Supporting information

Luquine Jr. *et al.* Comparing school experiences across Brazilian, Swiss, and Uruguayan adolescents: a measurement invariance study. *BMC Psychology*, 2025.

TABLE OF CONTENTS

EDUCATION IN SWITZERLAND, URUGUAY, AND BRAZIL

The three study contexts represent dramatically different educational resource environments and academic outcomes. Switzerland invests $17,333 (PPP) per student annually with teacher starting salaries of $76,318, while Brazil spends only $3,583 per student with teacher salaries of $20,261 (Table A1). These resource disparities translate into substantial PISA performance gaps: Swiss students score 508 in mathematics compared to 409 in Uruguay and 379 in Brazil. The countries also differ structurally: Switzerland operates an early-tracking system, Uruguay maintains a centralized system with limited technical pathways, and Brazil manages a fragmented dual system with extreme public-private quality differences (Table A2). Student-teacher ratios range from 9.8 in Switzerland to 16.7 in Brazil, reflecting varied classroom learning conditions that may influence student experiences and measurement validity.

Table A1: Country sociodemographic and education characterization.

	Switzerland	**Uruguay**	**Brazil**
Sociodemographic			
GDP per capita, adjusted by USD PPP (2023) [1]	$89,546	$34,426	$21,107
Human Capital Index (2020) [2]	76% productivity potential	60% productivity potential	55% productivity potential
Gini index [3]	Moderate inequality (33.8%)	Moderate inequality (40.9%)	Extreme inequality (51.6%)
Academic performance (PISA 2022) [4]			
Mathematics (average 472)	508	409	379
Reading (average 476)	483	430	410
Science (average 485)	503	435	403
Educational facts			
Government expenditure on educational institutions per student (USD PPP) (2021) [5]	$17,333	Limited data available	$3,583
Private school enrollment (2022) [6]	12%	11%	14%
Students per teacher (secondary)[7]	9.8 (2017)	12.7 (2010)	16.7 (2017)
Teacher starting salary in upper secondary education public institutions (USD PPP) [5]	$76.318	Limited data available	$20,261

USD PPP – purchasing power parity (PPP) international US dollars conversion factor.

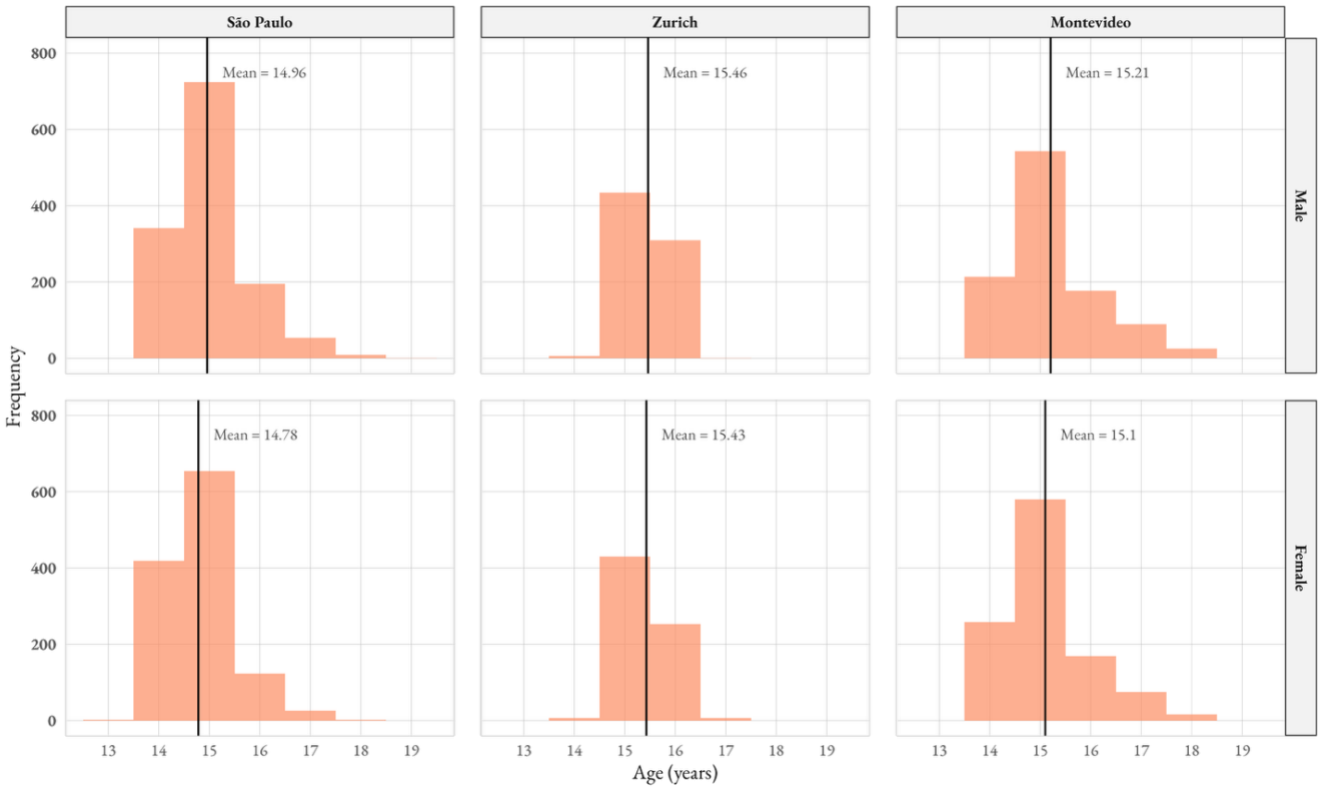
Table A2: Comparison of educational system operation and school outcomes in Switzerland, Uruguay, and Brazil.

Characteristic	Switzerland	Uruguay	Brazil
System structure	Early tracking system (age 12) with high stratification, significant cantonal variation, and limited track mobility [8, 9]	Centralized three-level system with modest technical pathways and persistent secondary completion challenges [10, 11]	Highly stratified dual system with fragmented governance and inverse public-private accessibility [12]
Tracking/streaming	Formal early tracking (ages 10-13) with limited permeability between academic and vocational tracks [8, 9]	Later differentiation with some flexibility between academic-technical pathways from secondary level [10, 11]	Informal tracking through school quality stratification and early grade retention patterns [13]
School day structure	Half-day primary with afternoon activities; full-day secondary with cantonal variation	Traditional morning sessions with some full-day pilot programs in secondary education; only 3% of schools operate full-time (7 hours/day) [10]	Varies dramatically by sector and region
Language considerations	Multilingual system with German, French, Italian, and Romansh instruction options	Monolingual Spanish system with limited indigenous language recognition	Portuguese dominance with 274+ indigenous languages marginalized in formal education
Governance	Federal coordination with 26 cantonal education systems maintaining strong autonomy	Nationally centralized with four education councils but limited local input	Complex federalism with federal standards; upper secondary education is offered by 26 states and 5,570 municipal primary systems, which are responsible for most students
Teacher education	Universities and Federal Institutes of Technology with theory-practice integration, requiring continuing professional development	State-administered entrance exams for theory and practice; strict degree requirements create two-tiered workforce of permanent vs. temporary teachers [14]	Teacher training based on university degrees, but strong regional disparities persist regarding course availability; employment opportunities vary due to complex governance structures
Educational inequality by parental education	About 21% performance variance explained by family background [15]	Mother's education remains significant predictor of academic achievement [10]; residential segregation reproduces educational inequality through school assignment [16]	Extreme impact: richest quintile children 4x more likely to complete secondary than poorest quintile [17]
Educational segregation	Early tracking creates ability-based segregation; high between-school segregation through tracking system and residential patterns [15]	Moderate territorial segregation with limited school choice and centralized assignment [10]	Extreme public-private segregation with minimal social mixing between school types [12, 16]
Access to higher education	Dual pathways: 25% academic track to universities, 65% vocational with applied sciences access. Tracking may limit social mobility for some [15]	Free public university system with no entrance barriers; limited capacity creates competition [18]	Inverse accessibility: elite public universities serve wealthy, mass private serves less affluent. Federal universities largely inaccessible to lower SES: only 20% from bottom two income quintiles [17]

SAMPLE DESCRIPTION

Our analysis included adolescents from three cities: São Paulo, Brazil (n = 2,680), Zurich, Switzerland (n = 1,447), and Montevideo, Uruguay (n = 2,204). The gender distribution was similar across sites, with females comprising 48% of participants in São Paulo, 48% in Zurich, and 51% in Montevideo (p = .088). Mean age distribution varied across the three samples (p < .001), with Zurich participants being the oldest (m = 15.44 years, sd = 0.36), followed by Montevideo (m = 15.15 years, sd = 0.91), and São Paulo (m = 14.88 years, sd = 0.69). São Paulo showed the widest age range (12.85-18.91 years), while Zurich had the narrowest distribution (14.10-16.90 years), reflecting differences in grade retention policies and age-grade correspondence across educational systems. Despite these demographic variations, all samples successfully captured the target population. Mean ages fell in the expected range for 9th-grade students across the three countries. Below, we describe each sample separetaly capturing their unique socieconomic features.

Figure A1: Histogram distribution and mean age (year) of student by city and sex.



