SURVEYS ON STUDENTS: INVALSI NATIONAL AND INTERNATIONAL TESTS

VI Seminar "INVALSI data: a tool for teaching and scientific research"

edited by
Patrizia Falzetti

FrancoAngeli

INVALSI

INVALSI PER LA RICERCA
STUDI E RICERCHE

INVALSI PER LA RICERCA

La collana Open Access INVALSI PER LA RICERCA si pone come obiettivo la diffusione degli esiti delle attività di ricerca promosse dall'Istituto, favorendo lo scambio di esperienze e conoscenze con il mondo accademico e scolastico.
La collana è articolata in tre sezioni: "Studi e ricerche", i cui contributi sono sottoposti a revisione in doppio cieco, "Percorsi e strumenti", di taglio più divulgativo o di approfondimento, sottoposta a singolo referaggio, e "Rapporti di ricerca e sperimentazioni", le cui pubblicazioni riguardano le attività di ricerca e sperimentazione dell'Istituto e non sono sottoposte a revisione.

OPEN ACCESS
la soluzione FrancoAngeli

Il presente volume è pubblicato in open access, ossia il file dell'intero lavoro è liberamente scaricabile dalla piattaforma **FrancoAngeli Open Access** (http://bit.ly/francoangeli-oa).

FrancoAngeli Open Access è la piattaforma per pubblicare articoli e monografie, rispettando gli standard etici e qualitativi e la messa a disposizione dei contenuti ad accesso aperto. Oltre a garantire il deposito nei maggiori archivi e repository internazionali OA, la sua integrazione con tutto il ricco catalogo di riviste e collane FrancoAngeli massimizza la visibilità, favorisce facilità di ricerca per l'utente e possibilità di impatto per l'autore.

Per saperne di più: Pubblica con noi

SURVEYS ON STUDENTS: INVALSI NATIONAL AND INTERNATIONAL TESTS

VI Seminar "INVALSI data: a tool for teaching and scientific research"

edited by
Patrizia Falzetti

INVALSI

FrancoAngeli

 ISBN 9788835164302

The opinions expressed are solely of the author(s). In no case should they be considered or construed as representing an official position of INVALSI.

Index

Introduction

by Patrizia Falzetti

INVALSI, as part of the National System of Evaluation of the Education and Training System (SNV), conducts periodic and systematic tests on students' knowledge and skills. Albeit with some modifications over time, these standardised tests have been objectively measuring for about 20 years students' achieving and learning in some main skills in Italian, Mathematics and English domains.

In addition to conducting the National Survey, INVALSI coordinates and ensures the participation of Italy in certain main international surveys[1] in education promoted by OECD (Organisation for Economic Cooperation and Development) and IEA (International Association for the Evaluation of Educational Achievement) which, both of them, carry out specific tests on some students' literacies and skills.

At the end of each survey, INVALSI makes useful databases available for studying and analysing the Italian education system – with an international comparison as well – and, on the occasion of the VI Seminar "INVALSI data: a tool for teaching and scientific research" (Rome, from 25th to 28th November 2021), the potential of their use became evident.

This volume collects some papers presented there. In detail, the research in chapters 3, 5, 6 and 7 use data from both national and international surveys. Chapter 3 focuses on the gender gap in Mathematics, also documented by the results of OECD-PISA survey in which for the Italian case it is wider than the international average. The exploratory analysis conducted aims to determine whether it is possible to limit some categories of item in which the gender gap is particularly notable, and to interpret these categories with

[1] For details on international surveys https://invalsi-areaprove.cineca.it/index.php?get=static&pag=indagini_internazionali_in_evidenza.

the theory of the educational didactic contract. In chapter 5, data from a sub-sample of students participating in the TIMSS 2019 survey – grade IV – are used to investigate the sense of school belonging of students with special educational needs by comparing it with the one of students without special educational needs. The starting assumption is represented by the assorted studies that have highlighted its important role in school inclusion: students who feel accepted and supported in the classroom context are more motivated to learn and more encouraged to participate in school activities. In Chapter 6, researchers investigated Italian ICILS 2018 survey data in order to study the role of schools in reducing the digital divide, a specific topic which became more crucial during the Covid-19 pandemic when digital skills were found to be a discriminating factor with respect to learning. Finally, the research in Chapter 7 aimed to compare the Italian students' results in Mathematics in OECD-PISA survey 2018 in the second year of upper secondary school with those obtained by the same students in INVALSI national tests in the same year. The analyse starts from the premise that these two surveys have different focuses and characteristics but also have points in common, such as the assessment of students' basic skills.

Chapters 1, 2 and 4, on the other hand, focus only on data from INVALSI surveys. Here, topics investigated are about the specificity of the case of the autonomous province of Trento; the relationship between economic, social and cultural status and school performance regarding English language skills; the impact that the age of the student has on the results of INVALSI tests.

The book is therefore full of insights on the possible uses of national and international surveys. We hope that from it reading, researchers, teachers and all stakeholders could find further stimuli to better investigate the Italian education system thanks to INVALSI data and beyond.

