transformed into *contextual* fixed effects, which had a more practical interpretation for this study and allowed us to test whether the characteristics of the educational context had an *incremental* contribution to the prediction of math (controlling for the original values of each *L*1 predictor among students).

Table 3

Results of the multilevel model predicting 2022 math scores

Predictor	Estimate	SE	Cohens' *d*	Partial *r*
Level 1				
Intercept	**536.41**	11.44	-	-
Sex				
Male	Ref.	-	-	-
Female CMC	**−34.05**	0.72	−3.24	-
Family Income Quintile				
Quintile 1	Ref.	-	-	-
Quintile 2 CMC	−0.70	0.91	−0.01	-
Quintile 3 CMC	**−3.23**	1.02	−0.02	-
Quintile 4 CMC	**−7.15**	0.91	−0.06	-
Quintile 5 CMC	**−7.28**	1.06	−0.05	-
Education Father				
Primary	Ref.	-	-	-
Secondary CMC	**6.25**	0.95	0.05	-
Higher Education CMC	**17.31**	1.19	0.11	-
Education Mother				
Primary	Ref.	-	-	-
Secondary CMC	**3.43**	1.01	0.02	-
Higher Education CMC	**8.20**	1.22	0.05	-
Ranking CMC / 10	**3.46**	0.03	-	0.43

Level 2				
School Type				
Technical	Ref.	-	-	-
Humanist	**18.40**	1.46	0.19	-
School in PACE				
No	Ref.	-	-	-
Yes	**−9.93**	2.56	−0.21	-
Sex				
Prop. Female − .50	**14.06**	5.16	-	0.07
Family Income Quintile				
Prop. Quintile 2 − .50	**32.39**	13.00	-	0.06
Prop. Quintile 3 − .50	17.39	15.76	-	0.03
Prop. Quintile 4 − .50	−5.67	11.89	-	−0.01
Prop. Quintile 5 − .50	**87.91**	11.15	-	0.20
Education Father				
Prop. Secondary − .50	−1.01	11.79	-	−0.00
Prop. Higher Education − .50	**32.86**	15.24	-	0.05
Education Mother				
Prop. Secondary − .50	**31.94**	12.90	-	0.06
Prop. Higher Education − .50	**66.63**	16.77	-	0.10
[SM Ranking – 600] / 10	−0.37	0.21	-	−0.05
R^2 measures				
Fixed effects at level 1				

Total	.14
Within	.21
Fixed effects at level 2	
Total	.23
Between	.77

Note. Prop. = Proportion; SM = school mean; Bold estimates had associated *p*-value < .01; n students = 72,968; n schools = 1,405.

Our best-fit model had an estimated fixed intercept equal to 536.40, which was the expected math score for reference students in the reference school. At *L*1, female biological sex (male was the reference) had a significant fixed effect, in which the average performance of females on math was estimated to be 34.05 below the average performance of their male peers in the same school. According to its effect size, this difference between groups was large (Cohens' $d = -3.24$). In addition, biological sex was the only *L*1 predictor whose random slope improved the relative fit of the model. This means that the estimated female disadvantage in math varied significantly across schools.

We can use the random slope variance for biological sex to construct a 95% confidence interval for its fixed effect. This interval reflects the estimated variation of the slope across schools. Concretely, 95% of the schools were predicted to have gender gap in math between −48.81 and −19.29, always in favor of males.

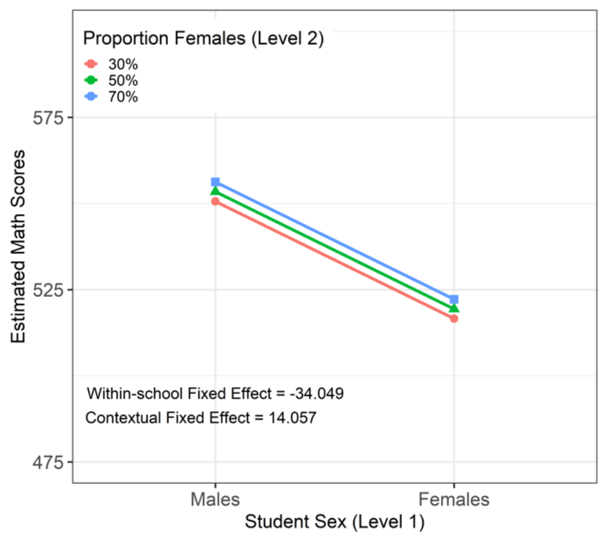
The proportion of females in the school had a positive incremental effect on math scores after controlling for students' biological sex. For instance, if we were to move a female student to an all-female school, her predicted math score would be higher by 14.06 (as compared to being moved to an all-male school and holding the rest of the variables constant).