Zurich

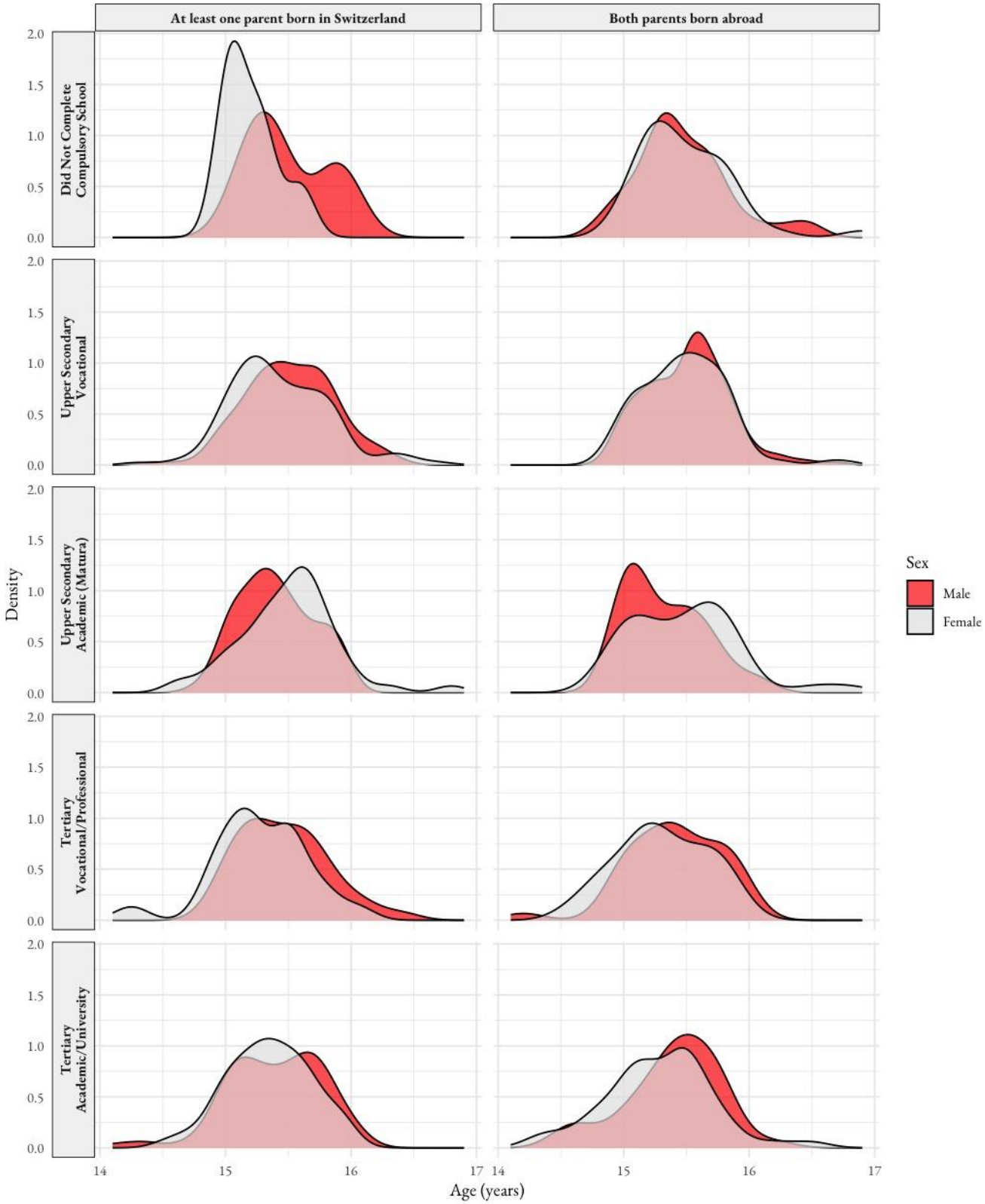
In Zurich, approximately half of participants have both parents born abroad (49.4%). Family educational attainment varied significantly by migrant status ($p < .001$), with Swiss-background families showing higher rates of tertiary academic education (28% vs 14%) and lower rates of incomplete compulsory schooling (1.5% vs 12%).

Table A3: Sociodemographic characterization by migration background (z-proso).

Characteristic	***n***	**At least one parent born in Switzerland *n* = 714**	**Both parents born abroad *n* = 697**	**p-value[3]**
Sex[1]	1,411			0.3
Male		379 (53%)	351 (50%)	
Female		335 (47%)	346 (50%)	
Age (years)[2]	1,411	15.4 (0.4)	15.5 (0.4)	0.2
Family education level[1]	1,340			<0.001
Did not complete compulsory school		10 (2%)	78 (12%)	
Upper secondary vocational		295 (43%)	350 (53%)	
Upper secondary academic (Matura)		105 (15%)	76 (12%)	
Tertiary vocational/professional		83 (12%)	60 (9%)	
Tertiary academic/university		192 (28%)	91 (14%)	

Note: [1] *n* (%), [2] mean (*sd*), [3] Pearson's Chi-squared test; Wilcoxon rank sum test.

Figure A2: Family education level by migration background (z-proso).



Montevideo

The m-proso sample showed relatively high parental educational attainment, with approximately one-third having tertiary-educated parents and the majority completing secondary education (56-61%). No ethnicity/racial variable was available for the m-proso sample.

Table A4. Sociodemographic characterization by sex (m-proso).

Characteristic	*n*	Male *n* = 1,074	Female *n* = 1,110	p-value[3]
Age (years)[1]	2,146	15.2 (0.9)	15.1 (0.9)	0.011
Family education level[2]	2,155			0.060
Primary		102 (10%)	125 (11%)	
Secondary		642 (61%)	609 (56%)	
Terciary		316 (30%)	361 (33%)	

Note: [1] mean (*sd*), [2] *n* (%), [3] Wilcoxon rank sum test; Pearson's Chi-squared test.

São Paulo

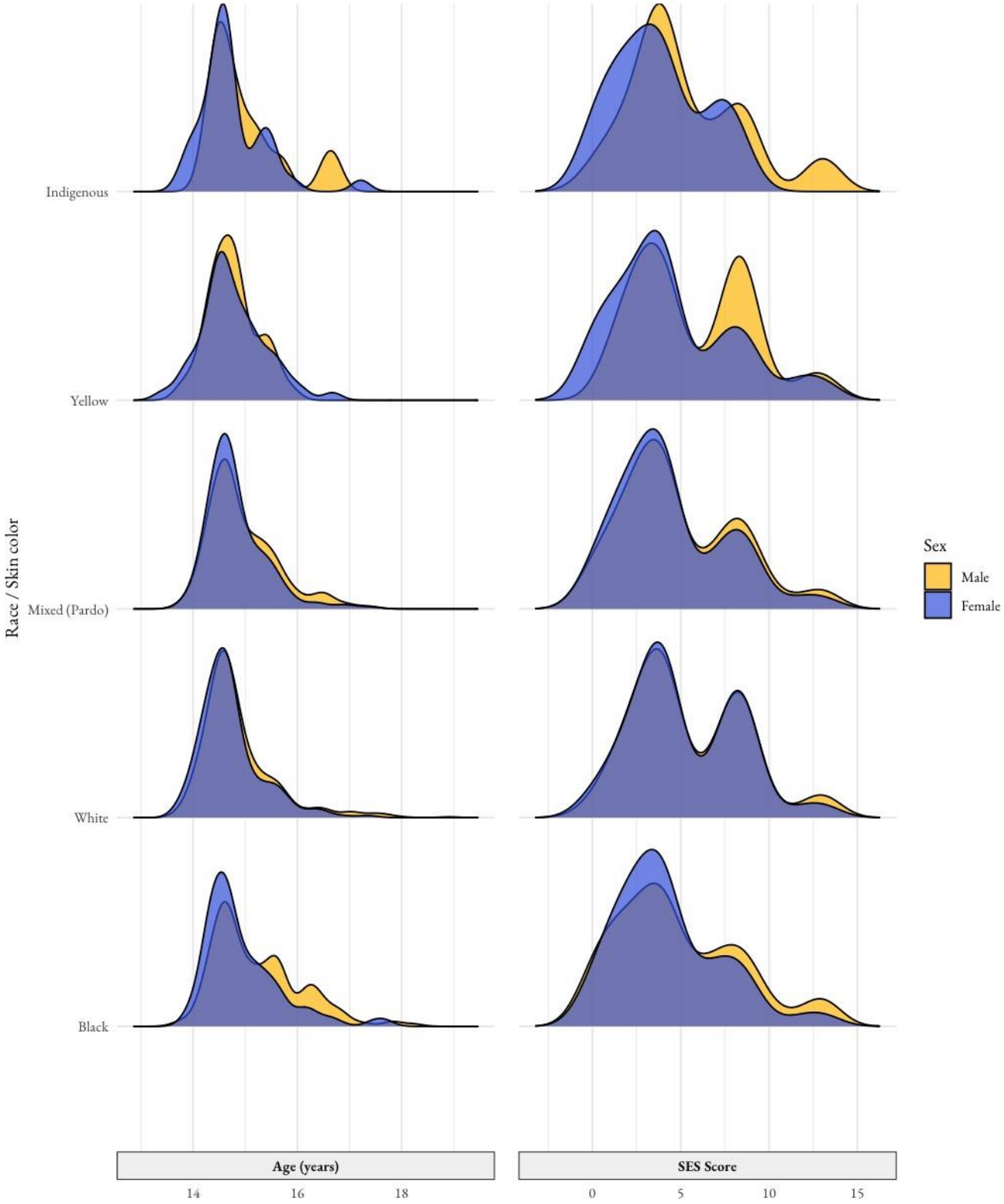
Considering the Brazilian race classification based on self-reported skin-color, São Paulo participants presented with significant disparities for socioeconomic status ($p < .001$), with White students showing the highest SES scores ($m = 5.4$) and Indigenous students the lowest ($m = 4.5$).

Table A5: Sociodemographic characterization by race/skin color (sp-proso).

Characteristic	*n*	Black *n* = 340	White *n* = 1,191	Mixed (Pardo) *n* = 955	Yellow *n* = 101	Indigenous *n* = 70	p-value[3]
Sex[1]	2,588						0.039
Male		185 (57%)	591 (51%)	503 (54%)	43 (43%)	31 (46%)	
Female		138 (43%)	574 (49%)	429 (46%)	57 (57%)	37 (54%)	
Age (years)[2]	2,595	15.1 (0.8)	14.8 (0.7)	14.9 (0.6)	14.8 (0.6)	14.9 (0.7)	<0.001
SES Score[2]	2,520	4.6 (3.3)	5.4 (3.1)	4.7 (3.2)	5.1 (3.4)	4.5 (3.0)	<0.001

Note: [1] N (%), [2] mean (sd), [3] Pearson's Chi-squared test; Wilcoxon rank sum test.

