Early Educational Inequalities: Socio-Emotional and Cognitive Skill Gaps among Italian Primary School Children *

Daniela Piazzalunga[1], Águeda Solís Alonso[1], Sara Agostini[1], Gabriella Berloffa[1], Pietro Biroli[2], Maria Laura Di Tommaso[3], Laura Franchin[1], Silvia Mendolia[3], Luca Piccoli[1], Matteo Ploner[1], Lucia Savadori[1]

[1] *University of Trento* [2] *University of Bologna* [3] *University of Torino*

February 2026

Abstract

This paper investigates inequalities in cognitive and socio-emotional skills among Italian primary school children, with a focus on gender and migration background. Using original data from over 2,500 pupils in grades 3–5 across three regions, we measure mathematics performance, academic self-concept, growth mindset, prosocial behavior, socio-emotional skills, and gender norms. We document systematic gender gaps: girls outperform boys in socio-emotional and non-cognitive domains, while boys achieve higher scores in mathematics and math self-concept. Children with a migratory background perform worse in both cognitive and most non-cognitive outcomes. We further examine the role of one's own and teachers' gender-related attitudes and instructional practices. Children's stronger adherence to gender stereotypes is associated with worse non-cognitive skills, notably socio-emotional skills, growth mindset, academic self-concept, and mathematics performance, with more pronounced associations for boys. At the teachers' level, gender stereotypes are associated with lower self-efficacy and academic self-concept, particularly among girls, whereas more inclusive teaching practices are linked to higher mathematics achievement and improved socio-emotional outcomes. These associations are not explained by observable teacher characteristics. Overall, the findings highlight the importance of gender norms and teacher beliefs in shaping early educational inequalities and inform policies aimed at reducing skill gaps in primary education.

Keywords: socio-emotional skills, gender norms, education, inequality, children, teachers
JEL codes: I24, J24, J16, J15

* *Corresponding author*: Daniela Piazzalunga. e-mail: daniela.piazzalunga@unitn.it.
Funding: This paper is part of the GENESIS project and is funded by the European Union – Next Generation EU under the GRINS – Growing Resilient, Inclusive and Sustainable Project (GRINSPE00000018 – CUP E63C22002140007), through the *'Bando a Cascata'* "Gender Attitudes and Children's Socio-Emotional Skills (GENESIS)". The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, and the European Union cannot be held responsible for them.
Acknowledgements: GENESIS is conducted in coordination with the project "Tackling Inequalities in Educational Outcomes: Experimental Evidence from Italian Primary Schools" (https://cubeprogramme.unito.it/en/home-en/). Ethical approval was obtained from the Universities of Padova, Turin, and Trento. We thank all researchers involved, the field managers, enumerators, and the participating school principals, teachers, and pupils. We also thank participants of the "Socio-Emotional and Education" workshop (Trento, September 2025) and the GRINS Spoke 3 workshop (Naples, January 2026) for valuable comments. In drafting this document, generative artificial intelligence tools were used to enhance writing quality. All statements have been verified and processed by the authors, who assume full responsibility for the content of the document.

1 Introduction

Gender differences in educational outcomes remain pervasive across many countries and levels of schooling. Girls are more likely than boys to complete upper secondary education, enroll in and graduate from university, and they tend to achieve higher overall grades. At the same time, boys typically outperform girls in mathematics test scores and are overrepresented in math-intensive and STEM fields of study, patterns that persist across contexts and cohorts (Contini et al., 2017, Bertocchi & Bozzano, 2020, Encinas-Martín & Cherian, 2023). These early disparities matter: gender differences in educational performance and attainment translate into unequal labor market prospects in adulthood.

A growing body of research suggests that gender-specific stereotypes may contribute to the emergence of these educational gaps (see Bertrand, 2020). Stereotypical beliefs about boys' and girls' abilities, behaviors, and aspirations can shape expectations from parents, teachers, and peers, ultimately influencing children's academic engagement and performance. Understanding the mechanisms through which stereotypes affect learning is therefore crucial to explaining persistent gender disparities.

The economic literature has increasingly emphasized the role of non-cognitive skills (encompassing socio-emotional skills, personality traits, and preferences), in driving both school achievement and labor market outcomes (Brunello & Schlotter, 2011, Cunha et al., 2010, Deming, 2017, Heckman et al., 2006, Heckman & Kautz, 2012, Huggett et al., 2011). These skills, including perseverance, self-efficacy, empathy, and social behavior, are strong predictors of later life success. However, empirical evidence on gender differences in non-cognitive and socio-emotional skills remains mixed and often inconclusive, with substantial variation across contexts, age groups, and measurement tools (Hyde, 2005, 2014).

Socio-emotional skills are shaped by home, school, and social environments (Del Bono et al., 2016, Fiorini & Keane, 2014, Meroni et al., 2022, Carneiro et al., 2013), and they respond to investments and targeted interventions (Sorrenti et al., 2025, Kosse et al., 2020). Importantly, these skills can also be influenced by children's exposure to stereotypes, which may affect their beliefs, aspirations, and behaviors in ways that reinforce traditional gender roles and other types of prejudices (Alan et al., 2018, Alesina et al., 2024, Carlana, 2019).

In the Italian context, the evidence is particularly scarce. Despite longstanding gender gaps in educational outcomes, little is known about gender differences in socio-emotional skills among Italian children, a stage of development during which such skills are highly malleable. Italy is characterized by larger gender gaps in both educational and labour market outcomes compared to other OECD countries. It has the largest gender gap in math among 15-year-old students: according to PISA 2022, the mean score difference in math between boys and girls is 9 score points in OECD countries and 22 score points in Italy. Understanding the main causes of these differences in educational outcomes is thus crucial to designing effective policies and interventions to reduce adult gender gaps and related inequalities.

This paper contributes to filling this gap by examining gender differences in socio-emotional

skills among children in Italy and exploring how these differences relate to gender norms and stereotypes. Results indicate

Our analysis reveals systematic gender gaps in both cognitive and non-cognitive domains. Girls outperform boys in socio-emotional and other non-cognitive skills, whereas boys achieve higher scores in mathematics and report a stronger math self-concept. In addition, children with a migratory background perform worse in cognitive outcomes and in most non-cognitive dimensions, pointing to persistent inequalities early in primary school. We then investigate the role of children's own gender-related attitudes as well as teachers' gender attitudes and instructional practices. A stronger adherence to traditional gender stereotypes among children is associated with weaker non-cognitive skills, particularly socio-emotional skills, growth mindset, and academic self-concept, and with lower mathematics performance. These associations are especially pronounced for boys. Teachers' gender stereotypes are negatively related to students' self-efficacy and academic self-concept, particularly among girls. Instead, more inclusive teaching practices are associated with higher mathematics achievement and better socio-emotional outcomes.

The remainder of the paper is organized as follows. Section 2 describes the data collection process and sample selection, and defines the main variables. Section 3 examines unconditional and conditional differences in cognitive and non-cognitive skills, focusing on gender and migratory background and, to a lesser extent, regional differences, thereby providing evidence on skill gaps. Section 4 investigates the role of children's own gender attitudes, while Section 5 analyzes whether teachers' gender attitudes and self-reported practices are associated with students' development. Finally, Section 6 concludes.

2 Data and descriptive statistics

2.1 Data collection

This study draws on data collected within the broader research project "Tackling Inequalities in Educational Outcomes: Experimental Evidence from Italian Primary Schools" (*L'Unione Fa La Scuola*), implemented by researchers from the Universities of Bologna, Padova, Torino, and Trento, and carried out in the context of the GRINS (Growing, Resilient, Inclusive, and Sustainable) partnership. The broader project was designed to evaluate the effectiveness of two teacher professional development interventions aimed at promoting educational equity in primary education. In particular, the two interventions targeted teachers' pedagogical practices by addressing implicit biases in school environments and strengthening instructional quality in mathematics, with the broader objective of reducing disparities in educational outcomes.[1]

Within this framework, data collection involved both pupils and teachers in grades 3, 4, and 5 across three Italian regions (Campania, Friuli Venezia Giulia, and Piemonte) representing Southern, North-Eastern, and North-Western Italy, respectively. The data collection

[1] For more details, see the website: https://cubeprogramme.unito.it/en/home-en/

followed a pre- and post-intervention structure consistent with the randomized controlled trial (RCT) design of the original project, allowing for the measurement of baseline conditions and subsequent changes associated with the interventions.

The GENESIS research project, which constitutes the focus of this paper, builds on the data infrastructure developed in the broader initiative by extending children's data collection to additional domains, namely non-cognitive (socio-emotional) skills and gender norms. While GENESIS is embedded within the broader experimental framework, the present analysis relies exclusively on pre-treatment data collected in autumn 2024, prior to the implementation of the interventions, to avoid potential confounding effects from the treatment.

The analytical sample comprises pupils taught by teachers who voluntarily participated in the RCT. Teacher recruitment was coordinated at the regional level in collaboration with education authorities. In Spring 2024, the Regional Boards of Education (*Uffici Scolastici Regionali*) of Campania, Friuli Venezia Giulia, and Piemonte issued a formal call to all primary schools in their respective regions, providing information on the objectives, design, and requirements of the study. Schools were required to formally opt into the project by the end of the 2023–24 school year through an official approval by the School Governing Council (*Consiglio di Istituto*) in May 2024. This step ensured institutional authorization and formal commitment at the school level to participate in the research activities. Following approval, the project was incorporated into each school's School Educational Plan (*Piano dell'Offerta Formativa*), the official policy document outlining the educational initiatives and activities adopted by the school. In September 2024, minor adjustments to the teachers' sample occurred due to routine teacher reassignments across schools or grade levels, including movements of non-permanent teachers and grade changes among both permanent and non-permanent staff.

Data collection at the pupil level was conducted in Autumn 2024. Students enrolled in the classes of participating teachers completed the baseline (pre-test) survey during regular school hours. The survey was administered using a paper-and-pencil format, with trained enumerators present in classrooms to ensure standardized administration procedures and to provide assistance when needed.

The pupil questionnaire collected information on students' mathematics achievement, socio-emotional skills, and gender attitudes, as well as basic socio-demographic characteristics. The overall duration of the assessment did not exceed two hours and included a scheduled break to account for students' age and attention span. Given both the nature of the RCT interventions and the age of the pupils, slightly different versions of the questionnaire were administered to students in grade 3 and to those in grades 4 and 5. Specifically, children enrolled in grade 3 had a longer version of the mathematics test than children in grades 4 and 5, whereas the non-cognitive questionnaire administered to students in grades 4 and 5 included a larger number of items measuring the same underlying constructs, as well as a small set of additional questions that are analyzed separately. For the purposes of the present analysis, we restrict attention to items that are common across age groups, with the exception of the cognitive test as mentioned earlier, unless otherwise specified.

2.2 Measuring children's cognitive and non-cognitive skills

This section describes the children's skills measures used in this paper, capturing both cognitive and non-cognitive skills. Specifically, the non-cognitive domain includes children's academic self-perception, socio-emotional skills, prosocial behavior, gender norms, and beliefs about the malleability of abilities (growth mindset). Taken together, these measures provide a comprehensive assessment of both the cognitive and non-cognitive dimensions of children's development.[2]

We measure cognitive ability using students' performance on a standardized mathematics assessment. The test was designed to align conceptually with the national INVALSI framework. Students in grade 3 completed a more comprehensive assessment consisting of 27 questions (33 items), developed by experts in mathematics education and test design, and covering all major content domains. Students in grades 4 and 5 completed a shorter, age-appropriate version of the test. To account for differences in test design and cognitive maturity across grades, raw test scores are normalized within each grade cohort using grade-specific deciles for descriptive purposes and subsequently standardized within grade for the regression analysis.

To assess non-cognitive skills, we rely on validated scales from the educational and psychological literature, which were, where necessary, adapted to ensure suitability for the age group and context of the present study. Note that some indices are directly measured using Likert scales, while others are constructed as the sum or the average of underlying items or subscales. The choice of using four or five response options in the Likert scales, as well as whether to use sums or averages when constructing composite indices, follows the specifications of the original validated instruments.