As some readers may infer, the interpretation of the contextual fixed effects of proportions is hard to follow. As Yaremych et al. (2023) mentioned, the hypothetical contexts for proportions are not always intuitive or possible. For this reason, we divided the contextual fixed effects by 10 to interpret them as the change in the predicted math for each 10% more students with a given characteristic in the hypothetical new school context. Returning to the contextual fixed effect of biological sex, if we were to move a female student to a new school, the model predicted a higher math by 1.40 for every 10% more females in that new school context.

In Figure 1 (and similarly structured Figures 2, 3, and 4), the slopes along the x-axis represent the within-school fixed effect of biological sex across different types of schools (in terms of student composition). In contrast, the distance between the slopes on the y-axis conveys its contextual fixed effect. Concretely, Figure 1 shows the negative fixed effect of biological sex at *L*1. However, as the proportion of females in the educational context increases, so do the predicted math (i.e., a positive contextual fixed effect).

Figure 1

Within and contextual fixed effects of sex on 2022 math scores

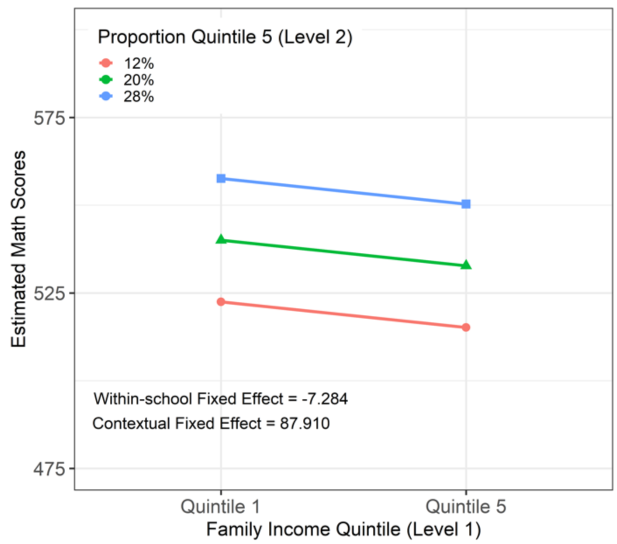


The per capita income quintile of the student's family had significant fixed effects (quintile 1 was the reference). At *L*1, the average math of students from quintile 1 was higher than that of their peers from other quintiles in the same school (as shown by the negative regression coefficients in Table 3). The largest differences were between students in quintile 4 and quintile 5 relative to students in quintile 1, in the same school. The size of each fixed effect of family income quintile was small.

At *L*2, the contextual fixed effects of the per capita income quintile of the student's family were significant for quintiles 2 and 5. This means that only the proportion of students in quintiles 2 and 5 (relative to the proportion of students in quintile 1) had an incremental contribution in predicting math. For instance, considering the fixed effect of quintile 5, if we were to move a student to a new school (holding the rest of his/her attributes constant), for every 10% more peers with families in quintile 5 in that school, their expected math would be higher by 8.79. Figure 2 illustrates these differences between first- and fifth-income quintile students in schools with different proportions of fifth-quintile students.

Figure 2

Within and contextual fixed effects of family income quintile on 2022 math scores

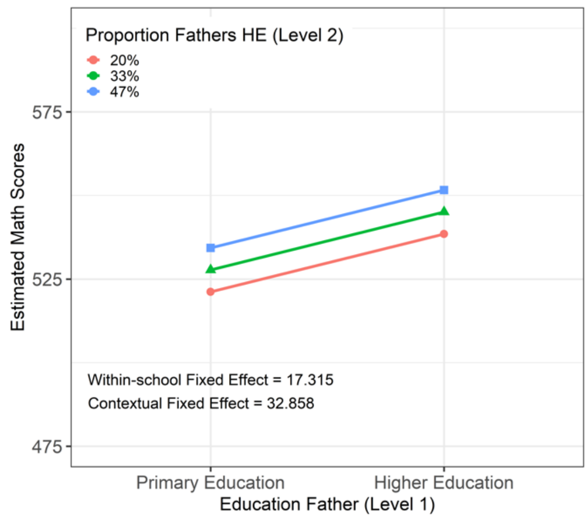


The father's educational level had significant fixed effects at *L*1, although of low and moderate effect sizes. Students from the same school with fathers who achieved secondary and higher education rather than only primary education were predicted to have higher average math by 6.25 and 17.31, respectively. In addition, the fixed effect of parents with higher education was significantly larger than that of parents with secondary education within schools (Est = 11.06, SE = 0.86, $p < .01$).