Figure A3: Age and SES score distribution by race/skin color and sex in the sp-proso sample.



SAMPLING PROCEDURE

z-proso – Zurich

The Zurich Project on the Social Development from Childhood to Adulthood (z-proso) was launched in 2004 as a combined randomised field experiment and cohort study [19]. The study employed a cluster-randomized design targeting first-grade children from the city of Zurich public schools in 2004. Using stratified random sampling based on school size and district location, 56 schools with 1,675 children were selected, with an intentional slight overrepresentation of low socioeconomic status districts. Schools were organized in 14 matched quadruplets and randomly assigned to four conditions: parenting program (voluntary for parents), social skills training (mandatory for teachers), combined, or control.

Initial recruitment achieved 81.2% child participation (n = 1,360), though socioeconomically disadvantaged and immigrant families were underrepresented despite comprehensive multilingual recruitment efforts including materials in ten languages. A second major recruitment phase at ages 13-15 successfully re-contacted the entire eligible sample (n = 1,675), significantly reducing initial demographic participation biases. This yielded high participation rates of 81.6% (age 13) and 86.4% (age 15, n = 1,147).

m-proso – Montevideo

The Montevideo Project on the Social Development of Children and Adolescents (m-proso) targeted all 9th-grade students in Montevideo, Uruguay in 2013 [20]. The study employed a stratified cluster-randomized sampling approach with classes as randomization units to obtain approximately 2,000 participants. Sampling was stratified across three school types that reflected different socioeconomic backgrounds: private high schools (32% of target sample), public high schools (63%), and technical schools (4%). Within each stratum, a systematic sampling procedure was applied where classes were selected at fixed intervals from an ordered list of all schools.

The sampling frame consisted of all 9th-grade classes in Montevideo (211 private school classes across 99 schools, 324 public school classes across 53 schools, and 22 technical school classes across 7

schools). To account for potential attrition, researchers targeted a raw sample of approximately 2,500 students. The final selection yielded 90 classes from 85 different schools: 33 classes from private schools, 50 classes from public schools, and 7 classes from technical schools.

According to school records, 2,690 students were registered in these classrooms. Non-attendance on survey day was 17.4% overall, with notable differences across school types: 12.1% in private schools, 19.2% in public schools, and 24.0% in technical schools. This absenteeism rate aligns with previous Uruguayan studies showing dropout rates of 19-25% among 15-year-olds, suggesting systematic differences between attendees and non-attendees that may impact study results.

The realized sample slightly overrepresented public high schools and technical schools while underrepresenting private schools compared to the target population (34.3% private, 58.4% public, 7.2% technical vs. population distribution of 32.7%, 63.2%, and 4.1%, respectively). In total, 2,204 students participated (82.6% of the targeted population), but in the original m-proso analysis 20 questionnaires were eliminated due to excessive missing data (n = 2,184).

sp-proso – São Paulo

The São Paulo Project on the Social Development of Children and Adolescents (sp-proso) targeted 9th-grade students enrolled in city of São Paulo public and private schools in 2017 [21]. This grade was selected due to students' expected age (14-15 years), a period associated with peak risk behaviors, and before increased dropout rates typically occur in high school. The study employed a stratified cluster sampling approach with three strata: state public schools, municipal public schools, and private schools. Classes served as the sampling units. Based on 2015 school census data, 156 different schools were initially selected, with 128 designated for initial data collection and 28 as reserves. The estimated sample size was 2,849 students, with an anticipated 10% non-response rate, yielding 3,300 target students.

All reserve schools were utilized due to higher-than-expected student absences. From the 61 private schools selected, 29 declined participation (47.5% refusal rate), while public schools had much lower

refusal rates (one state school and seven municipal schools). The final sample included 119 schools: 39 state schools (32.8%), 48 municipal schools (4.3%), and 32 private schools (26.9%). Of 2,816 students present on data collection days, 96 (3.4%) declined participation or lacked parental consent, and 18 (.6%) were excluded due to severe cognitive disabilities or reading limitations. This resulted in a 4% total loss among present students. The final sample comprised 2,702 questionnaires, representing 94.8% of the initial target. The realized sample distribution showed 68.1% of students from public schools and 31.9% from private schools, compared to the actual population distribution of 79% and 21%, respectively, indicating a slight overrepresentation of private school students. For analysis purposes, an additional 22 questionnaires were excluded from the sample due to a response of less than 20% of the questionaire (n = 2,680).

TRANSLATION AND ADAPTATION PROCEDURE

Both the m-proso and sp-proso studies adapted the questionnaire originally developed for z-proso wave 6 (2013), ensuring cross-cultural comparability while accommodating local contexts. For m-proso, the German original was translated by a qualified native Spanish-speaking translator experienced in social science projects. When the German questionnaire contained scales originally developed in English, the English version was consulted to maintain equivalence. The translation underwent several validation steps: review by a second translator, consistency checks, and examination by two German-speaking z-proso researchers.

The sp-proso employed a more comprehensive approach with four independent translators working simultaneously: two translating from German to Portuguese, one from English to Portuguese, and one from Spanish to Portuguese. Researchers compared these translations, resolved discrepancies through consultation with the originals and translators, and finalized an initial Portuguese version. This version was then reviewed by the principal investigators of both z-proso and m-proso, with all queries discussed with the sp-proso principal investigator. Translations to the items comprised in the school experience scale are show in Table A6.

Pre-testing

In Montevideo, the questionnaire was reviewed by five qualified local informants (two school directors, one teacher, and two sociologists with youth crime research experience). Three pre-tests were conducted: two small initial tests (with 3 and 8 adolescents) to estimate completion time and identify design issues, followed by a larger school-based pre-test with 121 students from both private (n = 58) and public schools (n = 63). Similarly, in São Paulo, pre-testing involved 114 students across five schools (two private, two state public, one municipal public). The pre-testing assessed question comprehension, response structure understanding, and completion time, followed by post-administration discussions with students to evaluate comprehension.

Adaptations

Both studies made adaptations to accommodate local contexts and time constraints. The m-proso expanded some scales (e.g., the morality scale increased from 5 to 14 items) and added new ones (e.g., school legitimacy scale), while removing or shortening others to keep the questionnaire within 80-90 minutes for completion. For sp-proso, the pre-test indicated completion times exceeding two hours, necessitating questionnaire reduction. These exclusions were decided collaboratively with z-proso and m-proso researchers to preserve comparability and maintain the psychometric properties of the scales. Despite these modifications, both projects maintained the majority of instruments identical to those used in Zurich, enabling valid cross-cultural comparisons across the three sites.

Table A6: School experience's scale items final translation in English, German, Spanish, and Portuguese.

Subscale	Item label	z-proso		m-proso	sp-proso
		School (English)	**Die Schule (German)**	**La escuela (Spanish)**	**Escola (Portuguese)**
Bond to class	Sense of community in class	We have a really good sense of community within the class.	Wir haben eine richtig gute Klassengemeinschaft.	Tenemos un grupo muy bueno en mi clase.	Nós somos bem unidos na nossa turma.
	Get along with classmates	I have a good relationship with the other adolescents in my class.	Ich komme mit den Jugendlichen in meiner Klasse gut aus.	Me llevo bien con la gente de mi clase	Eu me dou bem com os adolescentes da minha turma.
	Classmates are nice to me	The other adolescents in my class are nice to me.	Die anderen Jugendlichen in meiner Klasse sind nett zu mir.	La gente de mi clase se porta bien conmigo	Os outros adolescentes da minha turma são legais comigo.
Bond to teacher	Teacher treats me fairly	My teacher treats me fairly.	Meine Lehrerin/mein Lehrer behandelt mich gerecht.	Mi profesor/a es justo/a conmigo.	No geral, meus professores me tratam de forma justa.
	Get along with teacher	I have a good relationship with my teacher.	Ich komme mit meiner Lehrerin/meinem Lehrer gut aus.	Me llevo bien con mi profesor/a.	No geral, eu me dou bem com meus professores ou minhas professoras.
	Teacher helps when needed	My teacher helps me when necessary.	Wenn es nötig ist, hilft mir meine Lehrerin/mein Lehrer.	Cuando es necesario, mi profesor/a me ayuda	No geral, meus professores me ajudam quando eu preciso deles.
Future orientation	Working towards interesting job	When I grow up I want to have an interesting job, and I'm doing everything now to work towards that goal.	Ich möchte später eine interessante Arbeit haben und mache jetzt schon alles dafür, was ich kann.	Me gustaría tener un trabajo interesante más adelante, y ahora trato de hacer todo lo que puedo para conseguirlo	Eu gostaria de ter um emprego interessante quando for mais velho e agora faço de tudo para conseguir isso.
	Try hard at school for future job	I try hard at school to have a good job later in life.	Ich gebe mir in der Schule Mühe, um später einen guten Job zu haben.	En el centro de estudios me esfuerzo para poder tener um buen trabajo después	Eu me esforço na escola para conseguir um bom emprego no futuro.
	Doing well at school is important	Doing well at school is important to me.	Gute Schulleistungen sind für mich sehr wichtig.	Para mí es muy importante que me vaya bien en los estudios	O bom desempenho escolar é importante para mim.
School commitment	Likes going to school	I enjoy going to school.	Ich gehe gern in die Schule.	Me gusta ir al centro de estudios.	Eu gosto de ir à escola.
	Likes doing homework	I enjoy doing my homework.	Ich mache gern Hausaufgaben.	Me gusta hacer la tarea domiciliaria.	Eu gosto de fazer lição de casa.
	Finds school useless (r)	I think school is useless.	Ich finde die Schule nutzlos.	Creo que estudiar no sirve para nada.	Eu acho que a escola não serve para nada.
School difficulties	Often has bad grades (r)	I make a lot of mistakes in my homework	Ich mache bei den Hausaufgaben viele Fehler.	Cometo muchos errores en la tarea/deberes que hago en casa	Eu erro muito nas lições de casa.
	Makes mistakes in homework (r)	I often have bad grades.	Ich habe oft schlechte Noten.	A menudo tengo malas notas	Eu frequentemente tiro notas ruins.
	Struggles to follow lessons (r)	I often struggle to follow what is going on in the lesson.	Ich habe oft Schwierigkeiten, dem Unterricht zu folgen.	A menudo tengo dificultad para seguir la clase	Eu frequentemente tenho dificuldades para acompanhar a aula.