Academic self-concept (*Autoefficacia scolastica percepita* (ASCP), based on Caprara & Gerbino, 2001) captures students' perceptions and evaluations of their own academic abilities. It is measured through a set of items assessing how competent children perceive themselves to be in six school subjects: mathematics, science, Italian, history, English, and physical education. For each subject, students are asked, "How good do you feel you are at . . . ". Responses are recorded on a Likert scale ranging from 1 ("not at all good") to 5 ("very good").[3] An overall academic self-concept score is constructed by summing responses across the six items, yielding a composite index ranging from 6 to 30. In addition, two subject-specific indices are derived for mathematics and Italian (each ranging from 1 to 5), allowing for a more focused analysis of children's confidence in core learning domains.

Socio-emotional skills are measured using the Social-Emotional Health Survey (SEHS) (Furlong et al., 2020). This instrument consists of nine items in which children report their level of agreement with statements capturing key socio-emotional dimensions. Responses are provided on a Likert scale ranging from 1 ("not at all true") to 4 ("very true"). Several sub-

[2] All questionnaires are available from the authors upon request.

[3] To facilitate children's responses, all Likert-scale items were presented using smiley faces.

scale scores are computed to capture specific dimensions of socio-emotional functioning, as well as an overall index:

- **Self-efficacy** measures children's beliefs in their ability to successfully perform specific tasks or achieve particular goals (range 1–4).[4]
- **Persistence** measures the tendency to sustain effort toward a goal despite difficulties or obstacles (range 1–4).
- **Emotional regulation** measures the ability to monitor, manage, and modify emotional responses in order to adapt to situational demands and pursue goals effectively (range 1–4).
- **Empathy** measures the ability to understand and emotionally respond to the feelings and perspectives of others (range 1–4).
- **Self-confidence** is constructed as the average of self-efficacy and persistence (range 1–4).
- **Emotional competence** is constructed as the average of emotional regulation and empathy (range 1–4).
- **Socio-Emotional Health (overall)** is the composite index, constructed as the average of responses across all items (range 1–4).

Prosocial behavior captures children's tendency to engage in voluntary actions that benefit others and support positive social relationships. It is measured using the Positive Experience at School Scale (PEASS) (Furlong et al., 2013). This instrument consists of five items assessing how frequently children display prosocial behaviors in the school context, including rule-following, cooperation with teachers, helpfulness toward peers, and attentiveness to others' needs. Responses are recorded on a Likert scale ranging from 1 ("almost never") to 4 ("very often/always"). A composite prosocial behavior index is constructed by summing responses across the five items, yielding a total score ranging from 5 to 20.

The **Growth Mindset Scale for Children (GM-C)** measures children's beliefs about the malleability of intelligence and abilities, as opposed to viewing them as fixed traits. For this project, we constructed an adapted measure based on two short scenarios (four items in total), following Muradoglu et al. (2022). In each scenario, a student who initially struggles in a subject (Italian or mathematics) changes schools and receives additional instruction in that subject.[5] Children are asked to indicate how skilled the student might become, choosing among four response options ranging from "not skilled at all" to "very skilled". Responses are converted into numerical scores (0, 0.33, 0.66, and 1) and averaged to construct an index

[4] Although related, academic self-concept and self-efficacy capture distinct constructs. Academic self-concept refers to perceived competence in specific school subjects and is assessed using a single item per subject ("How good do you feel you are at ..."). Self-efficacy refers to beliefs about successfully performing specific tasks ("Can I do this task successfully?") and is measured through multiple items.

[5] Compared to the original version of the scale, which included six response categories, the adaptation used in this project simplifies the scale to four categories while preserving the underlying conceptual framework.

reflecting growth mindset, defined as the belief that abilities can improve through effort and learning.[6]

The measures described above refer to children's cognitive and non-cognitive outcomes. In addition, we collect information on children's gender norms, which is one of the explanatory variables of interest. The measure of children's adherence to **gender norms** is based on an adaptation of the COAT-AM (Children Occupations, Activities and Traits) module (Liben et al., 2002). This measure explores how children associate certain activities with either the male or female gender, including examples of activities typically perceived as masculine (e.g., "playing video games", "being good at mathematics"), feminine (e.g., "cooking", "taking care of children"), and gender-neutral (e.g., "being creative", "studying hard"). In this project, an adaptation of the original COAT-AM module was employed, updating a few activities.[7] Scores are computed by comparing children's responses with a predefined classification reflecting common gender stereotypes. Two main indices are derived from this analysis:

- Stereotype Adherence Index, representing the percentage of responses consistent with traditional gender-role classifications
- Flexibility Index, calculated as the proportion of "both boys and girls" responses, indicating the ability to move beyond rigid gender associations.

A higher score on the first index suggests stronger stereotyping, whereas a higher score on the second reflects a more flexible and less conventional view of gender roles. Given that the two indices are highly correlated, in the analysis we only considered gender norms conformity (i.e., Stereotype Adherence Index), also referred to as 'gender attitudes' in the text. Note that children's gender attitudes are used as explanatory variables of interest in Section 4.

While the descriptive analysis relies on raw scores, all outcome variables are subsequently standardized using the full-sample mean and standard deviation in the regression analysis in order to facilitate the interpretation of estimated effects.

Table 1 provides a brief summary of the definition, ranges, and references of the main cognitive and non-cognitive skills, as well as a short description of other socio-demographic variables used in the analysis, based on the children's questionnaire.

[6] Also for the growth mindset scale, originally we constructed two indices, but we then used only for for ease of interpretation.

[7] In the questionnaire, we also collect information on personality traits associated with the male or female gender, which are not included in this paper for the sake of brevity.

Table 1: Definitions of Key Variables

Variable	Definition	Reference	Min	Max
Math test score	Index of the child's skills in math (expressed in grade-specific deciles)			
Academic self-concept				
Academic self-concept (Italian)	Index of child's perceived competence in Italian		1	5
Academic self-concept (Maths)	Index of child's perceived competence in Math	ASCP (Caprara & Gerbino, 2001)	1	5
Academic self-concept (overall)	Index of child's perceived academic competence in Italian, Math, Science, History, English, P.E.		6	30
Socio-emotional health				
Empathy	Index of the capacity to understand and respond to the feelings and experiences of others		1	4
Emotional regulation	Index of the ability to manage own emotions (handling mistakes, accepting limits)		1	4
Persistence	Index of the tendency to continue working toward solving problems despite difficulty		1	4
Self-efficacy	Index of belief in own ability to successfully handle challenges and accomplish tasks	SEHS (Furlong et al., 2020)	1	4
Self-confidence	Average of persistence and self-efficacy		1	4
Emotional competence	Average of empathy and emotional regulation		1	4
Socio-Emotional Health (overall)	Average of empathy, emotional regulation, persistence and self-efficacy		1	4
Pro-social behavior	Index of the frequency of engagement with prosocial behaviors at school	PEASS (Furlong et al., 2013)	5	20
Growth mindset	Index of the belief that intelligence is a malleable quality	GM-C (Muradoglu et al., 2022)	0	1
Gender norms conformity	Index of the child's agreement with traditional gender-role stereotyping of activities	COAT-AM (Liben et al., 2002)	0	1
Socio-demographic variables				
Female	Dummy variable. 1 female, 0 otherwise (male and "prefer not to answer")			
Foreign	Dummy variable. 1 if at least one parent is foreign; 0 otherwise			
Mum employed	Dummy variable. 1 if mother works; 0 otherwise (included 'don't know')			
Dad employed	Dummy variable. 1 if father works; 0 otherwise (included 'don't know')			
Siblings				
Only child	Categorical variables based on the number of brothers and sisters collected in the survey.			
One sibling				
Two siblings				
Three or more siblings				

Note: Table includes a short description and the main measure of reference for the key variables used in this study, together with the range of possible values. P.E. stands for Physical Education.

2.3 Teachers' variables: attitudes and teaching practices

In the final section of the paper (Section 5), we assess the association between teachers' gender attitudes, their teaching practices, and children's outcomes.

Gender Stereotypes in Teaching (Della Giusta & Muratori, 2022) is a measure of own teacher's adherence to gendered beliefs about abilities and behaviors. It includes items asking teachers to attribute abilities or behaviors to students based on gender (e.g., predisposition to STEM subjects, humanities, or classroom behavior). Higher values indicate responses that are more conforming to gender stereotypes.

Inclusive Teaching Practices measures the extent to which teachers consider diverse pedagogical strategies (exams, collaborative teaching, activities to develop critical thinking, etc.) as inclusive (Ewing et al., 2018).

For both measures, the question-based instruments are aggregated using factor analysis, resulting in standardized indexes with a zero mean and unit variance. This technique is employed to combine multiple related survey items into a smaller set of underlying dimensions, reducing measurement noise and capturing the core latent constructs reflected in teachers' gender attitudes.

Additionally, for the subset of mathematics teachers in grade 3, we collected a questionnaire developed by Swan (2006) to measure their beliefs and teaching practices. **Beliefs** or orientation scales relate to teachers' beliefs about the nature of mathematics, learning, and teaching. In these questions, respondents are asked to allocate 100 points across three different options, according to their agreement with each statement. Each option is associated with one of three orientations: transmission, discovery, or connectionist. The final scales are calculated by averaging the points allocated to each orientation across the questions, and range between 0 and 100.

- The **transmission** orientation is teacher-centered, viewing mathematics as a fixed body of knowledge to be delivered efficiently. Teachers in this mode focus on explanation, demonstration, and procedural practice, while students primarily listen, imitate, and reproduce correct methods. Learning is measured by accuracy and mastery of procedures, and classroom interactions are largely controlled by the teacher.
- The **discovery** orientation emphasizes students' individual exploration of mathematical ideas. Here, learning is seen as an active process in which students experiment, reflect, and uncover principles for themselves, with the teacher supporting rather than directing.
- The **connectionist** orientation extends this further, highlighting the social and dialogic construction of knowledge. Students are encouraged to discuss, reason, and make connections between ideas, while teachers facilitate collaboration, pose challenging problems, and guide discussions. Both discovery and connectionist approaches are considered student-centered, but the connectionist orientation particularly values relational understanding and the co-construction of meaning.

For measuring **teaching practices**, also based on Swan (2006), we rely on items that quantify the frequency with which teachers engage in different practices and behaviors while teaching. Answers are summed, creating a scale for student-centered practices and one for teacher-centered practices. The student-centered practices are then reversed and aggregated to the teacher-centered practices scale, resulting in a unique measure, which we refer to as **Teacher-centered practices** scale. A higher number indicates that the way of teaching is more teacher-centered.

2.4 Sample selection and descriptive statistics

Table 2 summarizes the sample selection process by region, grade, and overall. The initial sample consists of all pupils enrolled in participating classes whose parents provided consent to take part in the study. The final sample is obtained by restricting the initial sample to pupils who were present in class on the day of data collection. As shown in the table, sample selection between the initial and final samples is limited across all regions and grades, likely due to absences related to illness. Overall, the final sample includes 2,538 pupils (from an initial sample of 2,772 children), corresponding to an absence rate of approximately 8.4%, with no major differences across regions or grades. Absence rates range from approximately 3% to 13%, with higher selection observed in grade 3 in Campania and in grade 5 in Friuli Venezia Giulia. Given that selection is driven primarily by pupil absence on the day of in-class data collection, these rates are consistent with expectations for school-based surveys, reflecting classical pupil absences, and a priori do not raise concerns about selective dropout. When considering regional differences, absence rates remain modest and range from 6.1% in Friuli Venezia Giulia to 7.1% in Piemonte and 9.7% in Campania.

Table 3 provides an overview of the final sample composition in terms of participating schools (*Istituto comprensivo*), school sites (*plesso*), classes, and pupils, also disaggregated by region and grade. The final sample covers 41 schools (17 in Piemonte, 18 in Campania, and 6 in Friuli Venezia Giulia - not shown) and 68 school sites across the three regions, for a total of 182 classes overall. Note that because some schools and school sites participated with classes in more than one grade, the total number of schools and school sites does not equal the sum across grades.

It can be observed that the majority of participating schools and pupils are located in Campania, followed by Piemonte, while Friuli Venezia Giulia contributes a smaller number of schools and pupils. This distribution reflects regional population size, with approximately 5.6 million inhabitants in Campania, 4.3 million in Piemonte, and 1.2 million in Friuli Venezia Giulia.