1. Better than whom? Methodological and substantial considerations emerged while going beyond a plain comparison of school performance in Trentino and in nearby provinces

by Gianluca Argentin, Chiara Tamanini, Loris Vergolini

The chapter aims to analyse the specificity of the case of the Autonomous province of Trento, questioning whether it is possible to speak of Trentino as an area of high learning. We analyse two dimensions related to school performance: the learning in Italian and Mathematics and social inequalities based on gender, migration background and parental education. The analytical strategy is based on the comparison of the province of Trento with the rest of the country, the North-East and the neighbouring provinces, controlling for a relevant set of observed characteristics to obtain estimates not biased by compositional effects. The results show that the advantage of Trentino disappears substantially once the comparison is made with contiguous geographical areas and that even in terms of equity the situation is similar with respect to the neighbouring provinces.

Il contributo si pone l'obiettivo di analizzare la specificità del caso della Provincia autonoma di Trento interrogandosi se effettivamente si può parlare del Trentino come di una zona con apprendimenti elevati. Nel capitolo analizziamo due dimensioni legate alle performance scolastiche: gli apprendimenti in Italiano e Matematica e le disuguaglianze sociali basate su genere, background migratorio e livello di istruzione dei genitori. La strategia analitica si basa sul confrontare la provincia di Trento con il resto del Paese, il Nord-Est e le sole province confinanti controllando per un insieme rilevante di caratteristiche osservate in modo da ottenere stime non influenzate da effetti di composizione. I risultati mostrano come il vantaggio del Trentino sparisca sostanzialmente una volta che il confronto viene fatto con realtà geografiche contigue e che anche in termini di equità la situazione è simile rispetto alle province confinanti.

1. Introduction

The growing availability or quantitative data coming from large-scale assessments, both at national and international level, raises several issues about their usage. Usually, criticisms towards statistical measures of pupils' achievement focus on the validity of standardized tests and on their repercussions on teachers' and school principals' behaviours (Koretz, 2009). The major risk underlined is that standardized assessments may severely reduce the set of school outcomes considered by actors in education and – especially when results imply high stakes – force them to put all their effort on a narrow set of actions directly connected to those measures. This clearly affect actors operating daily in the schools, such as teachers and principals, but also policy makers that, at different levels, take decisions about the reform of the education system, on the basis of what works or, better, on what seems to work.

Here we want to underline that the results coming from large-scale assessments are influential in the diagnosis of the education systems'. One of the mechanisms through which this process takes place is the attention paid to deviant mean scores, namely particular high or low performances of specific contexts. Detecting at local level average results superior or inferior to the ones of the overall population translates almost immediately in policy relevant questions. Two examples may clarify to the reader what we refer to. First, everybody knows that the Finnish education system displays very high pupils' performance: this makes it a widely investigated case study, in order to understand which features are underneath its effectiveness (Morgan, 2014). Second, scholars in Italy well know that pupils in Southern regions display performances lower than in the Northern ones: these results generated a wide set of analyses trying to explain to what extent this is consequential to different socioeconomic environments or to heterogenous schools' effectiveness (Bratti, Checchi and Filippin, 2007; Argentin *et al.*, 2017). Understanding what generates a difference in pupils' achievement, across nations or regions, is a form of comparison that seems to satisfy our need to understand what we can do to improve performance in education, learning lessons from contexts where things seem to go in the right direction. As we will show in the next pages, things are more complicated.

2. Research questions

One of the constant results detected by INVALSI national assessment year after year is the fact that Trento province obtain average performances higher than the national ones, but also superior to the ones of other (well performing) Northern regions. These comparisons have been widely reported for years in the media, both at national and local level, leading to a common consensus about the fact that the school system in Trento performs better than the one in other Italian regions. Also the negative consequences of COVID-19 pandemic on students outcomes seemed to be mitigated in this area.

The province of Trento, based on its special statute, enjoys a state of autonomy and has, among other perquisites, the official proxy to administer its school system. In the past, it has judiciously used this privilege, has invested in the school system introducing innovative practices sometimes taken as an example from the rest of the country. The fact that this province is autonomous in the management of its school system increases the attention to its results, also because several peculiar features characterize education in Trento context. Among many others, this province displays a larger amount of resources invested in education (CPT, 2019), more widespread ICT technologies in schools, more intense and persistent attention to teachers' professional development, a developed dual system where vocational track is larger[1] and more rooted then elsewhere, ecc. Which features are at the basis of this successful results? What makes Trentino a constantly high performer in the Italian school system? What can we learn from this territory for the national governance of the school system?

In this chapter we will not focus on these research questions, but we will a step back in order to understand if they are meaningful.

We wonder whether the school system in Trentino is really performing better than elsewhere. We focus on two dimensions of performance: average achievement and equity. More precisely, we try to: a) compare more carefully than usual pupils' average performance in Trentino with pupils' average performance in other contexts; b) estimate in Trentino and elsewhere the intensity of unequal performances related to ascriptive characteristics (gender, migration background and parental education).