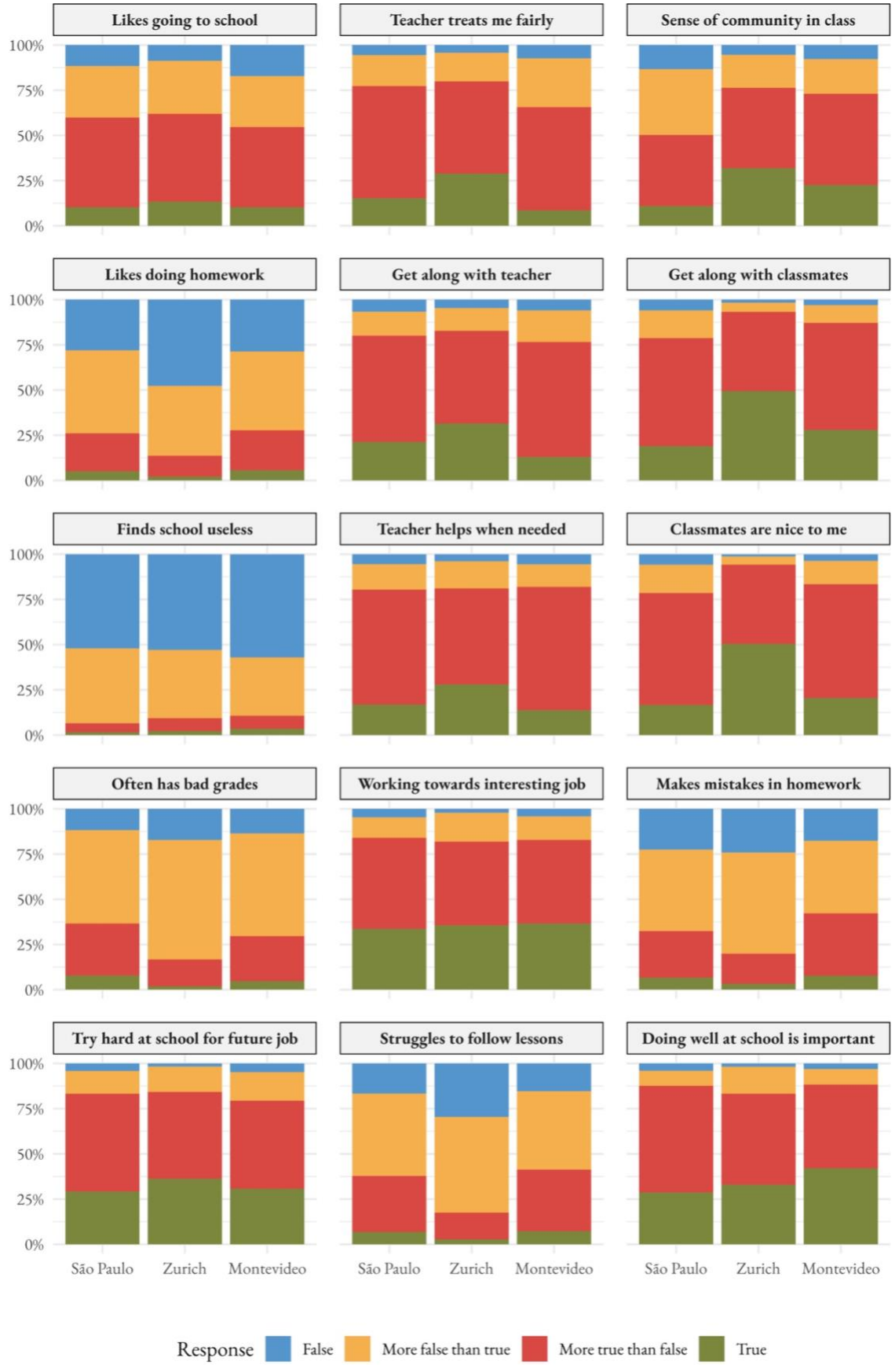
RESPONSE PATTERNS

The crude response patterns to the school experience scale revealed distinct patterns of endorsement (False, More false than true, More true than false, and True) across cities and dimensions of the scale (see Figure A4). For items related to bond to teacher and bond to class, (e.g. "Teacher treats me fairly", "Get along with teacher", "Sense of community in class"), students in all three cities showed predominantly positive responses. However, Zurich participants presented a higher proportion of "True" responses compared to São Paulo and Montevideo.

Items measuring future orientation showed strong positive endorsement across all cities. Notably, São Paulo and Montevideo students showed slightly stronger endorsement of these items than Zurich students, with higher proportions of "True" responses. For school difficulties items, which were negatively phrased, response patterns varied considerably. Zurich students reported fewer difficulties than São Paulo and Montevideo. The items from the school commitment dimension showed the greatest variability across cities. For "Likes doing homework", Zurich students showed distinctly lower enthusiasm compared to São Paulo and Montevideo. Conversely, on the negative item "Finds school useless", all three cities showed similar rejection rates of the sentence.

These response patterns highlight both cross-cultural commonalities and differences in how adolescents from these socioeconomically diverse urban contexts experience their school environments. Hence, our analysis underscores the importance of establishing measurement invariance before making substantive comparisons.

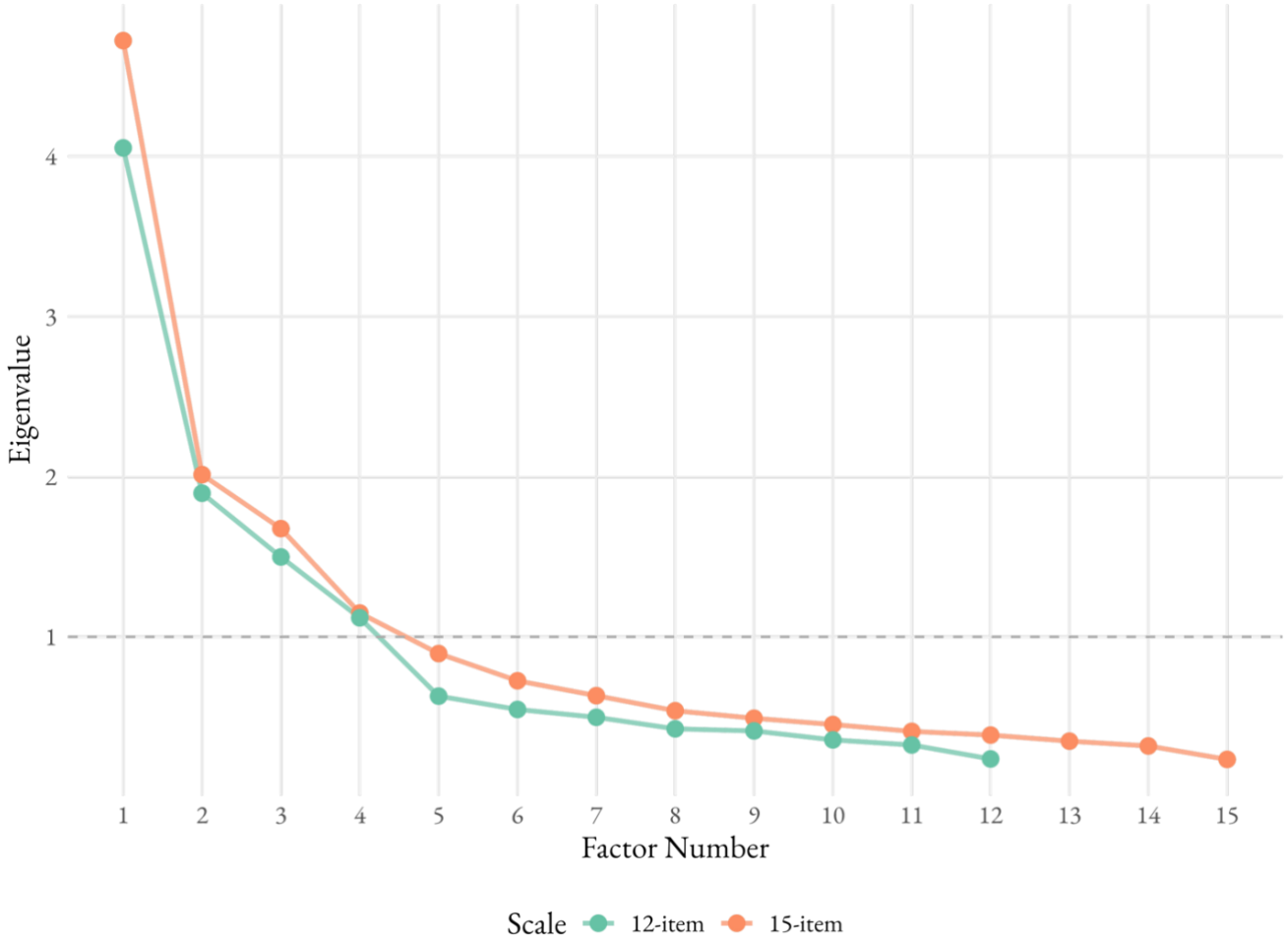
Figure A4: Proportion of response to items in the school experience scale by city.