Table 4 reports descriptive statistics for key children's socio-demographic characteristics, cognitive outcomes, and non-cognitive measures for the final sample, presented separately by region and for the full sample. The sample is balanced by gender, with approximately half of the pupils being female. In terms of family structure, about half of the children have one

Table 2: Children's sample selection

	Grade							
	3		4		5		Total	
Region	Initial	Final	Initial	Final	Initial	Final	Initial	Final
Campania	728	639	431	398	433	410	1602	1,447
Friuli-Venezia Giulia	218	207	105	100	55	48	378	355
Piedmont	563	518	134	126	95	92	792	736
Total	1,509	1,375	670	624	593	550	2,772	**2,538**

Note: "Initial" reports the number of enrolled students in participating classes whose parents/guardians provided consent. "Final" reports the subset of consented students who were present on the day of data collection and are included in the analysis.

Table 3: Final sample, by region and grade

Region	Campania			Friuli-Venezia Giulia			Piedmont			Total
Grade	3	4	5	3	4	5	3	4	5	
Schools	16	13	12	6	3	1	16	7	4	41
School sites	25	17	17	8	5	3	23	7	5	68
Classes	48	30	33	13	6	4	33	9	6	182
Children	639	398	410	207	100	48	518	126	92	2,538

Note: Some schools (and school sites) participated with classes in more than one grade. Therefore, the figures in the Total column for Schools and School sites report the number of unique participating schools/sites and do not equal the sum across grades. Totals for Classes and Children are additive.

sibling, while 17% are only children. Overall, 20% of pupils have at least one foreign parent. Nearly all children have a working father, and 74% have a working mother, with substantial regional variation in family background characteristics. The share of children with at least one foreign parent is higher in Piemonte (32%) and Friuli Venezia Giulia (26%) than in Campania (13%). Similarly, maternal employment rates are considerably higher in the Northern regions (85% in Piemonte and 88% in Friuli Venezia Giulia) compared to Campania (65%), while paternal employment rates are broadly similar across regions. Although the sample is not designed to be representative at the regional level, these patterns are consistent with broader regional differences.

With respect to outcomes, average mathematics performance, measured using grade-specific deciles, ranges from 1 to 10, with 5 indicating the grade-level median performance. It presents higher mean scores observed in Friuli Venezia Giulia (5.93) and Piemonte (5.40) relative to Campania (4.64). Across all non-cognitive outcomes, students exhibit relatively high levels, generally above the scale midpoints. In particular, scores are highest for prosocial behavior at school, academic self-concept, and self-efficacy, while emotional regulation is comparatively lower, though still at a high level. Growth mindset and gender norms conformity are not directly comparable with the other outcomes.[8] Modest regional differences also emerge in non-cognitive dimensions, in addition to cognitive skills. This will be further explored in section 3.4.

The descriptive evidence presented in Table 4 provides an initial overview of the distribution of children's socio-demographic characteristics and skills. While these patterns are informative, they do not shed light on the underlying sources of these differences, nor do they account for potential confounding factors or compositional effects. In the following section, we examine inequalities in cognitive and non-cognitive outcomes by gender and migratory background, considering both raw differences and estimates from multivariate regression models that control for individual, family, and school characteristics.

Before proceeding to the next section, we present the teacher sample selection process and the main summary statistics. Table 6 first summarizes the sample selection process for teachers. Starting from an initial sample of 189 teachers, the sample was slightly reduced due to accurate matching with classes and teachers' participation in the pre-intervention questionnaire. The final teacher sample consists of 177 individuals.

Table 6 also reports descriptive statistics for teachers in the final sample. The large majority of teachers are female, and their average age is 52 years. Fifty-five percent hold a university degree, reflecting changes in teaching requirements,[9] as this qualification became compulsory

[8] Non-cognitive skill measures are based on different constructs and scales. To facilitate descriptive comparisons across outcomes measured on different scales, we rescale each outcome using its minimum and maximum possible values and express scores as percentages of the scale range (scores not shown), according to the following transformation: $Y_j^{\text{pct}} = (Y_j - \min(Y_j))/(\max(Y_j) - \min(Y_j))$. These rescaled values are used for descriptive purposes only; all regression analyses rely on outcome-specific standardization. The growth mindset and gender norms conformity indices are constructed using methodologies that do not rely on summing Likert-scale responses; therefore, even this form of raw comparison is not appropriate for these outcomes.

[9] A university degree has been mandatory for teaching only since 2001.

only in 2001. Teachers seem to have more 'teacher-centered' practices in Campania and Friuli Venezia Giulia (mirrored by the higher average of Transmission orientation in Campania than in other regions).[10]

Table 4: Children's Descriptive Statistics

	Campania		Friuli-VG		Piedmont		Total	
Variable	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev
Socio-demographic								
Female	0.49		0.52		0.48		0.49	
Foreign	0.13		0.26		0.32		0.20	
Only child	0.15		0.21		0.20		0.17	
One sibling	0.52		0.53		0.49		0.51	
Two siblings	0.21		0.17		0.19		0.20	
Three or more siblings	0.13		0.09		0.12		0.12	
Mum works	0.65		0.88		0.85		0.74	
Dad works	0.97		0.98		0.96		0.97	
Cognitive skills								
Math test: schoolyear decile [1-10]	4.64	2.81	5.93	2.79	5.40	2.88	5.03	2.87
Non-cognitive skills								
Growth mindset [0-1]	0.63	0.15	0.61	0.18	0.62	0.16	0.63	0.16
Gender norms comformity [0-1]	0.36	0.26	0.36	0.26	0.34	0.26	0.35	0.26
Pro-social behavior at school [5-20]	17.53	2.32	17.08	2.48	17.37	2.36	17.42	2.36
Socio-emotional skills								
Empathy [1-4]	3.12	0.94	3.32	0.81	3.17	0.92	3.16	0.92
Emotional regulation [1-4]	3.00	0.73	2.98	0.73	2.98	0.74	2.99	0.73
Persistence [1-4]	3.16	0.57	3.14	0.60	3.13	0.58	3.15	0.58
Self-efficacy [1-4]	3.21	0.48	3.23	0.53	3.20	0.50	3.21	0.49
Emotional competence (emot. reg. + empathy) [1-4]	3.06	0.64	3.15	0.58	3.08	0.65	3.08	0.64
Self-confidence (self-efficacy + persistence) [1-4]	3.19	0.43	3.19	0.47	3.16	0.45	3.18	0.44
Average [1-4]	3.13	0.45	3.17	0.45	3.12	0.47	3.13	0.46
Academic self-concept								
Overall [6-30]	23.82	3.28	23.72	3.30	23.69	3.23	23.77	3.27
Italian [1-5]	4.09	0.88	3.90	1.01	3.87	0.96	4.00	0.93
Maths [1-5]	3.99	0.89	4.01	0.90	4.02	0.90	4.00	0.90
Number of observations	**1,447**		**355**		**736**		**2,538**	

Note: Note: A child is classified as "Foreign" if at least one parent is foreign-born. Cognitive skills are measured using the mathematics test administered during data collection. The growth mindset scale is based on Muradoglu et al. (2022) ; the gender-norm conformity index on Liben et al. (2002); socio-emotional skills on the Socio-Emotional Health Survey (SEHS) (Furlong et al., 2020); and academic self-concept on the Autoefficacia scolastica percepita (ASCP) scale (Caprara & Gerbino, 2001). Additional details on the measures and response scales are reported in Table 1. Standard deviations are not reported for dummy variables. For cognitive and non-cognitive skill measures, ranges are reported in brackets.

[10]We also investigated whether teachers' gender attitudes and teaching practices varies by observable characteristics, but differences in mean values across age groups and between teachers with and without a university degree are generally small and statistically insignificant, suggesting that gender-related attitudes and practices are not strongly stratified along these dimensions.

Table 5: Teachers' sample selection

	N. observations
Initial sample	189
Matched to a class	182
Matched + answered survey	177

Note: Note: "Initial sample" refers to the total number of teachers who initially agreed to participate in the intervention. "Matched to a class" refers to the subset of these teachers whose students completed the assessment and could be linked to the teacher. The final row reports the number of teachers who, in addition to participating and being successfully linked to student records, completed the pre-intervention teacher survey.

Table 6: Descriptive Statistics - Teachers

	Campania		Friuli		Piemonte		Total	
Variable	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev
Female	0.91		0.90		1.00		0.94	
Age	54.12	6.16	51.30	10.08	48.54	8.59	51.62	8.16
Married	0.85		0.81		0.88		0.85	
N. of children below 5	0.12	0.51	0.05	0.22	0.04	0.20	0.08	0.38
N. of children 5-18	0.47	0.79	0.43	0.68	0.72	0.86	0.55	0.80
N. of children above 18	1.24	1.02	0.71	1.06	0.84	0.91	1.01	1.01
Has a university degree	0.56		0.48		0.56	0.50	0.55	
Permanent position	1.00		0.95		0.90		0.96	
Years of experience teaching in primary	24.27	13.02	25.05	10.91	19.54	11.60	22.66	12.36
Years of experience teaching in current school	14.44	14.62	13.10	10.31	14.38	12.28	14.21	13.13
Inclusive teaching practices (factor) [0-1]	-0.06	0.82	0.09	0.62	0.23	0.80	0.04	0.79
Gender stereotypes in teaching (factor) [0-1]	0.09	0.92	0.02	0.83	-0.27	0.65	-0.05	0.82
Number of observations	**97**		**27**		**76**		**177**	
Teacher-centered practices* [25-125]	72.14	10.07	70.45	10.03	63.26	9.45	68.19	10.57
Transmission orientation* [0-100]	22.59	16.56	18.48	15.03	16.52	12.13	19.49	14.74
Discovery orientation* [0-100]	46.99	17.70	44.85	13.24	42.55	12.30	44.84	15.02
Connectionist orientation* [0-100]	30.42	12.44	36.67	10.33	40.93	13.28	35.68	13.35
Number of observations	**36**		**11**		**34**		**81**	

Note: Inclusive Teaching Practices is based on Ewing et al. (2018) and Gender Stereotypes in Teaching is based on Della Giusta & Muratori (2022). Both indices are constructed using factor analysis and hence are standardized with a zero mean and a standard deviation equal to one. *Teacher-centered practices scale and the Transmission, Discovery, and Connectionist orientation scales are based on Swan (2006), and are only available for the subset of 3rd-grade teachers. Standard deviations are not reported for dummy variables. For teachers' attitudes and beliefs, the range of measures' values is reported in square brackets.

3 Inequalities in children's socio-emotional skills

To provide an initial picture of inequality in skills, we present descriptive patterns in cognitive and non-cognitive outcomes across gender and migratory background. Figures 1–6 report unadjusted mean differences with 95% confidence intervals. We subsequently examine the extent to which these differences reflect compositional factors by estimating regression models with individual controls and school-site fixed effects. Finally, we complement this analysis by documenting regional variation in both cognitive and socio-emotional outcomes (Figures 9–11), which provides additional context for the patterns observed across groups.

3.1 Differences by gender

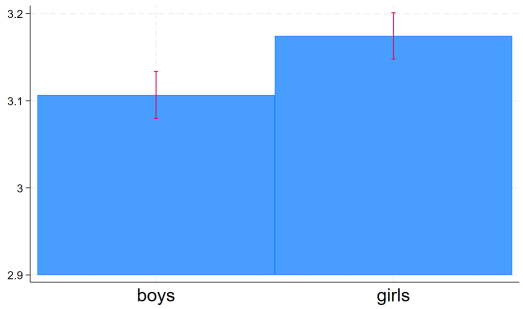
We now present descriptive evidence on the distribution of cognitive and non-cognitive skills by gender.[11] Figures 1–3 display the raw differences in mean values between boys and girls, along with the corresponding 95% confidence intervals. Overall, a clear gendered pattern emerges across most of the considered dimensions: girls tend to exhibit stronger non-cognitive skills (including socio-emotional skills and prosocial behavior), whereas boys perform better in mathematics-related outcomes.