3. Data and methods

Before focussing on our analytical strategy, some considerations about the common use of INVALSI assessment.

[1] According to ISTAT data about the 20% of students enrolled in upper secondary school in Trentino attend vocational training with the respect to the 8% at national level.

3.1. Preliminary considerations about INVALSI data and their use

After several years of controversies, nowadays national standardized tests are widely considered important to ensure transparency to the Italian educational system based on school autonomy and to identify improvement strategies in learning. The INVALSI tests can signal imbalances and macro-disequilibrium between various regions of the country, for example, those between the North and the South (INVALSI, 2021) and can show hidden fragilities between the implicit and explicit scholastic dispersion (Ricci, 2019); nonetheless, they do not have the objective to draw up rankings among teachers or scholastic institutions, and not even among regions or provinces. What happens in reality is that rankings are widely use and highlighted by press releases by the school authorities at the regional or provincial levels and resumed with clamour by the various media, especially when results are positive, such as in the case of Trentino, or negative.

The attention to rankings, in public debate, comes to the price of over-simplification: a grand quantity of contextual factors (i.e. the conditions in which schools operate, the background of the students, the alumni outcomes throughout the years) are not considered in the mean scores displayed for a list of regions or provinces. Even less attention is played to other crucial outcomes coming from INVALSI assessment, such as the added value of each school; the results of each student to the answers, item by item, of the questions and, the levels of the students expressed in a qualitative manner (De Simoni, 2018). INVALSI also puts at disposition practical guides to explain the level of competences that the students are expected to attain to demonstrate and develop the abilities on the basis of the INVALSI frameworks (Ricci, 2020). This is the part more important and significant to the scholastic world deriving from INVALSI assessment, but the one remaining submerged in the public discourse, much more focused on ranking and exceptionally high or low performances.

In the next pages, we rely on a dataset adding up several years of INVALSI assessments; we will try to take into account the complexity deriving by the choice to use these data to develop comparisons across regions, exploiting a part of the information available in the datasets usually not adequately considered in rankings.

3.2. Analytical strategy

In order to answer our research questions, we focus on INVALSI data collected at 8th grade[2], the end of lower secondary education, hence the last school level where Trento and other Italian regions are fully comparable and being it the ending point of a relevant part of the education pathway for students. In fact, in Trento the second cycle is based on a dual system in which the vocational track (*Formazione Professionale*) is eliminated and absorbed into the technical one (*Indirizzo Tecnico*) while a broad formation sector is formed from the Vocational Education and Training Courses.

We expect that, after 8 years of mandatory schooling, if the education system is different in Trentino, we should detect at that point its cumulated beneficial effects. Trying to establish a causal inference attributing better pupils' performance to the education system in Trento, we need to move beyond the usual plain averages comparison. In particular, we want to be sure that the higher education performance in Trento is not due to: a) estimates' random variation; b) different compositions of students' populations in Trento and other regions; c) the fact that Trento is a small and uniform territory, usually compared to much larger and heterogeneous regions (such as Veneto, Lombardy, etc.)[3].

Regarding the first element (point a), we may rely on large statistical samples[4] and in addition on the fact that we use INVALSI data for seven school year. These two elements allow us to rely not only on low statistical uncertainty (narrow interval confidence), but also on the longitudinal persistence/randomness of results coming from the comparison.

In order to reduce the risk of attributing to the school system in Trento an effect due to different composition of pupils in this province compared to the other ones (point b), we use OLS regression models controlling the difference between our province of interest with the other ones for four key covariates predicting students achievement, available in the INVALSI da-

[2] We consider scores corrected for cheating both for Italian and Maths.

[3] It should be noted that Trento is part of a region totally distinctive and composed of three entities: the province of Trento itself; the Province of Bolzano; and, the autonomous Region Trentino-Alto, Adige/Südtirol. Such a composition, connected to well-entrenched historical, geographical, and cultural factors, do not permit a homogenous comparative analysis (Marcantoni, 2019) with the other Italian regions.

[4] The overall samples adding up 7 years are based on 16,215 students for Trento, 275,787 students for the North-East and 94,231 students for the neighbouring provinces (due to slightly variations in the number of pupils filling the Italian/Math assessments, we report here the lowest value for each sample).

tasets, namely: parental education and occupational class, migration background and sex. When the comparison adds up in only one parameter all the school years, models include a dummy variable for each year, in order to capture sample/assessment specificities. The comparison regarding equity relies on the same OLS model, where we added interaction terms (one for each covariate). Finally, about the need to identify a proper geographical counterfactual to Trento province and its pupils (point c), we present a first comparison where the province of interest is compared to the North Eastern part of the country. Cleary this is not satisfying at all, but it seems less arbitrary than comparing Trento to single Regions defined on the basis of mere administrative boundaries. At least, North Eastern regions share common features in terms of socioeconomic contexts (Bagnasco, 1977). We would have liked working on the administrative boundaries, focussing our analyses on schools in towns located immediately in or out the province of Trento. Unluckily, the available data did not allow us to do a proper identification of schools. Hence, we decided to reduce the distance between Northern regions and Trentino, both in terms of socioeconomic contexts and their variability, identifying a control group based on pupils studying in only five neighbouring provinces (Belluno, Brescia, Vicenza, Verona and Bolzano, limited to the Italian speaking subpopulation). Figure 1 shows the location of these provinces compared to Trento (in black): the darker is the shade of grey, the larger is the percentage contribution of each province to the control group sample. In addition, in the map, also all the other provinces belonging to North-East are coloured in light grey.