EXPLORATORY FACTOR ANALYSIS

We conducted exploratory factor analysis (EFA) to examine the underlying factor structure of the school experience scale. Initial analyses were performed with all 15 items from five theoretical dimensions. Subsequently, we tested a reduced 12-item model omitting the "school commitment" items due to their less specific nature. The scree plot of eigenvalues (Figure A5) supported a multi-factorial solution, with pronounced breaks after the fourth factor in both versions. The first four eigenvalues for the 15-item scale (4.72, 2.01, 1.68, 1.15) and 12-item scale (4.05, 1.90, 1.50, 1.12) exceeded 1.0, while subsequent eigenvalues fell below this threshold, providing additional support for a four-factor solution according to Kaiser's criterion.

Figure A5: Scree plot with eigenvalues for the scale versions with 12 and 15 items.



The EFA revealed that a four-factor structure provided optimal fit for the 12-item school experience scale across the three sites (CFI = .981, RMSEA = .060), as shown in Table A7. While the five- and six-factor solutions demonstrated incrementally better fit indices, the four-factor solution offered the optimal balance between model parsimony and theoretical interpretability.

Table A7: Model fit indices for 15-item and 12-item EFA solutions.

Number of Factors	15-item Scale				12-item Scale			
	χ^2	df	CFI	RMSEA	χ^2	df	CFI	RMSEA
1-factor	17323	90	.508	.174	14573	54	.487	.207
2-factor	9808	76	.710	.146	8424	43	.692	.180
3-factor	4860	63	.842	.118	4078	33	.850	.143
4-factor	1475	51	.949	.075	320	24	.981	.060
5-factor	500	40	.976	.058	154	16	.990	.053
6-factor	256	30	.985	.052	56	9	.997	.039

The four-factor solution of the 12-item scale clearly represented the theoretically expected dimensions: bond to class (factor 1), bond to teacher (factor 2), school difficulties (factor 3), and future orientation (factor 4). All item loadings were strong (0.540-0.912) on their respective factors, with minimal cross-loadings. Table A8 presents a comparison of factor structures across different models. The five-factor solution for the 15-item scale demonstrated less clear factorial structures, with a distinct 5th factor emerging with strong loading from item "Likes going to school" but with item "Finds school useless" loading on the same factor as future orientation original items. In the four-factor solution, the school commitment items showed cross-loadings or merged with the future orientation factor, suggesting some conceptual overlap between these constructs.

Table A8: Item factor loading in four- and five-factor solutions with 15- and 12-item scales.

Item	15-item 5-factor					15-item 4-factor				12-item 4-factor			
	1st	2nd	3rd	4th	5th	1st	2nd	3rd	4th	1st	2nd	3rd	4th
Sense of community in class	0.683					0.667				0.681			
Get along with classmates	0.906					0.899				0.912			
Classmates are nice to me	0.803					0.802				0.795			
Teacher treats me fairly		0.732					0.798				0.784		
Get along with teacher		0.757					0.770				0.783		
Teacher helps when needed		0.540					0.528				0.540		
Likes going to school			1.012				0.388		0.307				
Likes doing homework			0.322						0.378				
Finds school useless (r)					0.368				0.394				
Often has bad grades (r)				0.614				0.616				0.618	
Makes mistakes in homework (r)				0.725				0.721				0.719	
Struggles to follow lessons (r)				0.767				0.749				0.749	
Working towards interesting job					0.686				0.668				0.675
Try hard at school for future job					0.830				0.813				0.864
Doing well at school is important					0.700				0.683				0.655

Note: Only factor loading > 0.3 are showed. (r) indicates reversed items.

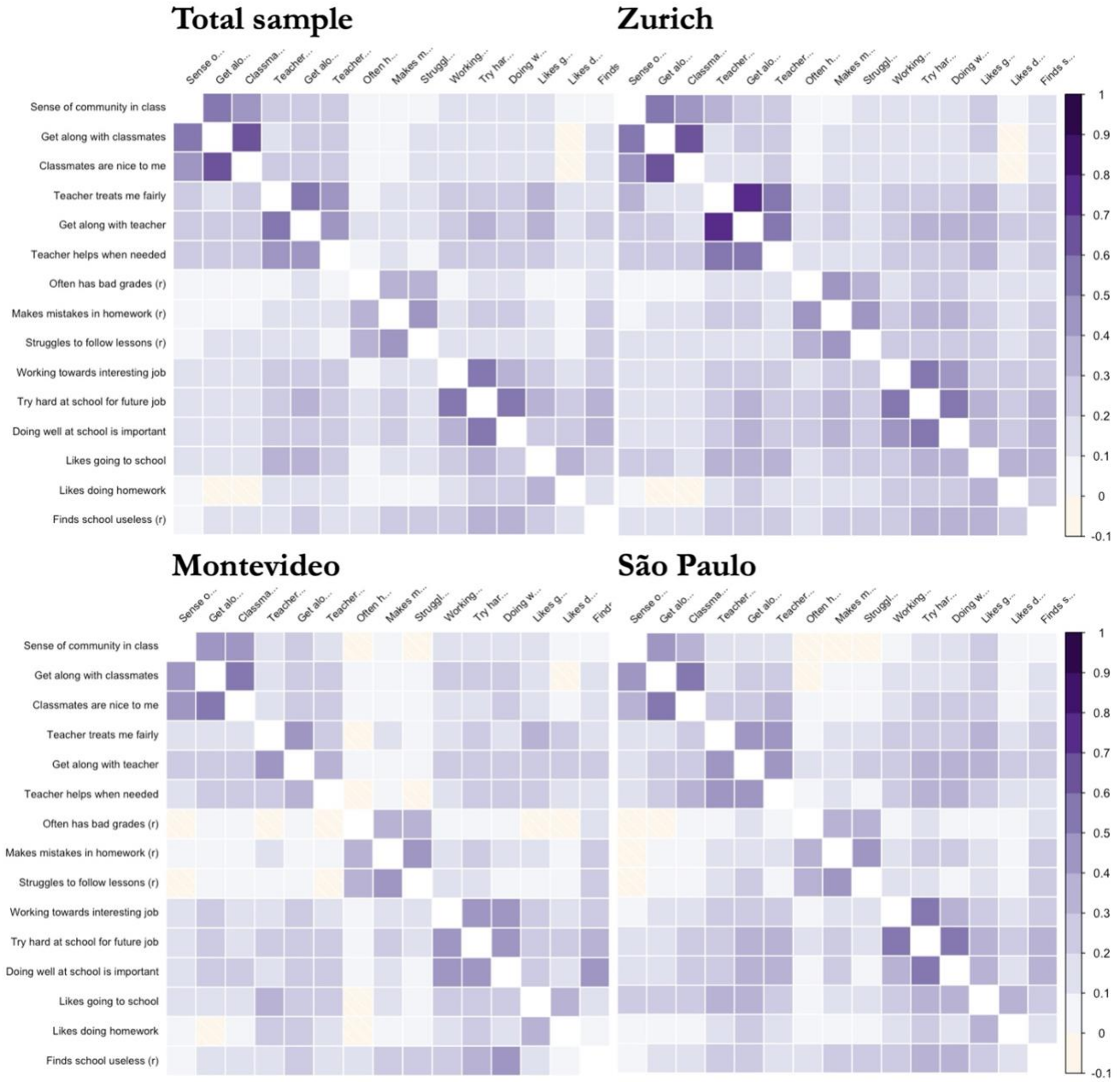
CORRELATION ANALYSIS

Correlation matrices

We examined the inter-item correlations to assess the relationships between items both within and across subscales (Figure A6). The correlation matrix for the total sample revealed moderate to strong correlations among items within the same theoretical dimension, supporting the predicted theoretical subscales (darker shades around the diagonal line). The highest within-subscale item correlations were observed in the bond to class dimension (r = .456-.617), future orientation (r = .395-.511), and bond to teacher (r = .410-.548). School difficulties items showed moderate correlations (r = .368-.453), while school commitment items demonstrated the weakest inter-item correlations (r = .127-.360), further indicating cohesion issues within this dimension.

Cross-subscale correlations were generally lower than within-subscale correlations. However, some notable patterns emerged: items from the school commitment dimension showed moderate correlations with items from subscales future orientation (r = .245-.355) and bond to teacher (r = .250-.356), suggesting conceptual overlap between these dimensions. The correlation patterns were largely consistent across cities, though Zurich showed the strongest within-subscale correlations overall. Particularly for the school difficulties and school commitment dimensions, Montevideo and São Paulo demonstrated more modest correlations. These findings suggest that while the underlying structure of school experience is similar across contexts, the strength of relationships between components may vary by cultural and educational setting.

Figure A6: Correlation between item response in the total sample and by city.