At *L*2, the contextual fixed effect of fathers' secondary education was not significant; However, the proportion of students whose fathers attained higher education did have a significant effect on the prediction of math. This incremental contribution means that if we move one student to a new school, their predicted math would be higher by 3.28 for every 10% of peers whose parents achieved higher education in the new educational context (holding the rest of the attributes in the model constant). Figure 3 shows both the within-school and contextual fixed effects across illustrative schools.

Figure 3

Within and contextual fixed effects of education father on 2022 math scores

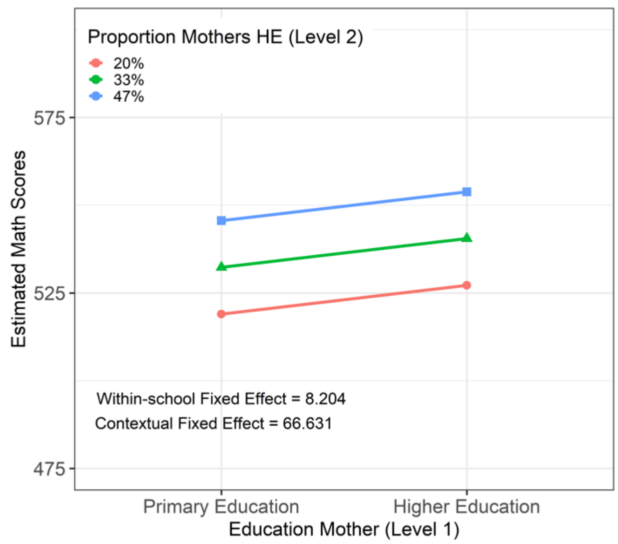


The fixed effects of the mother's education level were significant at both levels of the model. At *L*1, students whose mothers achieved secondary or higher education were predicted to have higher average math by 3.43 and 8.20, respectively, relative to their peers in the same school whose mothers had only primary education. In addition, the benefit of higher mother education in math relative to secondary education was significantly larger (Est = 4.77, SE = 0.83, $p < .01$).

At *L*2, both contextual effects of the mother's education level were statistically significant. The incremental contribution of both fixed effects was positive, so the model predicted higher math scores for students in schools with more educated mothers. For example, controlling for mother's education at *L*1, students moving to a new school would have higher math by 6.66 for every 10% of peers whose mothers attained higher education. Figure 4 shows both the within-school and contextual fixed effects across illustrative schools.

Figure 4

Within and contextual fixed effects of education mother on 2022 math scores



In the case of ranking scores, for every 10 points (the predictor was divided by 10) that students had above their school mean ranking score, the model predicted higher math by 3.46, with a large effect size. This predictor did not have a significant contextual effect (and so a figure was not included).

Concerning the two purely school-level variables that remained in the model, humanist schools were predicted to have significantly higher average math than technical schools. Non-PACE schools scored, on average, 9.93 points higher than PACE schools.

Regarding the effect sizes of our model, two R^2 measures were calculated as suggested in Rights and Sterba (2019). *L*1 fixed effects accounted for 14.40% of the total variance of math, whereas *L*2 fixed effects accounted for 23.50% of its total variance. Together, the fixed effects of the final model accounted for 37.90% of the total variance in math scores. In addition, *L*1 fixed effects accounted for 20.60% of the residual variance, whereas *L*2 fixed effects accounted for 77.3% of the random intercept variance.

Finally, we fitted our final model in the *validation* dataset. The unconditional ICC for math was .28. The fixed effects at *L*1 and *L*2 were similar in direction, magnitude, and statistical significance to what we found in the training dataset. As for the R^2 measures, *L*1 fixed effects accounted for 14.20% of the total variance of math and 19.50% of the residual variance; *L*2 fixed effects accounted for 21.40% of the total variance of math and 77.80% of the random intercept variance. Together, the fixed effects accounted for 35.60% of the total variance in math.

4. Discussion

In the present study, we found that about one-third of the variability in 2022 math could be attributed to differences between schools (as found in both the training and validation data sets). This finding is in line with previous evidence about the relevance of individual and school variables for admission test scores, especially in math and language. For example, Rodríguez and Padilla (2016) and Rodríguez et al. (2022b) have found that performance on admission tests in Chile is strongly correlated with students' socioeconomic and school characteristics.

Similarly, Botella and Ortiz (2018) have shown that the results of some standardized math and language tests used in primary and secondary education by the Chilean Ministry of Education to measure education quality also vary according to the socioeconomic profile of the students and the school they attended. In this sense, the differences in math found in this study could be considered an extension of those observed in primary and secondary education. As Faúndez et al. (2017) have argued, the results of the admission tests have shown strong correlations with the cultural, social, and economic capital of the students and their schools.