To conclude the methodological session, we present the main specification of the OLS models presented in the next section:

$$y_i = \alpha + \beta_1 Area + \beta_2 \boldsymbol{X} + \beta_3 \boldsymbol{W} + \varepsilon_i \tag{1}$$

Where represent the outcomes (i.e., Italian and Maths score), Area is a dummy variable that assumes value "1" for the province of Trento and "0" for the counterfactual area (i.e., North-East and neighbouring provinces)[5]. **X** stands for covariates at individual level (parental education and occupational class, migration background and sex) and the year of the survey, while **W** considers class and school size. As stated above, the equity issue will be considered looking at the interaction between the variable "Area" and the individual characteristics:

[5] As a benchmark we will also consider the rest of Italy as a comparison group.

$$y_i = \alpha + \beta_1 Area + \beta_2 \boldsymbol{X} + \beta_3 \boldsymbol{W} + \beta_4 Area \cdot \boldsymbol{X} + \varepsilon_i \quad (2)$$

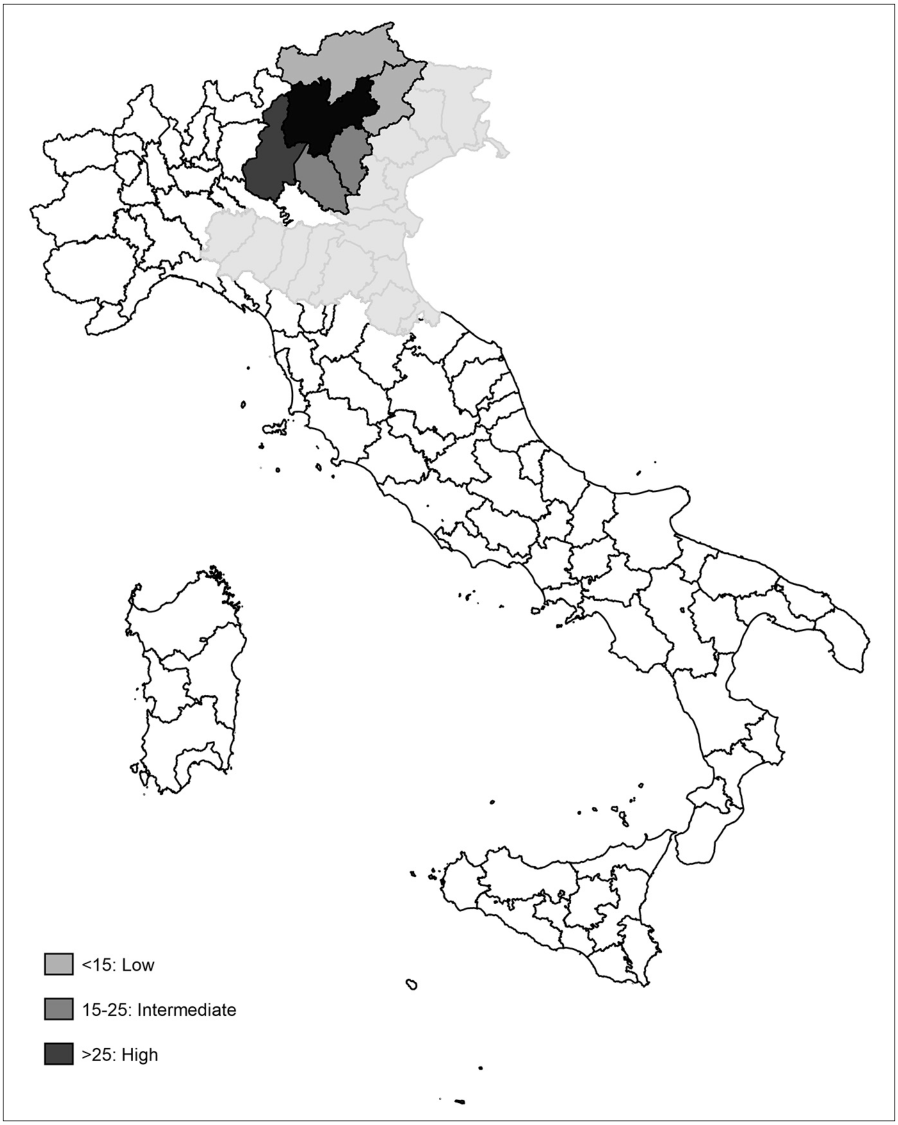

Fig. 1 – Trento province (in black), compared to the neighbouring provinces and to the overall North-Eastern region

Note: neighbouring provinces are coloured accordingly to their percentage contribution to the overall control sample size. The rest of North-East is coloured in pale grey regardless of sample size.