In particular, girls score on average 0.07 points higher than boys (on a 1-4 scale) in the aggregate Socio-Emotional Health Survey index (Fig. 1). A sizeable and statistically significant difference emerges in Prosocial Behavior at School, where girls outperform boys by 0.9 points (on a 5–20 scale) (Fig. 2b). A small positive gap in favor of girls is also observed for Growth Mindset (0.01 points on a 0–1 scale), although this difference is not statistically significant (Fig. 2a).

Girls also display a slightly higher overall academic self-concept, with an average advantage of 0.35 points (on a 6-30 scale) relative to boys, though this difference is not statistically significant. A larger gap is observed in Italian self-concept, while differences in mathematics self-concept are smaller. In contrast, boys perform better in mathematics achievement, scoring on average one decile higher than girls on the grade-standardized math test, and also reporting higher math self-concept.

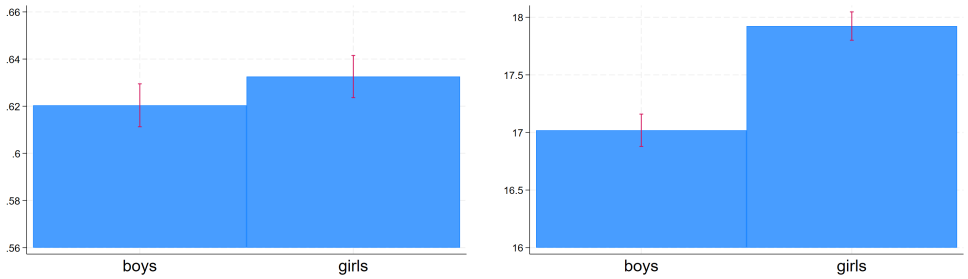
[11] The sample size may vary across figures due to item-level non-response. The maximum number of observations is 2,538.

Figure 1: Gender differences in socio-emotional health



Note: The figure reports the average score in socio-emotional health (measured with the SEHS - Furlong et al., 2020) by gender, with the associated 95% confidence intervals. The measure ranges from 1 to 4. The numbers are based on a sample of 2,191 children (1,106 boys and 1,085 girls).

Figure 2: Gender differences in growth mindset and pro-social behavior

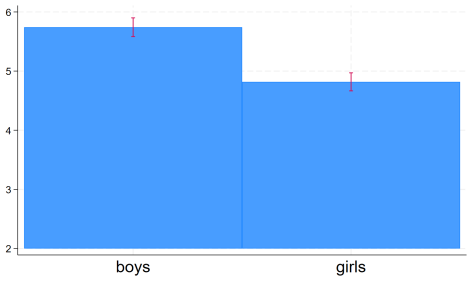


(a) Growth mindset

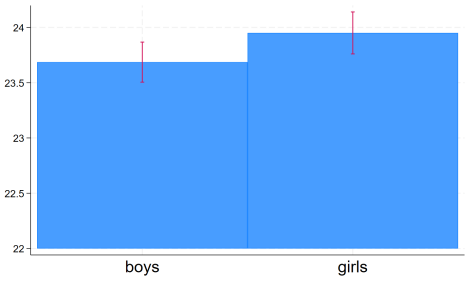
(b) Pro-social behaviour

Note: The figure reports the average score in growth mindset and pro-social behaviour by gender, with the associated 95% confidence intervals. Growth mindset, based on Muradoglu et al. (2022), ranges from 0 to 1, while Pro-social behavior (Furlong et al., 2013) is scaled from 5 to 20. The numbers are based on a sample of 2,346 (1,189 boys and 1,157 girls) for subfigure (a), and 2,290 (1,160 boys and 1,130 girls) for subfigure (b).

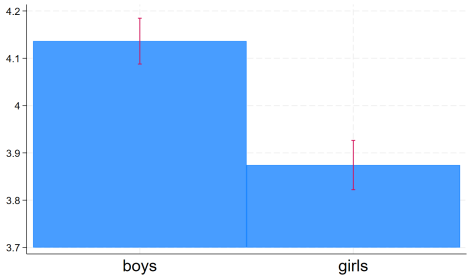
Figure 3: Gender differences in cognitive outcomes and academic self-concept



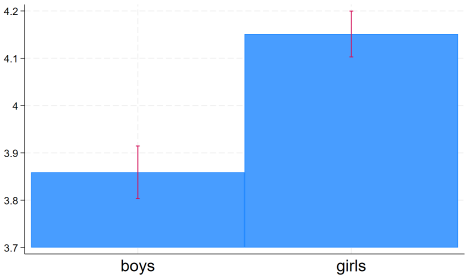
(a) Math test decile



(b) Academic self-concept overall



(c) Academic self-concept: Math



(d) Academic self-concept: Italian

Note: The figure reports the average score in cognitive outcomes (decile in the math test) and school self-efficacy by gender, with the associated 95% confidence intervals. Academic self-concept is measured on a 6-30 scale, while the specific subscales for Math and Italian self-concept are measured on a 1-5 scale. The numbers are based on a sample of 2,392 (1,216 boys and 1,176 girls) for subfigure (a), 2,298 (1,159 boys and 1,139 girls) for subfigure (b), 2,350 (1,189 boys and 1,161 girls) for subfigure (c), and 2,341 (1,184 boys and 1,157 girls) for subfigure (d).

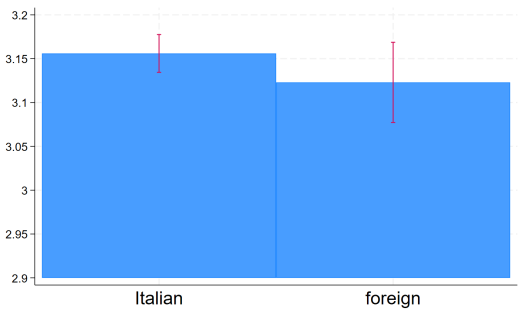
3.2 Differences by migratory background

We now turn our attention to differences by migratory background. In this case, differences consistently favor children with two Italian-born parents, although these gaps are not always statistically significant.

As shown in Figure 4, on average, native children score 0.03 points higher in the Socio-Emotional Health index than peers with at least one foreign-born parent. A similar pattern emerges for Growth Mindset (Fig. 5), with a modest advantage of 0.02 points, and for Pro-social Behavior at School, where the gap widens to 0.35 points in favor of native children.

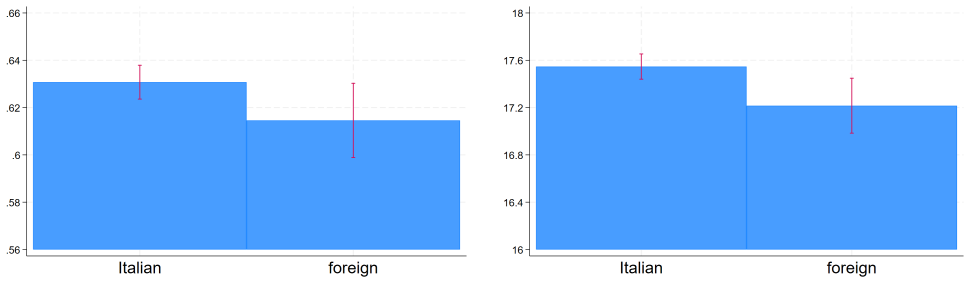
Similarly, Figure 6 indicates that differences also emerge in both perceived and actual academic performance. Native children rank, on average, 0.35 deciles higher than children with a migratory background on the mathematics test and score about 0.3 points higher on the Academic Self-Concept scale, indicating advantages both in objective achievement and in self-perceived school competence. The largest gap emerges in the Italian language self-concept.

Figure 4: Differences in socio-emotional health by migratory background



Note: Figure reports the average score in socio-emotional health (measured with the SEHS, Furlong et al. (2020)) of children with two Italian parents and children with at least a foreign parent, with the associated 95% confidence intervals. The measure ranges from 1 to 4. The numbers are based on a sample of 2,014 children (1,662 Italian and 412 foreign).

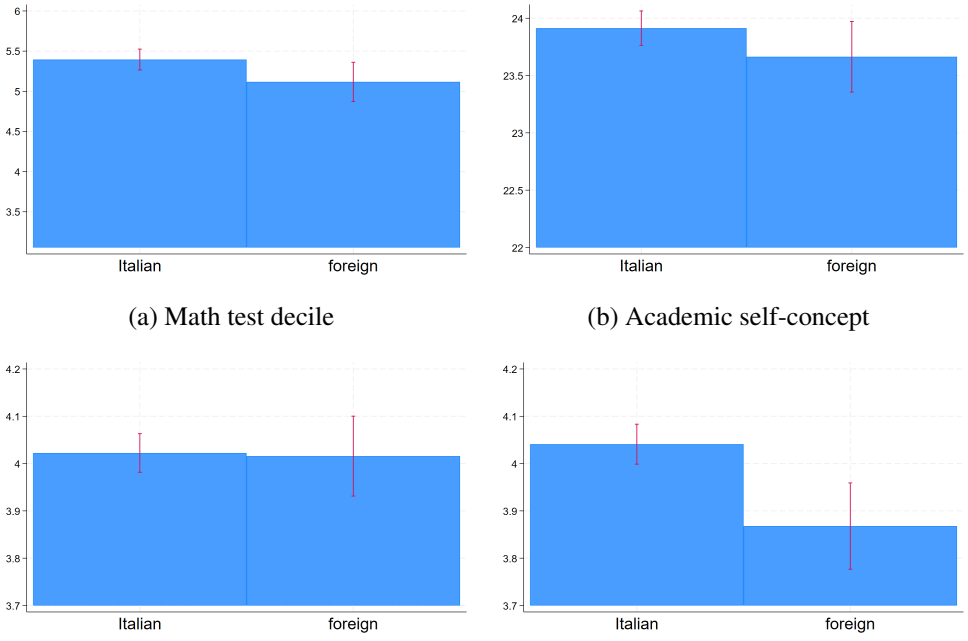
Figure 5: Differences in growth mindset and pro-social behaviour by migratory background



(a) Growth mindset (b) Pro-social behaviour

Note: Figure reports the average score in growth mindset and pro-social behaviour of children with two Italian parents and children with at least one foreign parent. Growth mindset, based on Muradoglu et al. (2022), ranges from 0 to 1, while Pro-social behavior (Furlong et al., 2013) is scaled from 5 to 20. The numbers are based on a sample of 2,229 (1,786 Italian and 443 foreign) for subfigure (a), and 2,173 (1,747 Italian and 426 foreign) for subfigure (b).

Figure 6: Differences in cognitive outcomes and academic self-concept by migratory background



(a) Math test decile (b) Academic self-concept

(c) Academic self-concept: Math (d) Academic self-concept: Italian

Note: The figure reports the average score in cognitive outcomes (decile in the math test) and school self-efficacy of children with two Italian parents and children with at least one foreign parent. Academic self-concept is measured on a 6-30 scale, while the Math and Italian self-concept subscales are measured on a 1-5 scale. The numbers are based on a sample of 2,268 (1,815 Italian and 453 foreign) for subfigure (a), 2,177 (1,749 Italian and 428 foreign) for subfigure (b), 2,232 (1,787 Italian and 445 foreign) for subfigure (c), and 2,220 (1,781 Italian and 439 foreign) for subfigure (d).

3.3 Differences conditional on individual and school characteristics

Next, we examine whether the observed raw differences by gender and migratory background persist after controlling for individual and school characteristics that may influence the outcomes of interest. To this end, we estimate the following regression model:

$$Y_{is} = \alpha + \beta_1 Boy_i + \beta_2 Migrant_i + \gamma_1 MomWorks_i + \gamma_2 N\_siblings_i + \gamma_3 Grade_i + \phi_s + \varepsilon_{is} \quad (1)$$

where Y_{is} denotes the standardized outcome variable for child i attending school site s. The coefficients of interest, β_1 and β_2, capture differences by gender and by migratory background,[12] conditional on maternal labor force participation, number of siblings, school grade, and school-site fixed effects ϕ_s. The school-site fixed effects absorb all unobserved characteristics that are constant within a school site, such as school quality (including teacher composition), peer composition, neighborhood characteristics, infrastructure, and other time-invariant local institutional factors. The error term ε_{is} captures idiosyncratic shocks.