Individual and school attributes of the students, such as gender, family income quintile, and school sector, provide relevant information to understand educational outcomes in primary and secondary school, in addition to the results in the admission tests (Rodríguez & Jarpa, 2015). On top of that, they have also been shown to be relevant predictors for analyzing academic performance during the first years of college. For instance, Kri et al. (2013) and Manzi and Carrasco (2021) have shown a substantive relationship between academic performance and socioeconomic individual and school attributes among first-year students in some public universities.

In short, different research findings have found strong correlations between educational outcomes and individual and school predictors. However, these analyses are usually based on descriptive statistics or single-level regression models. Our results have shown that individual and school attributes matter simultaneously, especially features of the educational context—which had incremental contributions in predicting math.

The heterogenous female disadvantage in math found here is in line with previous research. For example, Arias (2016) and Díaz et al. (2019) have shown that in recent years, the average performance of females has been lower than that of males on the math admission test. Moreover, Mizala (2018) and Vargas and Matus (2022) have shown that in the SIMCE, TIMSS, and PISA tests applied in Chile, the average performance of females in math has been below the average performance of males.

This gender gap in math has been analyzed from different perspectives. In the Chilean context, authors such as Cervini et al. (2015) and Mizala (2018) have argued that there are generalized stereotypes that link males with math and females with reading, which are implicitly accepted and shared among parents, teachers, and students. Moreover, authors such as Del Río et al. (2016) and Fernández and Hauri (2016) have argued that males and females have had differentiated access to social and cultural resources that encourage the learning and use of math.

Other studies have explored the relationships between motivation, logical intelligence, gender, and math achievement. They show that, for their respective samples and instruments, males perceive themselves as more prepared for math than females and tend to achieve better educational results in math tests than female students (Cerda et al., 2011). The present study showed that gender gap in math is heterogeneous across schools, although constantly in favor of males.

The contextual fixed effect of the proportion of females in the school allowed us to explore the incremental contribution of gender on math prediction. After controlling for the biological sex of the students, the higher the proportion of females in the school, the higher the predicted math.

The positive contribution of women in the educational context could be linked to studies on school climate and gender, as these have suggested that the proportion of women in the classroom and school is correlated with better teaching conditions for teachers and learning conditions for students, which may, in turn, facilitate math learning (Villalobos et al., 2016). In addition, some studies have suggested that women would be more willing to cooperate with their peers than men to study (Inglés et al., 2012), which could help learn and practice math in the classroom.

Using the family per capita income quintiles constructed by DEMRE, we found that the fixed effects of income among students in the same school were negative, such that students in quintile 1 were predicted to score higher than their peers in higher quintiles. This is a notable finding, given that other research has instead shown that as the economic group of the family increases, so do the individual scores in the admission tests (González et al., 2017; Rodríguez & Padilla, 2016).

A recurrent finding in Chilean educational research is that as household income increases, so do students' educational and learning opportunities, which translates into educational experiences with diverse levels of theory and practice of curricular content (Bellei et al., 2020). Given that the admission tests are elaborated according to these contents, it is not surprising to find that those who obtain the lowest scores are the students classified in the most vulnerable economic groups who studied in poor quality schools (Bellei et al., 2010; Jarpa & Rodríguez 2017; Rodríguez et al., 2020).

In our study, the effect of the family income quintile was contrary to the common literature findings; However, the proportion of fifth-quintile students in the school had a positive contextual effect on math prediction. Attending a school with more peers classified in the fifth quintile relates to significantly higher math scores. This finding aligns with other studies, especially those that analyzed the relationship between school economic classification and school averages in admission tests (Farías & Carrasco, 2012).

This apparent contradiction in the fixed effects of family income quintile at $L1$ and $L2$ could be a consequence of the estimation of conflated fixed effects in previous studies, which did not use models that correctly disaggregate intra- and inter-school relationships. This could be reviewed in future methodological reviews of the topic.

Parent education showed similar results to those found in other studies: Generally, the higher the parents' education, the higher the predicted math. In various studies, these predictors have shown strong correlations with students' interest in learning mathematics and language, study habits at home, performance on standardized tests, and educational choices upon entering higher education (Donoso et al., 2016).

A higher father's education significantly predicted higher math scores among students. While it could be argued that the more educated the father, the more motivated or supported students may be to learn math at home and school, it is not easy to find studies focusing specifically on the relationship between a father's education and student math achievement. In contrast, studies documenting the positive relationship between a mother's education and math learning and performance, or between parents' education and performance on standardized tests, are relatively easy to find in the literature.