4. Results

In this section we will show the main results emerging from our analysis. We will start presenting a set of descriptive statistics with the aim of stressing similarities and difference of the province of Trento with the neighbouring areas.

4.1. Descriptive evidence

In this subsection we compare the province of Trento with the different areas defined in the methodological section (i.e., the rest of Italy, the North-East and the neighbouring provinces) on the basis of the following indicators: the INVALSI score in Italian and Maths; the unemployment rate and the educational level of the resident population. The INVALSI scores represent the main outcomes of our analyses, while the other two indicators can provide useful information about the characteristics of the contexts under scrutiny. More precisely, unemployment rate is used as a proxy of the economic situation, while the educational level of the population serves as a proxy of the cultural environment. These are two macro characteristics that can affect the students' performance on the INVALSI scores (Tab. 1).

The INVALSI scores confirm what is known in literature (INVALSI, 2019) and that the province of Trento performs much better that the rest of the country (Tab. 1)[6]. The difference shrinks dramatically when Trento is compared with the North-East and with the neighbouring provinces. In fact, for both the scores Trento is just above Brescia, Bolzano and Verona and for Maths the score is also higher than the one observed in the province of Vicenza. These simple descriptive statistics point out that the province of Trento does not seem to outperform even the surrounding areas.

From Table 1 we can notice that the unemployment rate in the province of Trento is lower than in the rest of the country and in the North-East. The picture changes once we look at the neighbouring provinces: only the province of Brescia performs worse than Trento.

[6] It should be stressed that we are using INVALSI census data on the entire student population.

Tab. 1 – Descriptive statistics

Year and indicators	*TN*	*Neighbouring provinces* *BL*	*BS*	*BZ*	*VI*	*VR*	*North-East*	*Italy*
Italian score 2014-2019	200.5	202.9	199.4	192.6	202.3	200.4	200.8	196.1
Maths score 2014-2019	207.9	208.4	202.0	201.6	207.2	204.0	204.7	196.8
Unemployment rate 2014-2019	6.0	5.5	7.1	3.5	5.6	5.7	6.6	11.4
Educational level (2019)								
At least a secondary degree	82.4	78.5	67.3	80.0	73.7	75.2	75.4	71.4
Tertiary degree	28.8	23.5	20.6	21.7	22.2	24.4	25.7	24.6

Legend: TN = Trento; BL = Belluno; BS = Brescia; BZ = Bolzano; VI = Vicenza; VR = Verona.

Note: the descriptive statistics are calculated for the period preceding the COVID-19 pandemic.

Source: the Italian and Maths scores are calculated using census INVALSI data, while the unemployment rate and the educational level of the population has been obtained from the ISTAT data warehouse (http://dati.istat.it). More precisely, unemployment rate comes from the Labour Force Survey, while the educational level of the population from the Census.

The cultural environment is assessed through two indicators: the share of people aged 25-49 with at least a secondary degree and with a tertiary degree. For what concerns these two indicators, the province of Trento outperforms all the other areas of the country considered in our analyses (i.e., Italy, the North-East and the neighbouring provinces). It has to be stressed that in the past the educational level in the province of Trento was lower than the one observed in the rest of the country. Census data[7] shows that the province of Trento filled the gap with the other areas of the country only at the end of the Nineties.

To sum up, the province of Trento shows some differences with the areas of the country that could be mirrored in a compositional difference. For this reason, as explained in the methodological section, in the next subsection a set of controls have been added in the models.

4.2. Results from multivariate models: student attainment

The main aim of this subsection is to understand if the province of Trento performs much better of the other areas of the country following the analytical strategy depicted in the methodological section. The results are presented

[7] See the tables published by ISTAT: https://seriestoriche.istat.it (English version available).

in three different ways[8]. First, we examine the overall results pooling together all the waves (from 2014 to 2021). Second, we analyse the difference between the province of Trento and the neighbouring provinces year by year. Third, we look at the potential heterogeneous effects driven by ascriptive factors such as gender, migration background and parental education.

Table 2 reports a set of models regarding the scores on the Italian and Maths tests[9]. More precisely, we carried out three specific comparisons: Trento *vs* Italy; Trento *vs* North-East; Trento *vs* neighbouring provinces. For each comparison we estimated two nested models (for a total of six models). The first one (Model 1) controls for the year of the survey and for individual covariates, while in the second one (Model 2) are also added the class and school size[10].