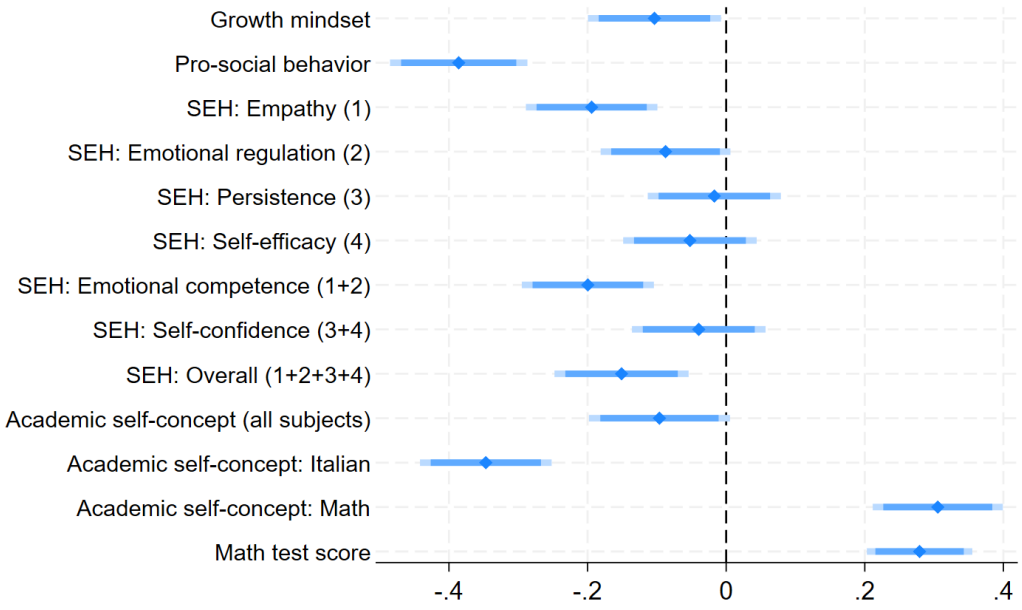
Figure 7 shows the estimated conditional gender differences. The results largely confirm the descriptive patterns documented earlier: gender gaps in socio-emotional and other non-cognitive outcomes persist and are estimated with greater precision once individual and school-level controls are included. Notably, the differences in growth mindset and academic self-concept, which were not statistically significant in the raw comparisons, become statistically significant after conditioning on these covariates.

Because the outcomes are standardized, comparisons across domains are more straightforward. The largest gaps in favor of girls are observed in prosocial behavior and self-concept in Italian (approximately 0.35 standard deviations (s.d.)). Gender differences in growth mindset and socio-emotional skills - including empathy, emotional regulation, and their composite measures - range between 0.1 and 0.2 s.d., again favoring girls. In contrast, persistence, self-efficacy, and their average (self-confidence) do not exhibit statistically significant gender differences. Consistent with prior evidence, a sizeable gap in favor of boys emerges in mathematics achievement and, correspondingly, in math self-concept (about 0.25 s.d.).

Figure 8 displays the conditional differences by migratory background, namely the estimates for β_2 in our sample. Controlling for child- and school-level characteristics does not substantially alter either the magnitude or the statistical significance of the native–migrant gaps presented earlier. Migrant students exhibit significantly worse skills (gaps of 0.20-0.15 s.d.) in prosocial behavior, persistence, self-efficacy, self-confidence, and Italian self-concept, and in math test scores (0.10). footnote Differences in growth mindset, emotional regulation, and overall academic self-concept are negative but not statistically significant. Notably, the gender gap in mathematics exceeds the foreign–native gap in the same domain. More broadly, disparities are more pronounced along the gender dimension.

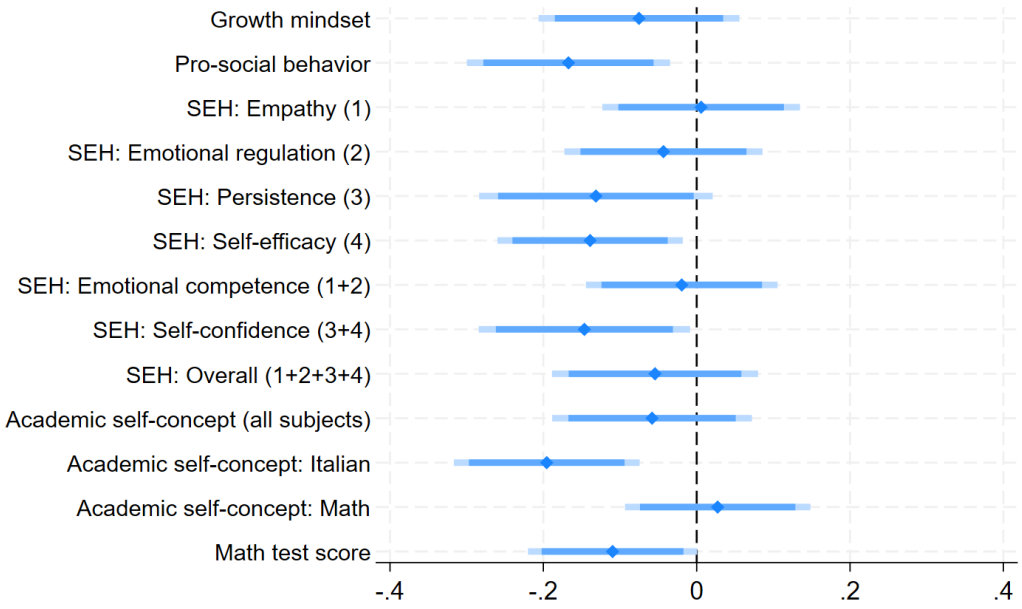
[12] *Migrant* is a dummy variable equal to 1 if at least one parent was born abroad and 0 otherwise, corresponding to the variable "Foreign" defined earlier in the text.

Figure 7: Gender differences in socio-emotional skills (boys - girls)



Note: The figure shows separate estimates for the gap between boys and girls in (standardized) socio-emotional and cognitive skills, controlling for migrant background, maternal employment, number of siblings, school grade, and school-site (*plesso*) fixed effects. Each coefficient and associated 90% and 95% confidence intervals reflect estimates for separate regressions in which the dependent variables are one of the (standardized) measured socio-emotional skills. SEH stands for 'Socio-Emotional Health'. Standard errors are clustered at the class level.

Figure 8: Differences in socio-emotional skills by migratory background (migrant - native)



Note: The figure shows separate estimates for the gap between children with children with at least one foreign parent ('migrant') and children with two Italian parents ('native') in (standardized) socio-emotional skills, controlling for a dummy indicating gender, a dummy indicating whether the child's mother is employed, number of siblings, school grade, and school site (*plesso*) fixed effects. Each coefficient and its associated 90% and 95% confidence intervals reflect estimates from separate regressions in which the dependent variable is one of the (standardized) measured socio-emotional skills. SEH stands for 'Socio-Emotional Health'. Standard errors are clustered at the class level.

3.4 Regional differences

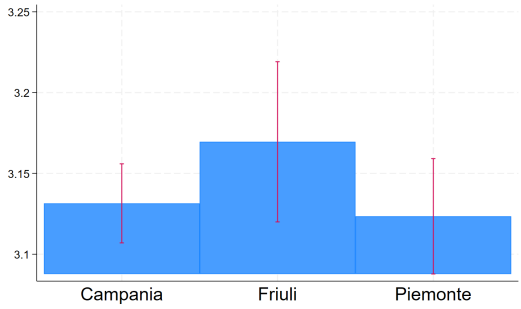
Finally, we also present descriptive evidence about regional differences across the various socio-emotional and cognitive outcomes considered.

Socio-emotional health (overall), displayed in Figure 9, is highest among students in Friuli Venezia Giulia, while no statistically significant differences are observed between Campania and Piemonte.

A different pattern is observed for Growth Mindset, which appears to be, on average, higher in Campania than in the other regions, with students in Friuli Venezia Giulia and Piemonte exhibiting broadly comparable mean values (10a). However, the relatively high variability observed in Friuli Venezia Giulia implies that these differences do not reach conventional levels of statistical significance. Pro-social behavior also shows a clear regional gradient (10b). Campania records the highest average levels, followed by Piemonte, with overlapping confidence intervals indicating that the difference between these two regions is not statistically distinguishable. Friuli Venezia Giulia, by contrast, exhibits the lowest average Pro-social behavior scores, significantly lower than Campania.

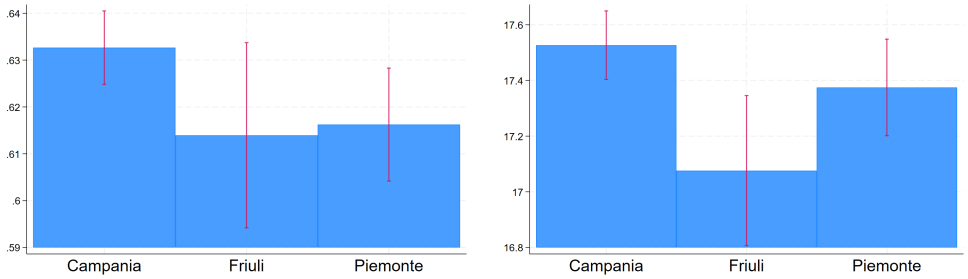
Academic self-concept also varies across regions, albeit to a more limited extent: it is marginally higher in Campania, followed by Friuli Venezia Giulia, while Piemonte displays the lowest average levels. In this case, however, the estimated differences are small in magnitude and not statistically significant. A distinct pattern emerges for cognitive performance as measured by the mathematics test. Here, Friuli Venezia Giulia exhibits the highest average score (5.8), followed by Piemonte (5.4), whereas Campania records the lowest mean performance (4.6). This ranking is consistent with well-documented evidence on geographical disparities in academic achievement (INVALSI data).

Figure 9: Regional differences in socio-emotional health



Note: The figure reports the average score in socio-emotional health of children across the three regions in the sample (measured with the SEHS, Furlong et al. (2020)), with the associated 95% confidence intervals. The measure ranges from 1 to 4. The numbers are based on a sample of 2,309 children (1,322 in Campania, 318 in Friuli-Venezia-Giulia, and 669 in Piemonte).

Figure 10: Regional differences in growth mindset and pro-social behaviour

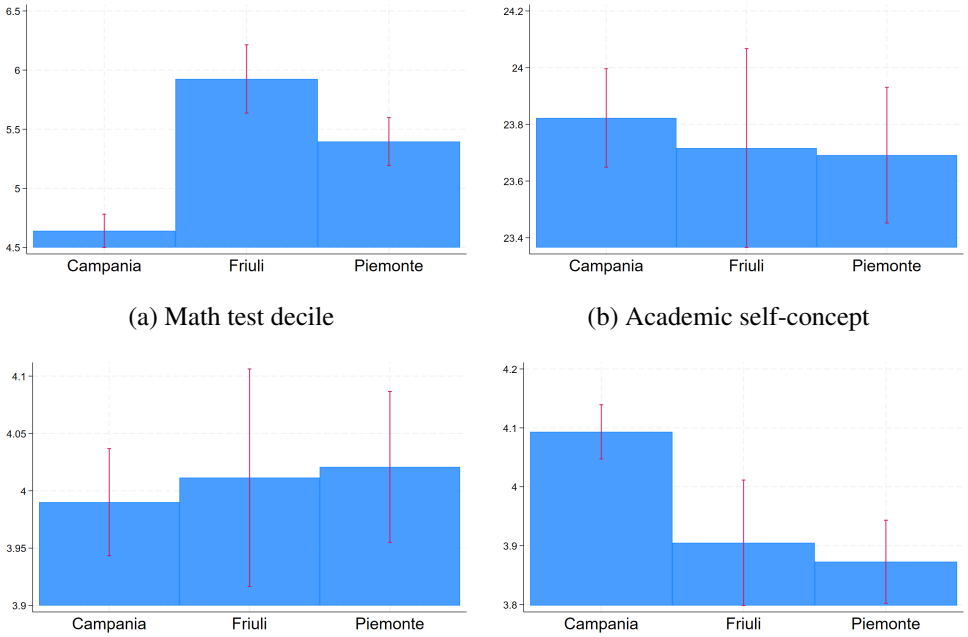


(a) Growth mindset

(b) Pro-social behaviour

Note: The figure reports the average score in growth mindset and pro-social behaviour of children across the three regions in the sample, with the associated 95% confidence intervals. Growth mindset, based on Muradoglu et al. (2022), ranges from 0 to 1, while Pro-social behavior (Furlong et al., 2013) is scaled from 5 to 20. The numbers are based on a sample of 2,446 (1,405 in Campania, 328 in Friuli-Venezia-Giulia, and 713 in Piemonte) for panel (a), and 2,413 (1,369 in Campania, 329 in Friuli-Venezia-Giulia, and 715 in Piemonte) for panel (b).