Regarding the contextual fixed effects of parent education, the proportion of mothers with higher education in school predicted significantly higher math scores. This finding can be linked to other studies that have also found a positive effect of the mother's education on learning and performance in math in high school and the admission process (Mizala, 2018). As other researchers have pointed out, educated mothers tend to encourage their children to study and learn math at home and school (Cervini & Dari, 2009; Del Rio et al., 2016), which can have additional effects on the student's educational community.

The likely greater involvement of mothers in their children's education also means that they are the family members who participate most in school activities and organizations (Pincheira, 2010; Sánchez et al., 2016). Given this background, we can assume that the presence of educated mothers in the school generate an adequate environment for student learning.

Finally, as shown in Table 3, the fixed effects of the predictors at *L*1 and *L*2 accounted for 35% of the total variability of the 2022 math scores, including approximately 75% of the differences between schools (in both the training and validation data sets). This is a substantial finding, given that most of the predictors represented non-academic attributes, which, in theory, should not be related to the math admission test scores.

4.1. Limitations of the Study

Before concluding, it is pertinent to mention two limitations that affect the scope of our findings. First, regarding missing data, approximately 133,895 students (47.57% of the original sample) were excluded due to missing information on selected predictors. These cases were removed to maintain consistency and comparability across regression models and to ensure accurate evaluation of relative fit. While their exclusion was based on missing predictor information rather than explicitly socioeconomic characteristics, we recognize that this substantial reduction in sample size could limit the generalizability of our findings. Future analyses might consider applying robust multiple imputation methods to retain more participants.

Second, the interpretation of the contextual fixed effects has some limitations, given that the school means and proportions derived from individual predictors were an imperfect representation of school predictors. More school predictors are needed to improve this.

5. Conclusion

Our results revealed significant relationships between socioeconomic attributes of students and their schools and 2022 math scores, which are evidence of potential problems in how math measures students' abilities and knowledge. Characteristics such as the biological sex of the students, their parents' education, and their family's economic income quintile were relevant predictors of scores. In addition, the characteristics of the educational context showed an incremental contribution to score prediction after controlling for students’ attributes. In other words, the composition of the school in terms of the proportion of females, educated parents, and high-income families were relevant contributors to the model’s predictions. Generally, the better these attributes were among students and in the educational context, the better the predicted scores.

6. References

Alessandri, F., & Peñafiel, A. (2022). *Análisis de brechas de puntajes en la PTU y prueba de transición.* Acción Educar. https://accioneducar.cl/wp-content/uploads/2022/08/PDF-VF-Analisis-de-brechas-de-puntajes-en-la-PSU-y-la-PDT.pdf

Améstica, L., Gaete, H., & Llinas-Audet, X. (2014). Segmentación y clasificación de las universidades en Chile: Desventajas de inicio y efectos de las políticas públicas de financiamiento. *Ingeniare. Revista Chilena de Ingeniería, 22*(3), 384-397. https://doi.org/10.4067/S0718-33052014000300009

Antivilo, A., Poblete-Orellana, V., Hernández-Muñoz, J., García, C., & Contreras, P. (2017). Factores individuales, sociodemográficos e institucionales en el acceso de los egresados de la educación media técnico profesional a las instituciones de educación superior. *Calidad en la Educación, 46*, 96-132. https://doi.org/10.4067/S0718-45652017000100096

Arias, Ó. (2016). *Brecha de género en matemáticas: El sesgo de las pruebas competitivas (evidencia para Chile)*. Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile.

Arias, Ó., Mizala, A., & Meneses, F. (2017). *Brecha de género en matemáticas: El sesgo de las pruebas competitivas (evidencia para Chile)*. CONICYT–Chile. https://conicyt.cl/-gendersummit12/wp-content/uploads/2017/12/Oscar-Arias.pdf

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637-/jss.v067.i01

Bellei, C. (2015). *El gran experimento: Mercado y privatización de la educación chilena*. LOM Ediciones.

Bellei, C., González, P., & Valenzuela, J. (2010). Fortalecer la educación pública: Un desafío de interés nacional. In C. Bellei, D. Contreras, & J. Valenzuela (Eds.), *Ecos de la revolución pingüina* (pp. 225-254). Universidad de Chile.