Tab. 2 – OLS estimates on the Italian and Maths scores according to different specifications. Selected parameters

	Italy		*North-East*		*Neighbouring provinces*	
	Model 1	*Model 2*	*Model 1*	*Model 2*	*Model 1*	*Model 2*
Italian						
Coeff.	0.076***	0.076***	-0.020**	-0.013	-0.018*	-0.015
SE	0.008	0.009	0.009	0.009	0.009	0.010
Maths						
Coeff.	0.243***	0.240***	0.076***	0.081***	0.067***	0.067***
SE	0.011	0.011	0.011	0.011	0.011	0.011

Legend: Coeff. = regression coefficient; SE = standard error. Standard error clustered at class level.

The first part of Table 2 is devoted to the score on the Italian test. When the province of Trento is compared with the rest of Italy, the regression coefficients are positive and statistically significant across all the specifications, meaning that Trento's students obtain on average better scores on the Italian test that

[8] To facilitate the reading of the chapter, we report only the main parameters, while the full models are displayed in the appendix.

[9] The score tests are standardized. In this way the regression parameters can be interpreted in terms of standard deviations.

[10] The class and school size are considered in an ad hoc model because their meaning is controversial. On one side, they could be interpreted as a policy tool that can be used to foster the learning process; at the same time, they are also the consequences of contextual demographic characteristics peculiar of each territory (i.e., the size of towns and youth concentration among them).

their peers in the rest of the country. The picture changes completely once the comparison is done with the North-East and with the neighbouring provinces. In these cases, the regression coefficients are negative, and they become statistically non-significant once the school and class size are considered (Model 2). The second part of Table 2 shows the results for the score on the Maths test. In this case, it emerges that the students of province of Trento outperform their peers in all the contexts considered independently from the model's specification. It has to be stressed that the difference is very small once the North-East and the neighbouring provinces are used as a comparison group.

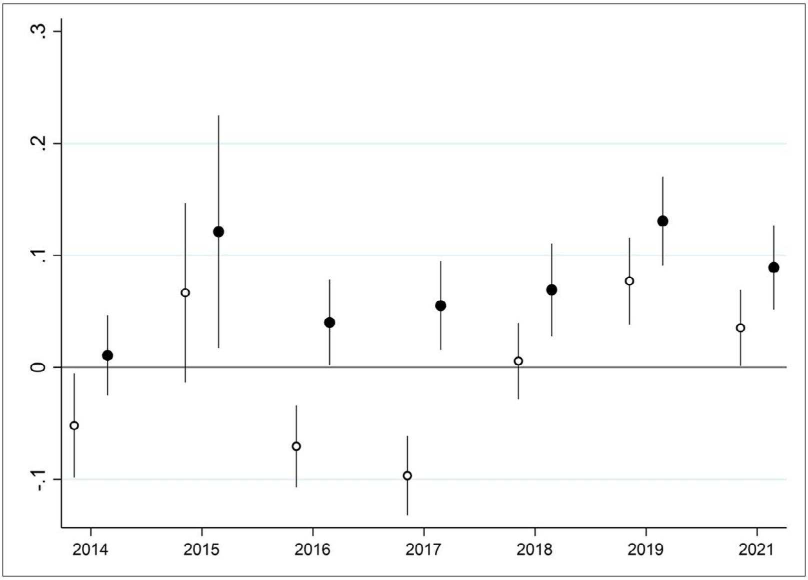

Fig. 2 – OLS models: regression coefficients (and 95% confidence intervals) of Italian and Maths scores for students in the province of Trento compared to students in the neighbouring provinces

Note: filled dots represent effects for the Maths score, while hollow circles represent estimates for the Italian score. The specification used is the one for Model 2.

The second set of analysis is summarised in Figure 2 where the regression coefficients, for Italian and Maths scores, are plotted according to the year of the survey. In this figure, we rely on the specification used in Model 2[11] and

[11] The size of the coefficients is very close in the two specifications. This implies that school-size and class-size are not a relevant factor in determining differences among geographical areas considered here.

we compare Trento with the neighbouring provinces. To us, this is the most relevant comparison to understand if the Trento's peculiarity really exists. The results for the Maths score are consistent over time. In fact, students in the province of Trento show always better results than the ones residing in the neighbouring provinces, despite their advantage is almost negligible in some years and relevant in others. The findings for the Italian scores are less stable: the province of Trento performs worse than the surrounding area in three (2014, 2016 and 2017) out of the seven years considered. In the last two years (2019 and 2021), the results change dramatically highlighting an overtaking of the students from Trento.

4.3. Results from multivariate models: equity

The last part of our empirical work is reserved to the discussion of another crucial aspect of education performance, namely its equity. We focus on the intensity of social inequalities detected in Trento and in the other regions at the end of lower secondary school: lower differences in students' results due to their ascriptive characteristics are here interpreted as a measure of local school systems' capacity to effectively reduce inequalities. More precisely, we consider the following factors: gender, migration background and parental education.