Figure 11: Regional differences in cognitive outcomes and academic self-concept



(a) Math test decile

(b) Academic self-concept

(c) Academic self-concept: Maths

(d) Academic self-concept: Italian

Note: The figure reports the average score in cognitive outcomes (schoolyear-specific decile in the math test) and school self-efficacy of children across the three regions in the sample, with the associated 95% confidence intervals. Academic self-concept is measured on a 6-30 scale, while the specific subscales for Math and Italian self-concept are measured on a 1-5 scale. The numbers are based on a sample of 2,673 (1,536 in Campania, 361 in Friuli Venezia-Giulia and 776 in Piemonte) for panel (a), a sample of 2,420 (1,375 in Campania, 342 in Friuli Venezia Giulia and 703 in Piemonte) for panel (b), a sample of 2,480 (1,409 in Campania, 349 in Friuli Venezia Giulia and 722 in Piemonte) for panel (c), and a sample of 2,467 (1,405 in Campania, 347 in Friuli Venezia Giulia and 715 in Piemonte) for panel (d).

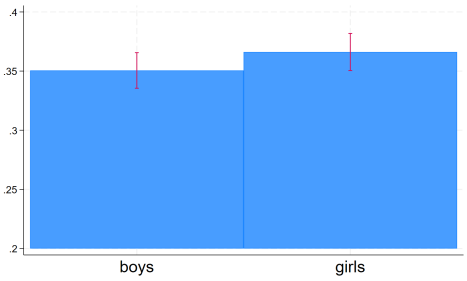
Two important caveats should be considered when interpreting these results. First, the socio-emotional measures are based on self-reported data. Consequently, children's responses may be influenced by their immediate reference group — namely, their classmates and peers — potentially introducing reference-group bias and contributing to patterns that may initially appear counterintuitive. Second, the sample used in this study is not representative of the regional population, and the findings should therefore be interpreted with caution. Nevertheless, it is reassuring that the regional differences observed in mathematics performance closely mirror those documented in the established and representative INVALSI assessments. Moreover, regional variation in maternal employment rates and in the share of students with a migratory background aligns with ISTAT's official statistics. These consistencies lend credibility to the patterns identified for cognitive outcomes and may also support the interpretation of differences in non-cognitive skills.

4 The influence of children's own gender norms

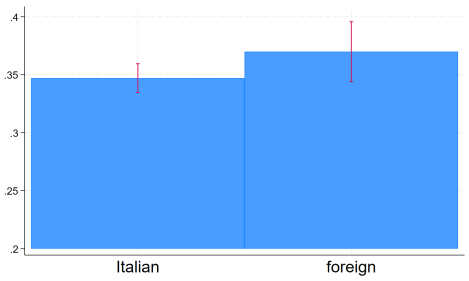
In this section, we examine the association between children's own gender attitudes and their cognitive and non-cognitive skills. The underlying hypothesis is that gender stereotypes — internalized through family, school, and the broader social environment — may play an important role in shaping children's socio-emotional and cognitive development, for example, by reinforcing gender categories and prescribing appropriate behaviors for boys and girls (Endendijk et al., 2014).

First, from a descriptive perspective, patterns of conformity to gender norms (our main explanatory variable) appear relatively stable across gender and migratory background, with only minor and statistically insignificant differences (Figures 12). Girls exhibit slightly higher conformity than boys, with a difference of 0.01 points on a 0–1 scale (i.e., 1 percentage point), indicating negligible variation between the two groups. Similarly, children with at least one foreign-born parent score 0.02 points higher on the gender norms conformity index compared to children with two Italian-born parents, though this difference is not statistically significant. Regional differences are also small (0.02 points), but statistically significant, with Piemonte displaying lower conformity than Campania and Friuli Venezia Giulia. This suggests that regional contexts may play a more important role in shaping gender norms conformity than other individual-level characteristics, such as gender or migratory background.

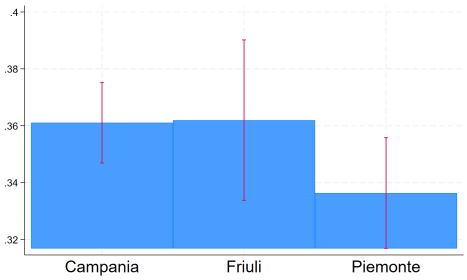
Figure 12: Differences in gender norms conformity



(a) By gender



(b) By migratory background



(c) By region

Note: The figure reports the average score as measured by our adaptation of the COAT-AM (Liben et al., 2002), which measures adherence to traditional gender norms in a scale from 0 (minimum conformity) to 100 (total conformity), with the associated 95% confidence intervals. The numbers are based on a sample of 1,126 boys and 1,096 girls for subfigure (a), 1,698 Italian and 413 foreign children for subfigure (b), and 1,312 children in Campania, 339 in Friuli-Venezia-Giulia, and 694 in Piemonte for panel (c).

We next investigate the relationship between children's gender attitudes and their cognitive and non-cognitive skills by estimating the following regression:

$$Y_{is} = \alpha + \beta\ GenderNormConf_i + \gamma X_i + \phi_s + \varepsilon_{is} \tag{2}$$

where Y_{is} denotes each of the cognitive and non-cognitive outcomes described in Table 1 (Section 2.2) for child i attending school site s. $GenderNormConf_i$ is the standardized index of adherence to gender stereotypes, our measure of gender norms based on Liben et al. (2002). The vector X_i includes child-level characteristics such as migratory background, maternal employment, number of siblings, and school grade. The term ϕ_s represents school-site (*plesso*) fixed effects. The error term ε_{is} captures idiosyncratic shocks. Standard errors are clustered at the class level. We estimate separate regressions for boys and girls, allowing for gender heterogeneity in the relationships of interest. β is our coefficient of interest and identifies the association between gender-norm conformity and children's outcomes for boys and girls, conditional on socio-demographic variables and time-invariant characteristics at the school-site level.

Figure 13 shows the estimated association between conformity to traditional gender norms and our indicators of socio-emotional health and related measures, controlling for a set of observable confounders (corresponding to the β parameter in Equation 2). Overall, the estimates point to a clear and systematic negative association between adherence to traditional gender norms and most dimensions of socio-emotional development. This inverse relationship is consistently stronger in magnitude for boys than for girls, suggesting that rigid gender attitudes may be particularly detrimental to boys' socio-emotional and academic outcomes.

Among girls, a one-standard-deviation increase in gender norms conformity is associated with a decrease in growth mindset, empathy, emotional regulation, emotional competence, and self-confidence (0.11 - 0.13 s.d.), as well as a decrease in math skills (0.11 s.d.). Smaller negative associations are also observed for persistence and self-efficacy (0.8 s.d.), whereas negative but non-significant associations emerge for pro-social behavior and academic self-concept. This last masks outcome heterogeneity: there is a zero effect on Italian self-concept, but a negative effect (-0.8 s.d.) on math self-concept.

For boys, the negative associations between conformity to gender norms and outcomes are generally larger and span a broader set of dimensions. A one–standard deviation increase in gender norms conformity is associated with a 0.14-0.15 standard deviation decrease in growth mindset, pro-social behavior, empathy, and emotional competence; a 0.10 s.d. reduction in self-confidence and in math skills. Emotional regulation and self-efficacy are also negatively correlated with gender norms conformity among boys, although these estimates are smaller in magnitude. Notably, we also observe a large reduction in overall and Italian academic self-concept (0.11 and 0.10 standard deviations, respectively), which was not affected for girls, but not in math self-concept.

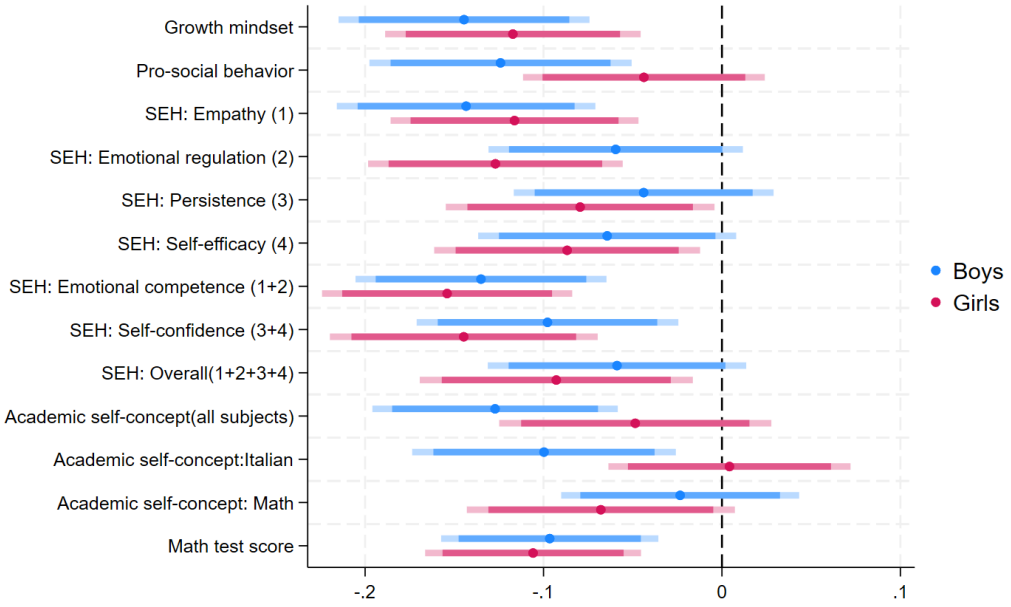
Overall, these results indicate that boys are more adversely affected by gender norms,

showing lower levels of non-cognitive skills such as growth mindset, pro-social behavior, empathy, and academic self-concept. Girls are more affected in terms of socio-emotional health (emotional regulation, persistence, and self-efficacy), except for empathy. The opposite effects in terms of academic self-concepts stand out.

Mathematical skills are similarly correlated with gender norms for both boys and girls, with statistically comparable coefficients. For both genders, a stronger conformity to stereotypical gender norms is associated with a poorer performance on the math test. This finding suggests that gender norms alone are not sufficient to explain the gender gap in math achievement.

While this empirical specification accounts for important factors that may correlate both with gender norms and socio-emotional skills, such as maternal employment (which is related to both parental gender norms, education, and socio-economic status) and school-site fixed effects (which capture shared aspects related to school climate and local gender norms), the results should be interpreted as correlational rather than causal. A fully causal interpretation would require richer information to better address residual sources of endogeneity, including parental education and socio-economic status, peer-group influences, teachers' beliefs and potential biases, as well as broader cultural and neighborhood characteristics that may jointly shape children's gender attitudes and socio-emotional development.

Figure 13: Socio-emotional skills and gender-norm conformity



Note: The figure shows separate estimates for the association between (standardized) gender norms and skills, controlling for migrant background, maternal employment, number of siblings, school grade, and school site (*plesso*) fixed effects, separately for boys (in blue) and girls (in red). Each coefficient and its associated 90% and 95% confidence intervals reflect estimates from separate regressions in which the dependent variable is one of the (standardized) measured socio-emotional skills. Standard errors are clustered at the class level. Estimates are based on a sample of a maximum of 1,777 children (848 boys and 929 girls) (the number of observations varies depending on the outcomes).

5 The role of teachers

We next move beyond children's observable characteristics to examine whether classroom environments help explain the skill gaps documented above. In particular, we study whether teachers' gender-related attitudes and self-reported practices are associated with students' cognitive and non-cognitive outcomes, and whether these associations differ for boys and girls. To do so, we link teacher questionnaire data to student outcomes and estimate regressions with student and teacher controls and school-site fixed effects, cautiously interpreting the results as conditional correlations rather than causal effects.