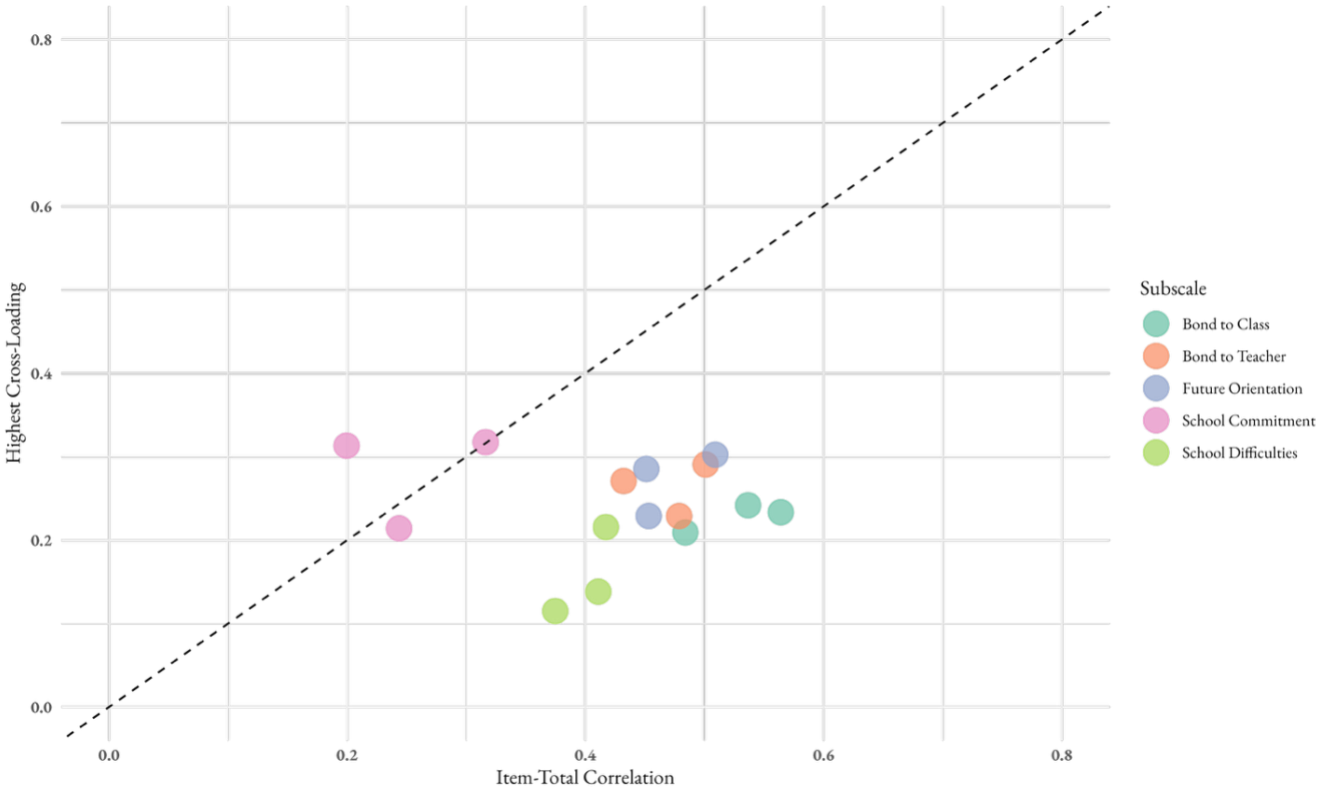
Discriminant validity

To assess the discriminant validity of scale items, we conducted a comparative analysis of item-total correlations and cross-loadings. This approach helps identify items that might be problematic due to stronger associations with other subscales than with their own. This analysis complements our exploratory factor analysis by providing an item-level assessment of the scale's dimensional structure based on correlation patterns. The approach is particularly useful for identifying items that might contribute to dimensional overlap or ambiguity in the measurement model.

First, item-total correlations were calculated for each item within its respective subscale. Item-total correlation represents the mean correlation between a given item and all other items in its intended subscale, indicating how well the item converges with its theoretical construct. Second, cross-loadings were calculated by determining the mean correlation of each item with items from other subscales. Cross-loading reflects the extent to which an item relates to constructs it was not designed to measure. For each item, we identified the highest cross-loading value (i.e., the strongest correlation with a competing subscale), representing potential discriminant validity issues.

Finally, we visually inspected discriminant validity by plotting each item's item-total correlation against its highest cross-loading value. In this visualization (Figure A7), items falling below the diagonal line demonstrate good discriminant validity (higher correlation with their own subscale than with other subscales), while items above the line indicate potential discriminant validity problems (stronger relationship with another subscale than with their own). The results indicated, as our former analyses, that school commitment items/subscale are less discriminant than all other items.

Figure A7: Item discriminant validity based on subscale cross-loading and total item correlation.



CRONBACH'S ALPHA RELIABILITY TEST

To assess the internal consistency reliability of the school experience scale and its dimensions, we calculated Cronbach's alpha coefficients for each subscale across the total sample and separately by city (Table A9). Cronbach's alpha measures the extent to which items within a scale consistently measure the same construct, with values above .7 generally considered acceptable. In the total sample, most subscales demonstrated satisfactory reliability: bond to class (α = .771), bond to teacher (α = .727), and future orientation (α = .728) all exceeded the recommended threshold of .70. The school difficulties subscale showed slightly lower but still adequate reliability (α = .668). However, the school commitment subscale exhibited poor internal consistency (α = .504), further supporting our decision to exclude it from subsequent analyses.

When examined by city, reliability coefficients showed some variation but followed similar patterns. Zurich typically demonstrated the highest reliability across all subscales (α = .592-.811), while Montevideo generally showed the lowest values (α = .448-.736). Despite these variations, the relative pattern of reliability across subscales remained consistent in all three cities, with school commitment consistently showing the lowest reliability. The overall scale reliability was good across all samples (α = .769-.828), indicating that the school experience scale as a whole demonstrates good internal consistency despite variations at the subscale level.

Table A9: Cronbach's alpha for each city and the total sample, by subscale.

Subscale	Total sample	Zurich	Montevideo	São Paulo
Bond to Class	.771	.785	.734	.736
Bond to Teacher	.727	.811	.636	.709
School Difficulties	.668	.671	.644	.659
Future Orientation	.728	.769	.715	.720
School Commitment	.504	.592	.448	.523
All items	**.797**	**.828**	**.769**	**.793**

THRESHOLD ANALYSIS

In measurement invariance testing with ordered categorical (ordinal) data, thresholds represent the points on the underlying continuous latent variable where respondents transition from one ordinal category to the next. These thresholds are critical for understanding response patterns across different groups and are the ordinal equivalent of item intercepts in continuous variable measurement models. In our case, each observed variable has three thresholds (t1, t2, and t3) corresponding to the transitions between the four response categories of our scale (False → More false than true → More true than false → True).

In our Figure A8 visualization, thresholds are displayed on the X-axis, while the Y-axis shows their estimated values. Lower threshold positions (negative values on the Y-axis) indicate that respondents need less positive school experience to select higher response categories, suggesting a tendency toward more positive responding. Conversely, higher threshold positions (more positive values on the Y-axis) indicate that respondents need stronger positive experiences to select more favorable response options, reflecting greater stringency in rating. The distances between thresholds reveal how distinctly respondents differentiate between adjacent response categories, with wider gaps suggesting more nuanced discrimination.

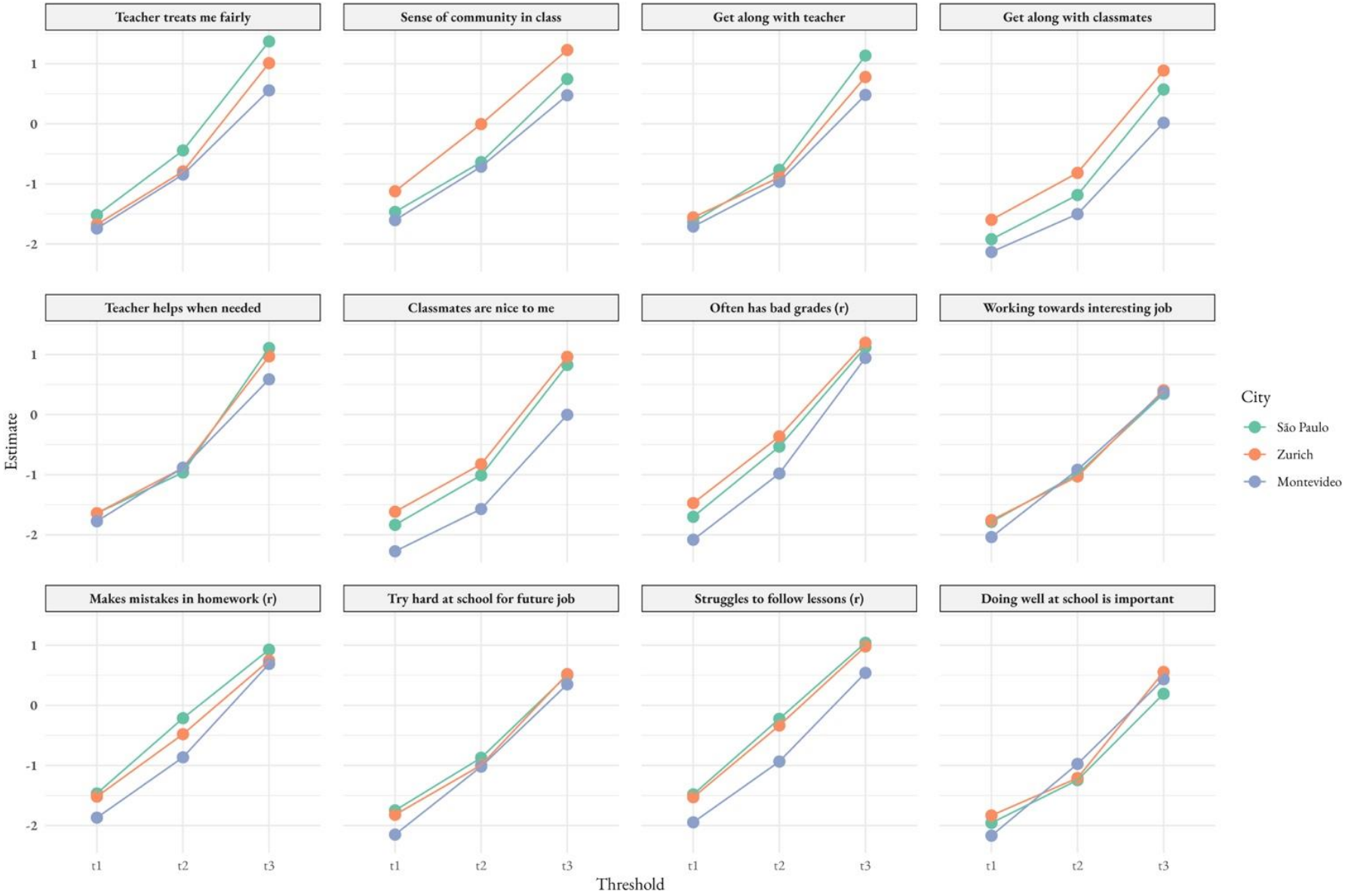
Examination of threshold patterns reveals city-specific response tendencies in how adolescents engage with the measurement scale. São Paulo respondents demonstrated more dispersed threshold distributions, with wider gaps between threshold points (particularly between t2 and t3), indicating greater discrimination between response categories and a more nuanced utilization of the full response scale. This pattern suggests São Paulo adolescents may employ more deliberate distinctions when evaluating their school experiences.