Bellei, C., Orellana, V., & Canales, M. (2020). Elección de escuela en la clase alta chilena: Comunidad, identidad y cierre social. *Archivos Analíticos de Políticas Educativas, 28*(5), 1-27. https://doi.org/10.14507/epaa.28.3884

Botella, M., & Ortiz, C. (2018). Efectos indeseados a partir de los resultados SIMCE en Chile. *Revista Educación, Política y Sociedad, 3*(2), 27-44. https://doi.org/-10.15366/reps2018.3.2.002

Canales, M., Opazo, A., & Camps, J. P. (2016). Salir del cuarto: Expectativas juveniles en el Chile de hoy. *Última Década, 24*(44), 73-108. https://doi.org/10.4067/S0718-22362016000100004

Catalán, X., & Santelices, M. (2014). Rendimiento académico de estudiantes de distinto nivel socioeconómico en universidades: El caso de la Pontificia Universidad Católica de Chile. *Calidad en la Educación, 40*, 21-52. https://doi.org/10.4067/S0718-45652014000100002

Cerda, G., Ortega, R., Pérez, C., Flores, C., & Melipillán, R. (2011). Inteligencia lógica y rendimiento académico en matemáticas: Un estudio con estudiantes de educación básica y secundaria de Chile. *Anales de Psicología, 27*(2), 389-398. https://revistas.um.es/analesps/article/view/123011

Cervini, R., & Dari, N. (2009). Género, escuela y logro escolar en matemática y lengua de la educación media: estudio exploratorio basado en un modelo multinivel bivariado. *Revista mexicana de investigación educativa, 14*(43), 1051-1078.

Cervini, R., Dari, N., & Quiroz, S. (2015). Género y Rendimiento Escolar En América Latina: Los Datos Del SERCE En Matemática y Lectura. *Revista Iberoamericana de Educación, 68*, 99-116. https://doi.org/10.35362/rie680206

Curran, P., Lee, T., Howard, A., Lane, S., & MacCallum, R. (2012). Disaggregating Within-Person and Between-Person Effects in Multilevel and Structural Equation Growth Models. In J. Harring, & G. Hancock (Eds.), *Advances in Longitudinal Methods in the Social and Behavioral Sciences* (pp. 217-254). Information Age Publishing Inc. https://revistas.um.es/analesps/article/view/123011

Del Río, M., Strasser, K., & Susperreguy, M. (2016). ¿Son las habilidades matemáticas un asunto de género?: Los estereotipos de género acerca de las matemáticas en niños y niñas de Kínder, sus familias y educadoras. *Calidad en la Educación,* (45), 20-53. http://dx.doi.org/10.4067/S0718-45652016000200002

Departamento de Evaluación, Medición y Registro Educacional. (2022). *Informe Técnico de las Pruebas de Admisión 2022*. DEMRE.

Departamento de Evaluación, Medición y Registro Educacional. (2023). *Informe Técnico de las Pruebas de Admisión 2023. PAES Regular.* DEMRE. https://demre.cl/estadisticas-/documentos/informes/2023-informe-resultados-paes-regular-admision-2023.pdf

Díaz, K., Ravest, J., & Queupil, J. P. (2019). Brechas de género en los resultados de pruebas de selección universitaria en Chile: ¿Qué sucede en los extremos superior e inferior de la distribución de puntajes? *Pensamiento Educativo: Revista de Investigación Educacional Latinoamericana, 56*(1), 1-19. https://doi.org/10.7764/PEL.56.1.2019.5

Donoso, P., Rico, N., & Castro, E. (2016). Creencias y concepciones de profesores chilenos sobre las matemáticas, su enseñanza y aprendizaje. *Profesorado: Revista de Currículum y Formación de Profesorado, 20*(2), 76-97. https://doi.org/10.30827-/profesorado.v20i2.10409

Enders, C. K. (2022). *Applied missing data analysis*. Guilford Publications.

Espinoza, O. (2017). Privatización de la educación superior en Chile: Consecuencias y lecciones aprendidas. *EccoS Revista Científica, 44*, 175-202. https://doi.org/-10.5585/eccos.n44.8070

Farías, M., & Carrasco, R. (2012). Diferencias en resultados académicos entre educación técnico-profesional y humanista-científica en Chile. *Revista Calidad en la Educación,* (36), 87-121. https://doi.org/10.31619/caledu.n36.118

Faúndez, R., Labarca, J., Cornejo, M., Villarroel, M., & Gil, F. (2017). Ranking 850: Transición a la educación terciaria de estudiantes con desempeño educativo superior y puntaje PSU insuficiente. *Pensamiento Educativo, 54*(1), 1-11. https://doi.org/-10.7764/PEL.54.1.2017.2

Fernández, M., & Hauri, S. (2016). Resultados de aprendizaje en La Araucanía: La brecha de género en el Simce y el androcentrismo en el discurso de docentes de lenguaje y matemática. *Calidad en la Educación, 45*, 54-89. https://doi.org/10.4067/S0718-45652016000200003