Figure 3 reports the predicted values, estimated according to a set OLS models (see equation 2 in the methodological section), for these factors in the province of Trento and in the neighbouring provinces. If the province of Trento were characterized by greater equity, we should observe a fanning-in conformation in which the distance between the different social groups tends to reduce passing from the neighbouring provinces to Trento. At a first glance, this kind of conformation is not apparent in Figure 3 and the lines representing the various groups seems to be parallel (i.e., the differences between the province of Trento and the neighbouring provinces are not statistically significant). At a closer look, we can notice a small and statistically significant inequalities reduction for the migration background for what concerns the score on the Italian test. At the same time, there is also a small significant increase in the inequalities in the province of Trento. In fact, the distance on the Maths score between students from well-educated families and those whose parents possess only a lower secondary degree is higher in the province of Trento than in the neighbouring provinces.

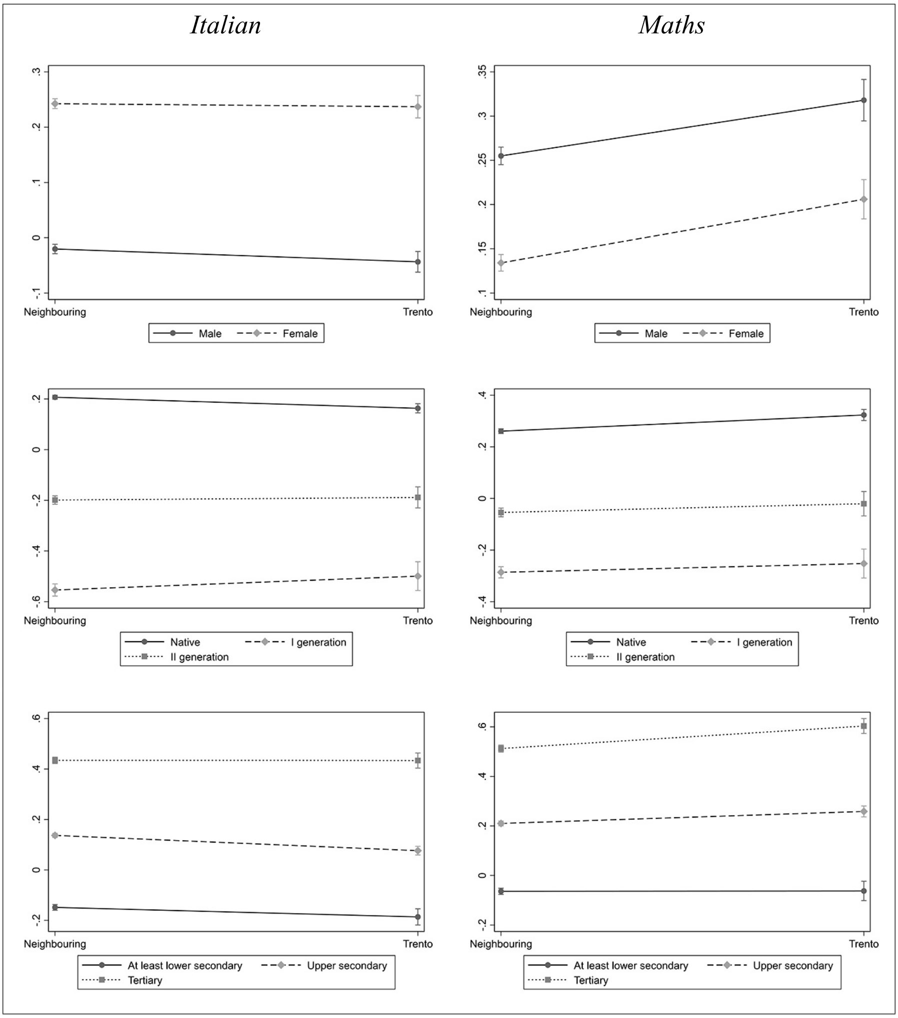

Fig. 3 – Predicted values of Italian and Maths score according to geographical area (province of Trento Vs. neighbouring provinces) and gender, migration background and parental education. The models follow the specification used for Model 2

5. Conclusions

Our chapter intends to draw attention to the potential issues regarding territorial comparison of students' achievement on standardized tests and to the policy implications deriving from the differences between Italian regions and provinces.

As explained in the previous sections, we compare the province of Trento with the whole country and with surrounding areas at the end of lower secondary schools (8th grade students). If Trentino is confronted with the surrounding provinces (Belluno, Vicenza, Verona, Bolzano and Brescia) one realizes that in these areas the average results in Maths and Italian between 2014 and 2019 are not extremely distinctive. As far as the province of Trento is concerned, upon which is focused this research, in some annual statistics for grade 8, the results for the Italian language are actually inferior to those of the surrounding areas, whereas results for Maths are constantly superior although not very significantly. Even in relation to fairness and other factors (gender, family background, migration) Trento does not significantly distinguish itself from the adjacent provinces. As stressed in the methodological section, the best identification strategy should have compared schools in municipalities very closed to the border as done by Battistin and Schizzerotto (2009). This could be the object for future research on this topic to better measure the potential gap in Italian and Maths scores between the province of Trento and the neighbouring provinces. Clearly, this approach would reduce the sample size and further limit the external validity of estimates, but internal validity would be stronger.