Zurich participants exhibited notably more symmetrical threshold distributions centered around zero, especially evident in bond to class and bond to teacher items. This symmetry indicates a balanced use of response categories without strong tendencies toward either extreme, possibly reflecting Swiss cultural norms favoring moderation in self-evaluation and a more calibrated approach to rating experiences. On the

other hand, Montevideo respondents displayed a distinct pattern with consistently more negative values for the first two thresholds (t1 and t2) across most items, coupled with smaller distances between thresholds. This compression toward the negative end of the latent continuum suggests students demonstrate a generalized tendency toward more positive self-reported school experiences, requiring relatively lower levels of the underlying constructs to endorse more positive response options. This pattern could reflect cultural response tendencies, social desirability effects, or genuinely more positive school experiences in the Uruguayan educational context.

Despite these interpretable cross-cultural variations in response patterns, our subsequent scalar invariance testing (which constrained thresholds to be equal across groups) showed acceptable fit, indicating that these differences did not substantially impact the overall measurement equivalence of the school experience scale across the three cities. For instance, future orientation items showed a pattern of increasingly negative thresholds in all three cities, with the final threshold (t3) being closest to zero. This indicates that students across all contexts required relatively lower levels of the underlying construct to endorse positive responses on these items, reflecting generally high future orientation across all samples.

Figure A8: Threshold distributions across cities for the school experience scale.



MISSING VALUES

To assess potential systematic patterns of non-response in the school experience scale, we conducted a comprehensive analysis of missing data. Understanding the extent and nature of missingness is crucial for determining appropriate analytical approaches and interpreting results. As shown in Table A10, missing data rates were consistently low across all items of the school experience scale, ranging from 1.04% ("Likes going to school") to 2.54% ("Struggles to follow lessons"). When examined by subscale, bond to teacher items showed slightly higher missingness (average 1.94%) compared to school commitment items (average 1.47%).

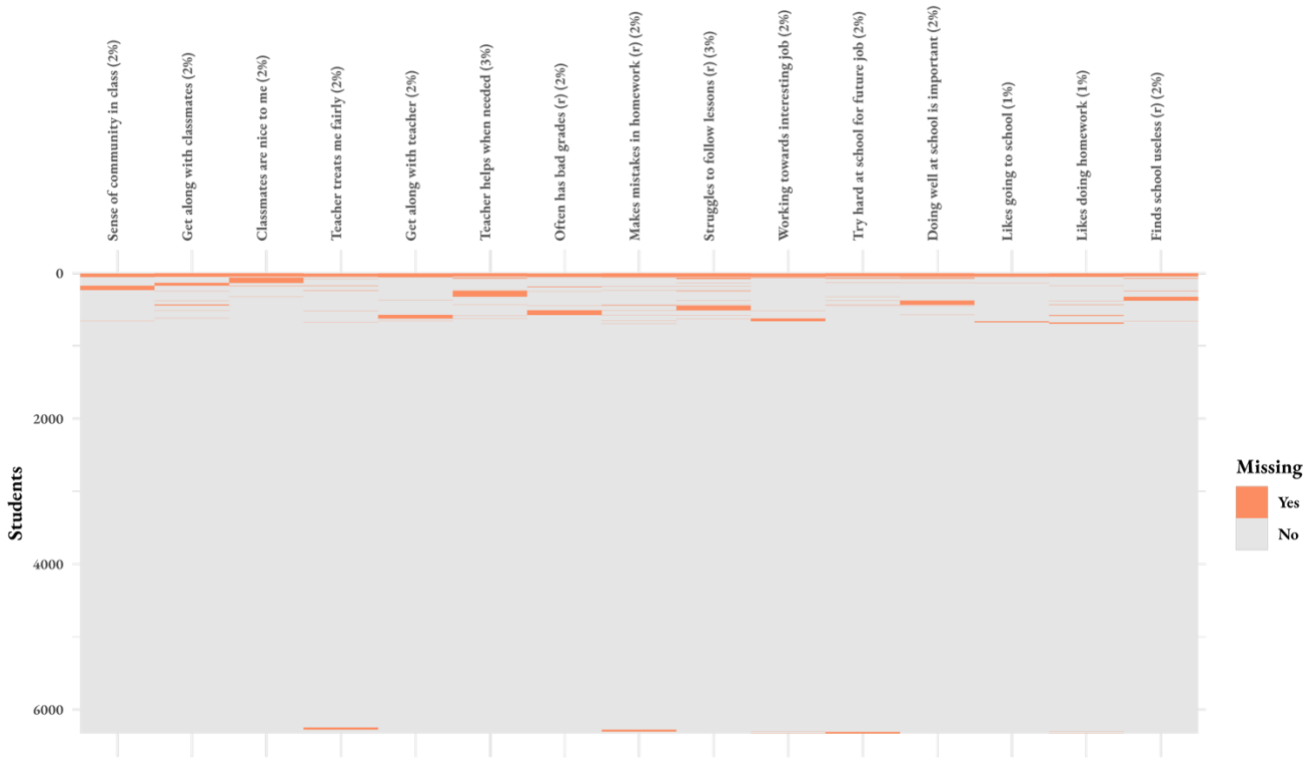
Table A10: Missing data by item and subscale.

Subscale	Item	Valid Cases (%)	Missing Cases (%)
Bond to Class	Classmates are nice to me	6,195 (97.85)	136 (2.15)
	Sense of community in class	6,215 (98.17)	116 (1.83)
	Get along with classmates	6,218 (98.22)	113 (1.78)
Bond to Teacher	Teacher helps when needed	6,171 (97.47)	160 (2.53)
	Get along with teacher	6,219 (98.23)	112 (1.77)
	Teacher treats me fairly	6,234 (98.47)	97 (1.53)
Future Orientation	Doing well at school is important	6,193 (97.82)	138 (2.18)
	Working towards interesting job	6,230 (98.40)	101 (1.60)
	Try hard at school for future job	6,230 (98.40)	101 (1.60)
School Difficulties	Struggles to follow lessons (r)	6,170 (97.46)	161 (2.54)
	Often has bad grades (r)	6,197 (97.88)	134 (2.12)
	Makes mistakes in homework (r)	6,219 (98.23)	112 (1.77)
School Commitment	Finds school useless (r)	6,205 (98.01)	126 (1.99)
	Likes doing homework	6,243 (98.61)	88 (1.39)
	Likes going to school	6,265 (98.96)	66 (1.04)
All items (complete cases)		**5,551 (87.68)**	**780 (12.32)**

Overall, the proportion of valid cases exceeded 97% for all items, indicating excellent data completeness. Although item-level missingness was minimal, complete-case analysis revealed that only 5,551 participants (87.68%) had responses for all 15 items simultaneously, with 12.32% missing at least one

item. This highlights how even small percentages of missing data at the item level can accumulate to more substantial missingness when considering all variables together. In Figure A9, we show how missingness is observed for each student.

Figure A9: Individual level missingness.

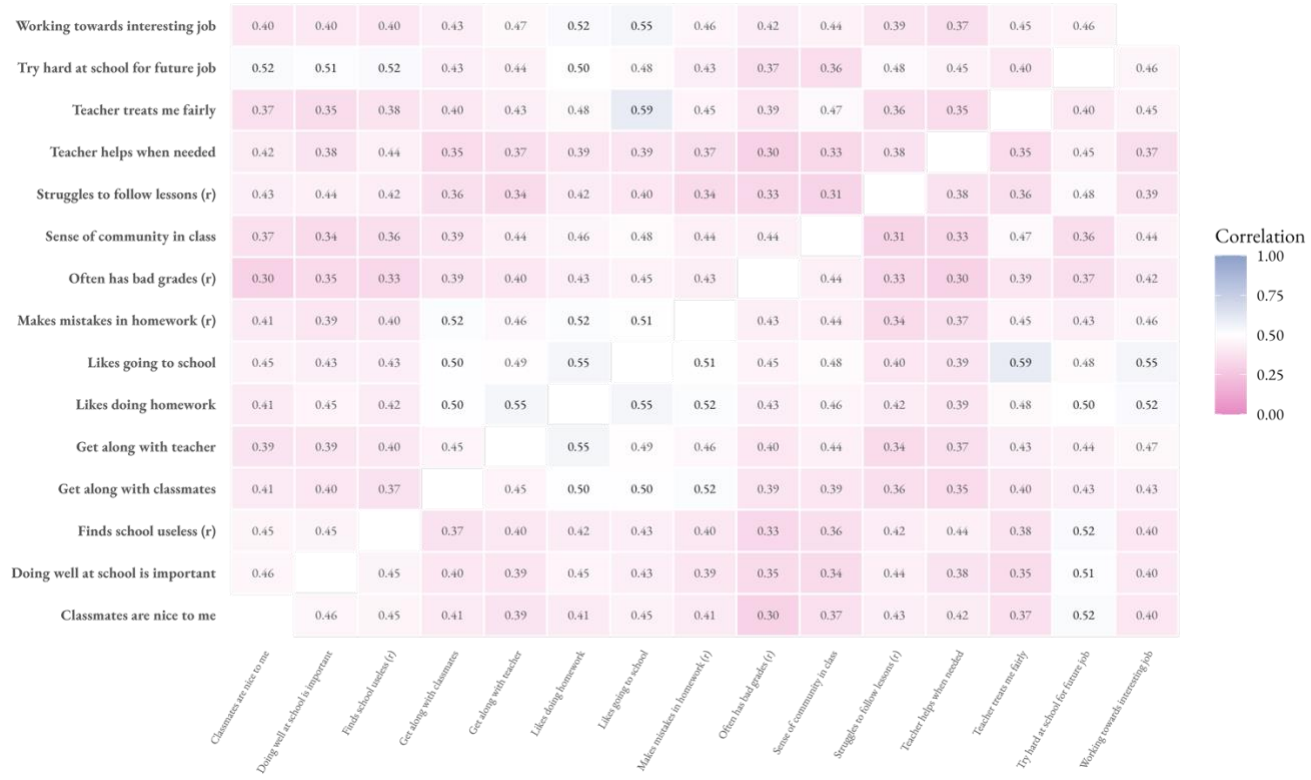


To examine potential systematic patterns in missing data, we conducted a correlation analysis of missingness indicators across all scale items (Figure A10). The correlation matrix of missing values revealed moderate correlations (r = .304-.595) among missingness indicators. The patterns observed indicate that when participants skipped one item, they were more likely to skip related items measuring similar constructs. Missingness correlation was particularly higher within school commitment (r = .418-.546) and future orientation (r = .404-.508) items.