Gayo, M., Otero, G., & Méndez, M. (2019). Elección escolar y selección de familias: Reproducción de la clase media alta en Santiago de Chile. *Revista Internacional de Sociología, 77*(1), e120. https://doi.org/10.3989/ris.2019.77.1.17.310

González, P., Arancibia, V., & Boyanova, D. (2017). Talento académico, vulnerabilidad escolar y resultados en la prueba de selección universitaria. *Estudios Pedagógicos (Valdivia), 43*(1), 171-191. https://dx.doi.org/10.4067/S0718-07052017000100011

González, R. (2017). Segregación educativa en el sistema chileno desde una perspectiva comparada. *El primer gran debate de la reforma educacional: Ley de Inclusión Escolar*, 48-91.

Hoffman, L. (2019). On the interpretation of parameters in multivariate multilevel models across different combinations of model specification and estimation. *Advances in Methods and Practices in Psychological Science, 2*(3), 288-311. https://doi.org/10.1177-/2515245919842770

Hoffman, L., & Walters, R. (2022). Catching up on multilevel modeling. *Annual Review of Psychology, 73*, 659-689. https://doi.org/10.1146/annurev-psych-020821-103525

Hox, J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel Analysis: Techniques and Applications.* Routledge. https://doi.org/10.4324/9781315650982

Inglés, C., Díaz, Á. G., Ruiz, C., Delgado, B., & Martínez, M. (2012). Auto-atribuciones académicas: Diferencias de género y curso en estudiantes de educación secundaria. *Revista Latinoamericana de Psicología, 44*(3), 53-64.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R.* Springer. https://doi.org/10.1007/978-1-4614-7138-7

Jarpa, C., & Rodríguez, C. (2017). Segmentación y exclusión en Chile: El caso de los jóvenes primera generación en educación superior. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud, 15*(1), 327-343. https://doi.org/10.11600/-1692715x.1512028032016

Kri, F., Gil, F., Maximo, G., & Lamatta, C. (2013). *Ranking de notas como predictor del éxito en educación superior. Estudio de Caso Universidad de Santiago. Informe Final.* Consejo Nacional de Educación.

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(13), 1-26. https://doi.org/-10.18637/jss.v082.i13

Lamb, S., & Fullarton, S. (2002). Classroom and school factors affecting mathematics achievement: A comparative study of the US and Australia using TIMSS. *Australian Journal of Education, 46*(2), 154-171. https://doi.org/10.1177/000494410204600205

Larroucau, T., Ríos, I., & Mizala, A. (2015). Efecto de la incorporación del ranking de notas en el proceso de admisión a las universidades chilenas. *Pensamiento Educativo: Revista de Investigación Latinoamericana (PEL), 52*(1), 95-118. https://doi.org/-10.7764/PEL.52.1.2015.8

León, V., & Salazar, A. (2014). Diferencias de género en matemática y lenguaje en alumnos de colegios adventistas en el sistema de medición de la calidad de la Educación (SIMCE) en Chile. *Apuntes Universitarios. Revista de Investigación, 4*(2), 81-106.

Manzi, J., & Carrasco, D. (2021). Validity Evidence of the University Admission Tests in Chile: Prueba de Selección Universitaria (PSU). In J. Manzi, M. García, & S. Taut (Eds.), *Validity of Educational Assessments in Chile and Latin America* (pp. 331-352). Ediciones Universidad Católica de Chile. https://doi.org/10.1007/978-3-030-78390-7_14

McNeish, D., & Kelley, K. (2019). Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. *Psychol Methods, 24*(1), 20-35. https://dx.doi.org/-10.1037/met0000182

Mizala, A. (2018). Género, cultura y desempeño en matemáticas. *Revista Anales, 14*, 127-150.

Muñoz, M., & Blanco, C. (2013). Una taxonomía de las universidades chilenas. *Calidad en la Educación, 37*, 181-213. https://doi.org/10.4067/S0718-45652013000100005

Muñoz, P., & Redondo, A. (2013). Desigualdad y logro académico en Chile. *Revista de la CEPAL, 109*, 107-123. https://doi.org/10.18356/5d9855fa-es

Organisation for Economic Co-operation and Development. (2019). *PISA 2018 Results (Volume II): Where All Students Can Succeed.* OECD Publishing. https://dx.doi.org10.1787/b5fd1b8f-en

Pincheira, L. (2010). La participación educativa de padre, madre y/o apoderado en el centro educativo: Mito o realidad. *REXE: Revista de Estudios y Experiencias en Educación, 9*(17), 107-114.