5.1 Teachers' gender attitudes and children's cognitive and non-cognitive skills

Building on the previous analysis, we extend our investigation to examine whether and how teachers' gender attitudes and practices are associated with students' cognitive and non-cognitive skills. To this end, this subsection focuses on teachers' gender stereotypes and their inclusive teaching practices. In the following subsection, we instead examine the role of teaching practices in greater detail, specifically whether they are more or less teacher-centered, as well

as beliefs about the nature of mathematics, learning, and teaching among the subsample of mathematics teachers in grade 3.

In this section, the sample of children is restricted to those classes with at least one teacher participating in the program who answered the teacher's questionnaire, and the maximum number of observations is thus reduced to 2,481 (instead of 2,538). Descriptive statistics for this smaller sample are very similar and available upon request. Teachers are 177 in 166 matched classes.

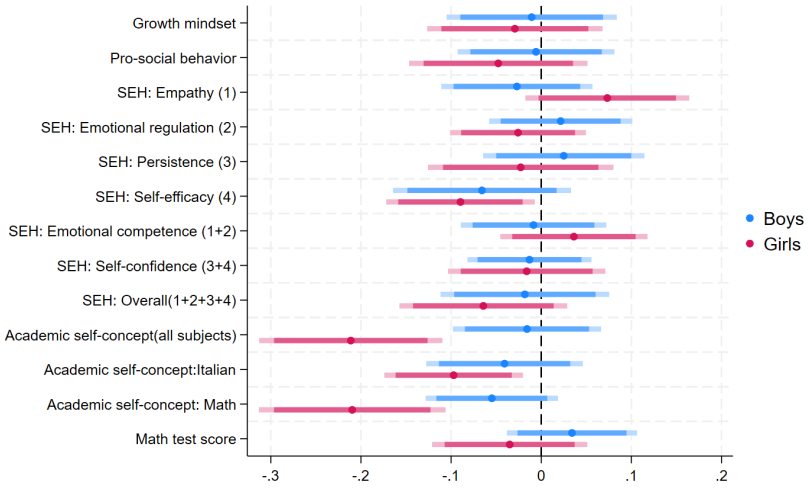
In order to analyze how teachers' gender attitudes impact their students' skills, we estimate the following regression, separately for boys and girls:

$$Y_{ijp} = \alpha + \beta T\_GenderBias_j + \gamma_1 X_i + \gamma_2 Z_j + \phi_p + e_{ijp} \tag{3}$$

Where Y_{ijp} stands for our set of cognitive and non-cognitive skill measurements, and $T\_GenderBias_j$ represents our main variable of interest, which captures the extent to which teachers attribute abilities or behaviors to students based on gender. X_i and Z_j refer to child-level (as above) and teacher-level observable characteristics (age, years of experience, an indicator for having a university degree, and an indicator for having a permanent contract) respectively, while ϕ_p represents school-site (*plesso*) fixed effects.

Our coefficients of interest, β, plotted in Figure 14, reflect the association between teachers' gender stereotypes and their students' skills, controlling for observed student characteristics and common school-specific factors, captured by the fixed effects.

Figure 14: Teachers' gender stereotypes in teaching and children's Socio-emotional skills



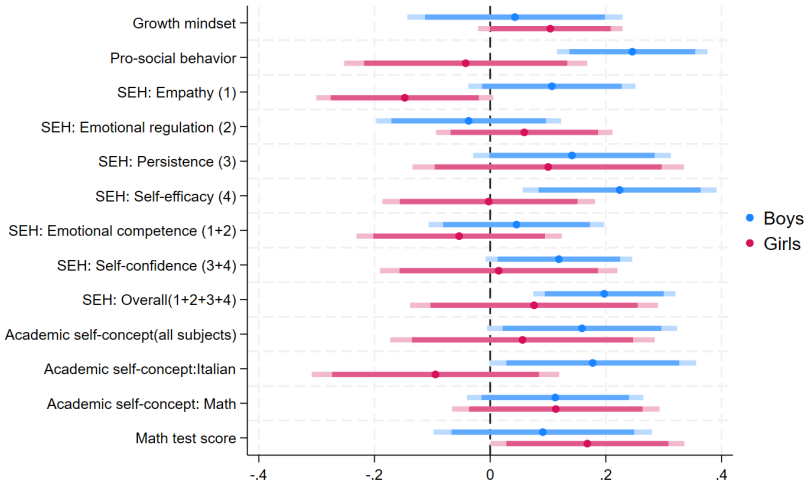
Note: The figure shows separate estimates for the association between (standardized) teacher's gender stereotypes index and children's socio-emotional skills, controlling for migrant background, maternal employment, number of siblings, school grade, teachers' characteristics, and school site (*plesso*) fixed effects, separately for boys (in blue) and girls (in red). Each coefficient and its associated 90% and 95% confidence intervals reflect estimates from separate regressions in which the dependent variable is one of the (standardized) measured socio-emotional skills. Standard errors are clustered at the class level.

The patterns differ markedly by student gender. Among girls, teachers' gender stereotypes are positively associated with empathy, but negatively associated with self-efficacy and academic self-concept—particularly in mathematics, though the pattern is also evident for Italian and overall self-concept. Among boys, the corresponding associations with socio-emotional skills are smaller and not statistically significant. For mathematics achievement, the coefficients have opposite signs (positive for boys and negative for girls); while neither estimate is statistically different from zero on its own, the difference between them is statistically significant, indicating a gendered divergence in how teacher stereotypes relate to math outcomes. Overall, these results are consistent with the view that teacher gender stereotypes may contribute to gender-specific patterns in socio-emotional development and academic confidence, with adverse implications for girls' math-related self-beliefs and performance.

We next examine teaching practices, a dimension that is not inherently gender-specific but may relate differently to boys' and girls' skill development. We estimate specifications analogous to equation 3, replacing the teachers' gender-stereotyping measure with a factor-based index capturing how inclusive teachers rate specific instructional practices and assessment methods.

The corresponding estimates for the association between this measure of inclusive teaching and children's skills are reported in Figure 15, separately for boys and girls. For boys, positive effects are observed across several non-cognitive skills, including prosocial behavior, persistence, self-efficacy, self-confidence, socio-emotional health (total), and academic self-concept, with an additional positive but non-significant impact on empathy and math performance. For girls, inclusive teaching is positively associated with growth mindset and math test scores, while other correlations are not significant. A notable exception is the negative association between teachers' inclusive teaching and lower empathy among girls.

Figure 15: Teachers' inclusive teaching and children's socio-emotional skills



Note: Figure shows separate estimates for the association between the (standardized) inclusive teaching index and children's socio-emotional skills, controlling for migrant background, maternal employment, number of siblings, school grade, teachers' characteristics, and school site (*plesso*) fixed effects, separately for boys (in blue) and girls (in red). Each coefficient and associated 90% and 95% confidence intervals reflect estimates for separate regressions in which the dependent variables are one of the (standardized) measured socio-emotional skills. Standard errors are clustered at the class level.

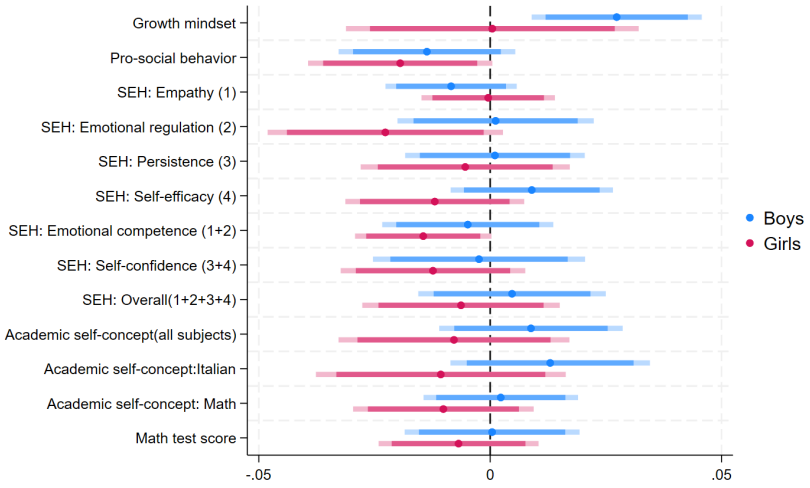
5.2 Teaching beliefs and practices and children's cognitive and non-cognitive skills

Finally, we focus on math teachers in grade 3 and their pupils, and assess whether teaching practices and beliefs are associated with changes in students' cognitive and non-cognitive outcomes. Again, we estimate equations similar to eq. 3, substituting teacher gender bias with outer explanatory variables of interest, namely teacher-centered practices (results shown in Figure 16) and beliefs about mathematics (described below in Figure 17).

As shown in Figure 16, for boys, a more teacher-centered approach to teaching appears to be positively associated with a growth mindset, suggesting that instructed teaching may reinforce beliefs about the malleability of ability. There are also positive, although not statistically significant, associations with self-efficacy and academic self-concept, especially in Italian. Notably, there is no observable relationship between this measure of teaching practices and mathematics outcomes: both math self-concept and actual performance in the math test show essentially a zero effect. Overall, these results suggest that teacher-centered practices may support internal motivational and self-perception aspects of boys' socio-emotional health, while having no effects on other aspects more related to external interactions.

For girls, instead, teacher-centered practices are consistently associated with lower socio-emotional health outcomes. In particular, a statistically significant negative relationship emerges between teacher-centeredness and pro-social behavior and emotional regulation, and emo-

Figure 16: Teacher-centered practices and children's socio-emotional skills



Note: The figure shows separate estimates for the correlation between the teacher-centered practices scale and children's socio-emotional skills, controlling for migrant background, maternal employment, number of siblings, school grade, teachers' characteristics, and school site (*plesso*) fixed effects, separately for boys (in blue) and girls (in red). Each coefficient and associated 90% and 95% confidence intervals reflect estimates for separate regressions in which the dependent variables are one of the (standardized) measured socio-emotional skills. Standard errors are clustered at the class level.

tional competence. The overall pattern points towards a mild negative relation between teacher-centered instructional practices and girls' socio-emotional development.

Focusing now on teachers' beliefs about the nature of mathematics, teaching, and learning, we analyze the correlation between the teachers' agreement with the three orientations developed by Swan (2006) (Transmission, Discovery, and Connectionist) and children's outcomes. These correlations differ markedly by gender (Fig. 17).

For boys, a stronger transmission orientation is associated with a statistically significant positive effect on growth mindset, in line with the patterns found above regarding structured and teacher-led teaching practices. There is essentially no association with either math self-concept or math test score. For girls, however, the patterns that emerge are more mixed. In particular, a more transmission-oriented teacher is associated with statistically significant negative effects on pro-social behavior and self-efficacy.

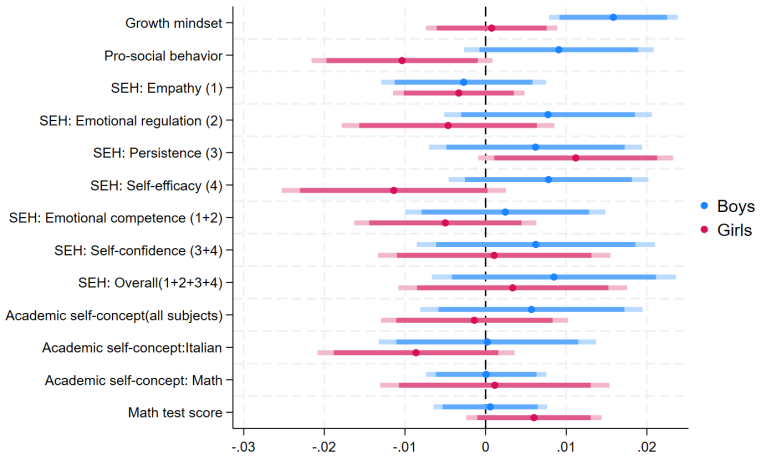
Regarding discovery-oriented teacher beliefs, the associations again differ by gender. For boys, the patterns appear mixed and with no significant associations. For girls, in contrast, discovery-oriented beliefs are more consistently linked to positive outcomes: statistically significant positive associations emerge with empathy, emotional competence, self-confidence, socio-emotional skills overall, and academic self-concept, particularly in Italian. Positive, but not statistically significant, associations emerge as well for pro-social behavior, math academic self-concept, and persistence. However, the coefficients for math test scores are negative (even if not statistically significant, possibly due to the small number of teachers ' observa-

tions). Taken together, these patterns suggest that discovery-oriented beliefs may help girls' socio-emotional development and academic self-perceptions, even if they do not translate into improvements in math achievement.