This is further evidenced by the discrepancy between low item-level missingness (< 3%) and the higher proportion of incomplete cases overall (12.32%), suggesting that missing values tended to cluster within certain participants rather than being distributed randomly throughout the sample. Despite this clustering, no clear patterns of differential missingness were observed across cities or demographic

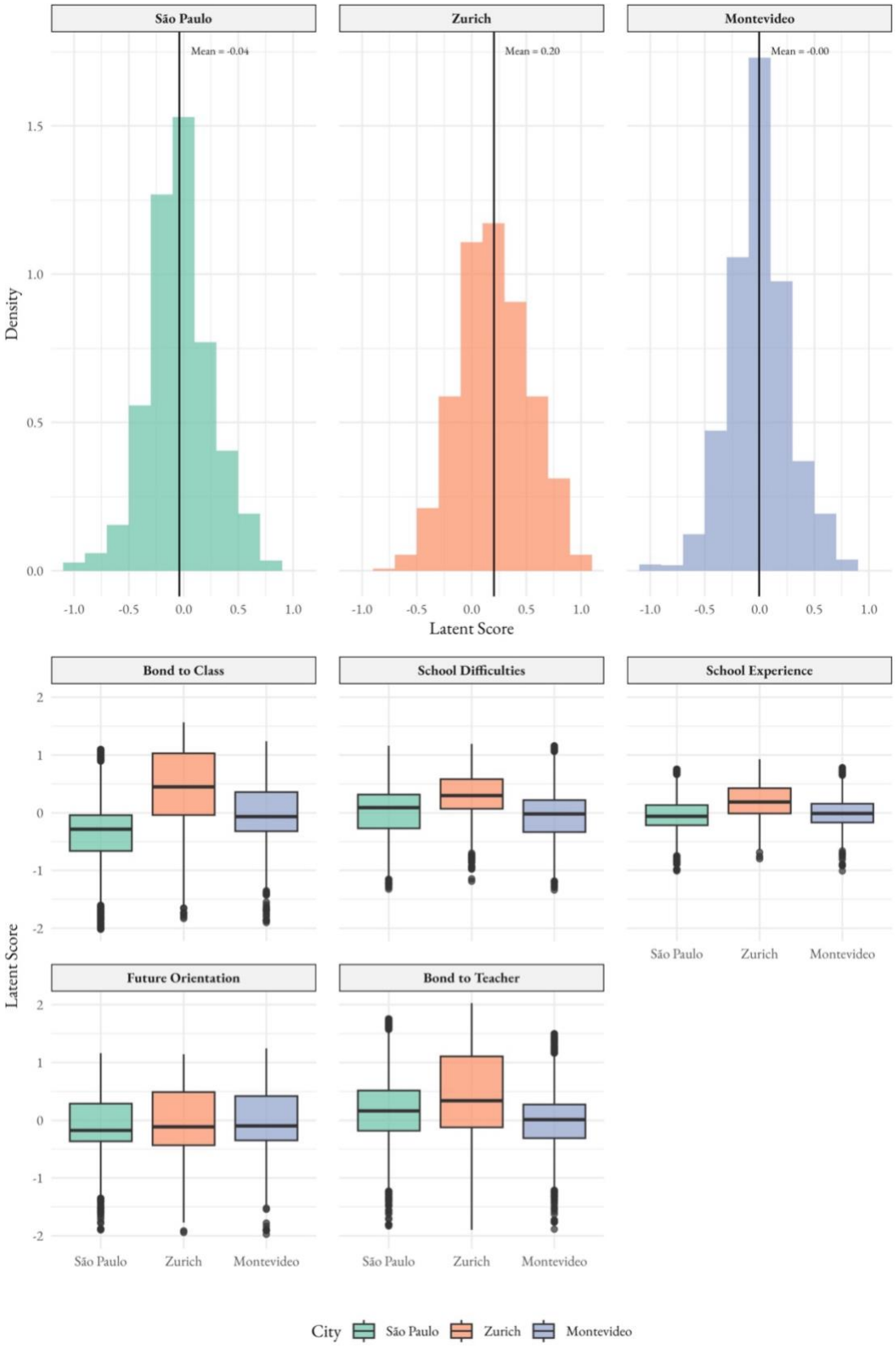
subgroups. Following methodological recommendations for handling ordinal data in confirmatory factor analysis [22, 23], we adopted a complete-case approach for our primary measurement models. Given that our complete sample still retained 87.68% of participants, we considered the risk of substantial bias as minimal due to this approach.

Figure A10: Correlation matrix for missing values between items.



ALTERNATIVE LATENT SCORES VISUALIZATION

Figure A11: Distribution of latent scores in the scalar model by city and scale dimension.



REFERENCES

1. World Bank, International Comparison Program. GDP per capita, PPP (current international $) - NY.GDP.PCAP.PP.CD. 2023.

2. World Bank. Human Capital Index (HCI) (scale 0-1) - HD.HCI.OVRL. 2020.

3. World Bank, Poverty and Inequality Platform. Gini index - SI.POV.GINI. 2023.

4. OECD. PISA 2022 results (volume I): the state of learning and equity in education. OECD; 2023.

5. OECD. Education at a glance 2023: OECD indicators. Paris: OECD; 2023.

6. UNESCO, Institute for Statistics. School enrollment, secondary, private (% of total secondary) - SE.SEC.PRIV.ZS. 2025.

7. UNESCO, Institute for Statistics. Pupil-teacher ratio, secondary - SE.SEC.ENRL.TC.ZS. 2020.

8. Schmutz R. Inequality of educational opportunity in Switzerland: exploring regional differences and institutional factors. Int J Educ Res. 2024;128:102495.

9. Edler S, Hadjar A. Educational trajectories within and beyond the core education phase in Switzerland: a sequence analysis based on SHP data 1999–2023. Front Sociol. 2025;10.

10. ANEP. Estado de Situación de la Educación. 2022.

11. Fernández Aguerre T, Cardozo Politi S, Marques Hill A. The education system of Uruguay. In: Jornitz S, Parreira Do Amaral M, editors. The Education Systems of the Americas. Cham: Springer International Publishing; 2022. p. 1–48.

12. Luquine Jr. CD, Peres MFT. Escolas, saúde e riscos na adolescência: Reflexões a partir do estudo sp-proso. Cad Pesqui. 2025;55:e11159.

13. Carnoy M, Marotta L, Louzano P, Khavenson T, Guimarães FRF, Carnauba F. Intranational comparative education: what state differences in student achievement can teach us about improving education—the case of Brazil. Comp Educ Rev. 2017;61:726–59.

14. Luschei TF, Carnoy M. Educational production and the distribution of teachers in Uruguay. Int J Educ Dev. 2010;30:169–81.

15. Felouzis G, Charmillot S. Tracking and educational inequality: a longitudinal analysis of two school reforms in Switzerland. Br J Sociol Educ. 2023;44:738–63.

16. Bogliaccini JA, Rodríguez F. Education system institutions and educational inequalities in Uruguay. CEPAL Rev. 2016;2015:85–99.

17. Wink Junior MV, Zanandréa Paese LH. Inequality of educational opportunities: evidence from Brazil. Economia. 2019;20:109–20.

18. Méndez L. University supply expansion and inequality of opportunity of access: the case of Uruguay. Educ Econ. 2020;28:115–35.

19. Ribeaud D, Murray A, Shanahan L, Shanahan MJ, Eisner M. Cohort Profile: The Zurich Project on the Social Development from Childhood to Adulthood (z-proso). J Dev Life-Course Criminol. 2022;8:151–71.

20. Trajtenberg N, Eisner M. Towards a more Effective Violence Prevention Policy in Uruguay. Montevideo: University of Cambridge, Universidade de la República, Administración Nacional de Educación Pública; 2015.

21. Peres MFT, Eisner M, Loch AP, Nascimento A, Papa CHG, Azeredo CM, et al. Violência, bullying e repercussões na saúde: resultados do Projeto São Paulo para o desenvolvimento social de crianças e adolescentes (SP-PROSO). São Paulo: Departamento de Medicina Preventiva/FMUSP; 2018.

22. Asparouhov T, Muthen B. Auxiliary Variables in Mixture Modeling: Using the BCH Method in Mplus to Estimate a Distal Outcome Model and an Arbitrary Secondary Model.

23. Wu H, Estabrook R. Identification of Confirmatory Factor Analysis Models of Different Levels of Invariance for Ordered Categorical Outcomes. Psychometrika. 2016;81:1014–45.