Posit team. (2023). *RStudio: Integrated development environment for R*. Posit Software, PBC. https://posit.co/

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage Publications.

Reardon, S. F., & Owens, A. (2014). 60 Years After Brown: Trends and Consequences of School Segregation. *Annual Review of Sociology, 40*, 199-218. https://doi.org/10.1146/-annurev-soc-071913-043152

Rights, J., & Sterba, S. (2019). Quantifying explained variance in multilevel models: An integrative framework for defining R-squared measures. *Psychological Methods, 24*, 309-338. https://doi.org/10.1037/met0000184

Rights, J., & Sterba, S. (2023). On the common but problematic specification of conflated random slopes in multilevel models. *Multivariate Behavioral Research, 58*(6), 1-28. https://doi.org/10.1080/00273171.2023.2174490

Rodríguez, C., Espinosa, D., Padilla, G., & Suazo, C. (2022a). Entre el talento académico y la segmentación socioeducativa: Admisión universitaria de estudiantes Top 10% ranking en Chile. *Revista Cubana de Educación Superior, 41*(2), 2-14. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0257-43142022000200004

Rodríguez, C., Espinosa, D., Padilla, G., & Suazo, C. (2022b). Trayectoria escolar y procesos de admisión universitaria en Chile: Entre el talento académico y la reproducción de brechas. *Estudios Pedagógicos, 48*(3), 227-241. https://doi.org/10.4067/S0718-07052022000300227

Rodríguez, C., & Jarpa, C. (2015). Capacidad predictiva de las notas en enseñanza media sobre el rendimiento en pruebas de selección universitaria: El caso chileno. *Aula Abierta, 43*(2), 61-68. https://doi.org/10.1016/j.aula.2015.03.002

Rodríguez, C., & Padilla, G. (2016). Trayectoria escolar y selección universitaria: Comportamiento del ranking como factor de inclusión a la educación superior. *Sophia, 12*(2), 195-206. https://doi.org/10.18634/sophiaj.12v.2i.376

Rodríguez, C., Padilla, G., & Espinosa, D. (2021). No todo es prueba de selección universitaria: El ranking como vía de inclusión a la universidad en Chile. *Sophia, 17*(2), 1-15. https://doi.org/10.18634/sophiaj.17v.2i.1026

Rodríguez, C., Padilla, G., & Suazo, C. (2020). Medición de calidad educativa en Chile: Lo que reportan los indicadores de desarrollo cognitivo, personal y social en la escuela. *Revista Pilquen, 17*(1), 34-48. https://revele.uncoma.edu.ar/index.php/-psico/article/view/2647

Sánchez, A., Reyes, F., & Villarroel, V. (2016). Participación y expectativas de los padres sobre la educación de sus hijos en una escuela pública. *Estudios Pedagógicos, 42*(3), 347-367. https://doi.org/10.4067/S0718-07052016000400019

Schleicher, A. (2019). *PISA 2018: Insights and interpretations*. OECD Publishing. https://www.educationduepuntozero.it/wp-content/uploads/2020/01/-tabelleOCSEFierli.pdf

Shaw, M., Rights, J., Sterba, S., & Flake, J. (2023). r2mlm: An R package calculating R-squared measures for multilevel models. *Behavior Research Methods, 55*(4), 1942-1964. https://doi.org/10.3758/s13428-022-01841-4

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

United Nations Educational, Scientific and Cultural Organization. (2020). *Global education monitoring report 2020: Inclusion and education—All means all*. UNESCO Publishing. https://unesdoc.unesco.org/ark:/48223/pf0000373718

Vargas, C., & Matus, C. (2022). Brechas persistentes de género en matemáticas en las pruebas nacionales chilenas SIMCE. *Estudios Pedagógicos, 48*(1), 389-400. https://doi.org/10.4067/S0718-07052022000100389

Vergara, G., & Peredo, H. (2017). Relación del desempeño académico de estudiantes de primer año de universidad en Chile y los instrumentos de selección para su ingreso. *Revista Educación, 41*(2), 95-104. http://dx.doi.org/10.15517/revedu.v41i2.21514

Villalobos, C., Wyman, I., Schiele, B., & Godoy, F. (2016). Composición de género en establecimientos escolares chilenos: ¿Afecta el rendimiento académico y el ambiente escolar? *Estudios Pedagógicos, 42*(2), 379-394. https://doi.org/10.4067/S0718-07052016000200022

Yaremych, H., Preacher, K., & Hedeker, D. (2023). Centering categorical predictors in multilevel models: Best practices and interpretation. *Psychol Methods*, 613-630. https://doi.org/10.1037/met0000434