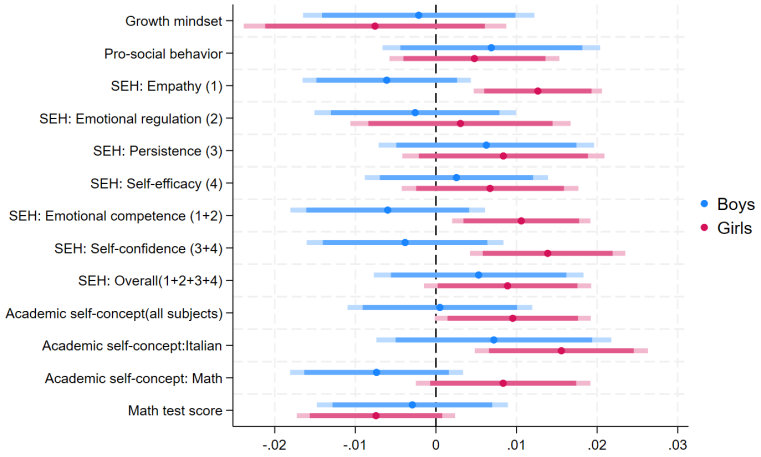
Turning to connectionist-oriented beliefs, the patterns seem more negative, with some gender variation. For boys, a connectionist orientation is associated with statistically significant negative effects on growth mindset, pro-social behavior, and persistence. Additional negative associations, though not significant, emerge for emotional regulation, self-efficacy, and academic self-concept (both overall and in Italian). At the same time, there is a statistically significant positive relationship with empathy. The association with math self-concept is positive but not statistically significant, while the coefficient for math test score is close to zero. For girls, connectionist teacher beliefs are associated with statistically significant negative effects on empathy, persistence, self-confidence, socio-emotional health overall, and math self-concept. In contrast, growth mindset and pro-social behavior show positive but non-significant associations. Similarly to boys, the relationship with math performance is very close to zero. Overall, this orientation appears to have limited links to mathematics outcomes for either gender, while effects on socio-emotional outcomes are prevalent and of different strengths across genders.

Overall, instructional practices and belief orientations have small and often not statistically significant correlations with children's outcomes, and clear effects on math performance are largely absent for both girls and boys. At the same time, some consistent gendered patterns emerge: more traditional teacher-centered and transmission-oriented approaches tend to be more positively aligned with outcomes related to motivation for boys, such as a growth mindset. Discovery-oriented beliefs appear more favorable for girls' socio-emotional development and academic self-perceptions, with limited/no association for boys. Finally, connectionist beliefs show mixed and prevalently negative negative associations, particularly for socio-emotional dimensions. This evidence suggests that the relationship between teaching approaches and student development is gender-specific, but effect sizes and statistical significance patterns call for cautious interpretation.

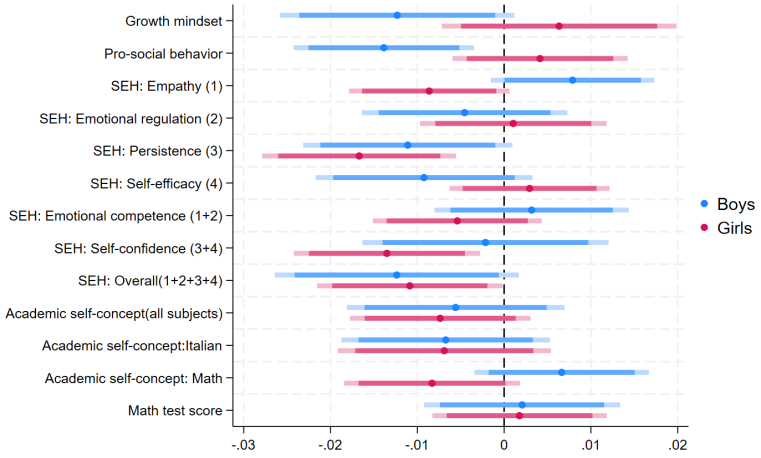
Figure 17: Teacher belief orientations and children's socio-emotional skills



(a) Transmission



(b) Discovery



(c) Connectionist

Note: The figure shows separate estimates for the correlation between the three scales of teachers' belief orientations and children's socio-emotional skills, controlling for migrant background, maternal employment, number of siblings, school grade, teachers' characteristics, and school site (*plesso*) fixed effects, separately for boys (in blue) and girls (in red). Each coefficient and its associated 90% and 95% confidence intervals reflect estimates from separate regressions in which the dependent variable is one of the (standardized) measured socio-emotional skills. Standard errors are clustered at the class level.

6 Conclusions

We present new evidence on early inequalities in cognitive and non-cognitive skills among primary school children in Italy, with a focus on gender and migratory background. Using original data from more than 2,500 pupils in grades 3–5, we document a consistent pattern: girls show stronger socio-emotional and prosocial outcomes and higher Italian self-concept, while boys score higher in mathematics and math self-concept. We also find that children with a migratory background tend to perform worse in mathematics and in several non-cognitive domains, indicating that multiple gaps emerge early in the school trajectory. We further document regional heterogeneity: socio-emotional outcomes vary across regions, and mathematics performance is highest in Friuli Venezia Giulia and lowest in Campania, in line with broader evidence on territorial gaps.

A second contribution concerns the role of gender norms and classroom factors. Children's stronger adherence to traditional gender norms is systematically associated with weaker socio-emotional outcomes, lower growth mindset, and lower academic self-perceptions, with more pronounced associations for boys in several dimensions. At the teacher level, gender stereotypes are linked to lower self-efficacy and academic self-concept among girls, while more inclusive teaching practices are generally linked to better outcomes, including higher mathematics achievement for girls and stronger socio-emotional indicators for boys. Taken together, these findings suggest that gender norms and instructional environments are relevant correlates of early skill formation.

Overall, the evidence in this paper highlights that inequalities in socio-emotional skills, academic self-perceptions, and mathematics achievement emerge early and are patterned by gender, migratory background, and the educational environments children face. These results point to concrete policy priorities. They suggest the value of primary-school interventions that combine cognitive and socio-emotional objectives, strengthen teacher training on gender bias and inclusive teaching, and provide targeted support for pupils with a migratory background. Overall, reducing early inequalities appears to require coordinated action on instructional practices, classroom climate, and students' confidence in their own learning potential.

The analysis has several important limitations. First, the evidence is correlational and based on pre-treatment observational data, so the estimated relationships should not be directly interpreted as causal effects. Second, several outcomes rely on children's self-reports and may be affected by reporting or reference-group bias. Third, the sample is not fully representative at the regional level, and some relevant family and contextual characteristics are only partially observed. Future research should therefore exploit the longitudinal and experimental components of the broader project to identify causal mechanisms, assess the persistence of effects over time, and determine which school practices are most effective in reducing early skill gaps.

References

Alan, S., Ertac, S., & Mumcu, I. (2018). Gender stereotypes in the classroom and effects on achievement. *Review of Economics and Statistics*, *100*(5), 876–890.

Alesina, A., Carlana, M., La Ferrara, E., & Pinotti, P. (2024). Revealing stereotypes: Evidence from immigrants in schools. *American Economic Review*, *114*(7), 1916–1948.

Bertocchi, G., & Bozzano, M. (2020). Gender Gaps in Education. In K. F. Zimmermann (Ed.), *Handbook of Labor, Human Resources and Population Economics* (pp. 1–31). Cham: Springer International Publishing. Retrieved from https://doi.org/10.1007/978-3-319-57365-6_26-1 doi: 10.1007/978-3-319-57365-6_26-1

Bertrand, M. (2020). Gender in the twenty-first century. *AEA Papers and Proceedings*, *110*, 1–24.

Brunello, G., & Schlotter, M. (2011). Non-cognitive skills and personality traits: Labour market relevance and their development in education & training systems. *IZA Discussion paper*, *5743*.

Caprara, G. V., & Gerbino, M. (2001). Autoefficacia emotiva: La capacita'di regolare l'affettività negativa e di esprimere quella positiva (Emotional self efficacy: Managing negative affect and expressing positive affect). In G. V. Caprara (Ed.), *La valutazione dell'autoefficacia* (pp. 35–50). Erickson.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, *134*(3), 1163–1224.

Carneiro, P., Meghir, C., & Parey, M. (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, *11*(suppl_1), 123–160.

Contini, D., Di Tommaso, M. L., & Mendolia, S. (2017). The gender gap in mathematics achievement: Evidence from Italian data. *Economics of Education Review*, *58*, 32–42.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, *78*(3), 883–931.

Del Bono, E., Francesconi, M., Kelly, Y., & Sacker, A. (2016). Early maternal time investment and early child outcomes. *The Economic Journal*, *126*(596), F96–F135.

Della Giusta, M., & Muratori, C. (2022). Un'indagine sulle tematiche di genere nella scuola. In M. Della Giusta, B. Poggio, & M. Spicci (Eds.), *Educare alla parita: Principi, metodologie didattiche e strategie di azione per l'equita e l'inclusione.* Pearson Academy.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*(4), 1593–1640.

Encinas-Martín, M., & Cherian, M. (2023). *Gender, Education and Skills*. OECD. Retrieved from https://doi.org/10.1787/34680dd5-en

Endendijk, J. J., Groeneveld, M. G., Van der Pol, L. D., Van Berkel, S. R., Hallers-Haalboom, E. T., Mesman, J., & Bakermans-Kranenburg, M. J. (2014). Boys don't play with dolls: Mothers' and fathers' gender talk during picture book reading. *Parenting*, *14*(3-4), 141–161.

Ewing, D. L., Monsen, J. J., & Kielblock, S. (2018). Teachers' attitudes towards inclusive education: a critical review of published questionnaires. *Educational Psychology in Practice*, *34*(2), 150–165.

Fiorini, M., & Keane, M. P. (2014). How the allocation of children's time affects cognitive and noncognitive development. *Journal of Labor Economics*, *32*(4), 787–836.

Furlong, M. J., Nylund-Gibson, K., Dowdy, E., Wagle, R., Hinton, T., & Carter, D. (2020). Modification and Standardization of Social Emotional Health Survey-Secondary–2020 Edition. *Grantee Submission*.

Furlong, M. J., You, S., Renshaw, T. L., O'Malley, M. D., & Rebelez, J. (2013). Preliminary development of the Positive Experiences at School Scale for elementary school children. *Child Indicators Research*, *6*(4), 753–775.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, *19*(4), 451–464.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*(3), 411–482.

Huggett, M., Ventura, G., & Yaron, A. (2011). Sources of lifetime inequality. *American Economic Review*, *101*(7), 2923–2954.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581.

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, *65*(1), 373–398.

Kosse, F., Deckers, T., Pinger, P., Schildberg-Hörisch, H., & Falk, A. (2020). The formation of prosociality: causal evidence on the role of social environment. *Journal of Political Economy*, *128*(2), 434–467.

Liben, L. S., Bigler, R. S., Ruble, D. N., Martin, C. L., & Powlishta, K. K. (2002). The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the society for research in child development*, i–183.

Meroni, E. C., Piazzalunga, D., & Pronzato, C. (2022). Allocation of time and child socio-emotional skills. *Review of Economics of the Household*, *20*(4), 1155–1192.

Muradoglu, M., Porter, T., Trzesniewski, K., & Cimpian, A. (2022). Developing and Validating a Growth Mindset Scale for Young Children (GM-C). *PsyArXiv Preprints*.

Sorrenti, G., Zölitz, U., Ribeaud, D., & Eisner, M. (2025). The causal impact of socio-emotional skills training on educational success. *Review of Economic Studies*, *92*(1), 506–552.

Swan, M. (2006). Designing and Using Research Instruments to Describe the Beliefs and Practices of Mathematics Teachers. *Research in Education*, *75*(1), 58-70. Retrieved from https://doi.org/10.7227/RIE.75.5 doi: 10.7227/RIE.75.5