A fact is that the territorial areas of the nearby provinces here taken into consideration have better results or are in line with those of North-East Italy, and significantly better than the rest of Italy. From this point of view, one could open an investigation about the excellent performance of this geographical area and we think that this is a relevant further research question for future research.

Taking into account the specificity of the province of Trento depicted in the previous sections, the provincial reflections made concerning the learning of the students and the organization of the scholastic world can be more properly focused. Expanding the look to other circumscribed territories, one can, for example, inquire about learning in the 8th grade so as to maintain a stable level of good results in Italian, without periodic decreases, and how to conserve, maybe by reinforcing, the existing constantly positives ones in Maths.

Our research also shows that in 2021, in the middle of the Covid-19 pandemic, the relative results of Trentino still surpass those of the neighbouring scholastic populations (both in Maths and in Italian). These results are in line with the fact that distance teaching have not caused in Trentino any loss of learning, as instead had taken place in other parts of Italy (Gavosto and Romano, 2021). This positive fact might also be the starting point for a future research project.

As mentioned before, our analyses are based only on the 8th grade, but it has to be stressed that average results of the INVALSI testing study show that in grades 10 and 13 that the learning of the Trentino students is particularly high and superior to those of students in the North-East of Italy. Even the results of the Vocational Education and Training Courses seem better than those of the rest of Italy (Tamanini, Oliviero and Covi, 2021). It would be interesting, therefore, to conduct additional research concerning this phenomenon to understand if a Trentino exception regarding upper secondary education really exists.

References

Argentin G., Barbieri G., Falzetti P., Pavolini E., Ricci R. (2017), "I divari territoriali nelle competenze degli studenti italiani: tra fattori di contesto e ruolo delle istituzioni scolastiche", *Politiche Sociali*, 4 (1), pp. 7-28.

Bagnasco A. (1977), *Tre Italie, la problematica territoriale dello sviluppo italiano*, il Mulino, Bologna.

Battistin E., Schizzerotto A. (2019), "Threat of grade retention, remedial education and student achievement: evidence from upper secondary schools in Italy", *Empirical Economics*, 56, pp. 651-678.

Bratti M., Checchi D., Filippin A. (2007), *Da dove vengono le competenze degli studenti? I divari territoriali nell'indagine OCSE PISA 2003*, il Mulino, Bologna.

CPT (Sistema Conti Pubblici Territoriali) (2019), *Agenzia per la Coesione Territoriale, Le entrate e le spese pubbliche nelle Regioni italiane – Dati aggiornati al 2017*, Temi CPT n. 12/2019

De Simoni M. (2018), *I livelli per la descrizione degli esiti delle prove INVALSI,* retrieved on February, 16, 2022 from www.invalsi.it.

Gavosto A., Romano B. (2021), "The impact of the COVID-19 pandemic on italian schools and universities: the challenge of distance learning", in G. Bellettini, A. Goldstein (*eds.*), *The Italian Economy after COVID-19*, Bononia University Press, Bologna.

INVALSI (2019), *Rapporto prove INVALSI 2019*, retrieved on February, 16, 2022 from www.invalsi.it.

INVALSI (2021), *I risultati delle prove INVALSI 2021*, retrieved on February, 16, 2022 from www.invalsi.it.

Koretz D. (2009), *Measuring Up. What Educational Testing Really Tells Us*, Harvard University Press, Cambridge (MA).

Marcantoni M. (2019), *Storia. Profili storici, istituzionali e finanziari dell'autogoverno Trentino*, IASA, Trento.

Morgan H. (2014), "Review of research: The education system in Finland: A success story other countries can emulate", *Childhood Education*, 90, 6, pp. 453-457.

Ricci R. (2019), *La dispersione scolastica implicita*, retrieved on February, 16, 2022, from www.invalsiopen.it.
Ricci R. (2020), *Risultati INVALSI assoluti o contestualizzati? Servono entrambi, per scopi diversi*, retrieved on February, 16, 2022, from www.invalsiopen.it.
Tamanini C., Oliviero M., Covi L. (2021), "Vocational Education and Training Courses in Trentino Using INVALSI", in P. Falzetti (*ed.*), *INVALSI data to investigate the characteristics of students, school, and society*, FrancoAngeli, Milano.

2. *Effect of economic, social and cultural status of students on English language learning and territorial differences*

by Paola D'Elia, Sabrina Girardi, Caterina Balenzano, Sergio Di Sano

The relationship between socioeconomic-cultural status and school performance is a much-debated research topic of international interest. In Italy, the effect of the status is partially associated with the differences between North and South. From the 2017/18 school year, INVALSI data allows investigating these differences also for the English language, which is the most spoken foreign language in the world, given its wide diffusion. Therefore it is of primary importance to investigate the factors that can favor or hinder its acquisition. The aim of this work is to investigate the effect of socioeconomic-cultural status (ESCS) on English language competence, comparing five Italian macro-regions: north-west, north-east, center, south, south and islands. Furthermore, we investigated whether this relationship is influenced by gender and immigrant status. The results of the INVALSI English tests (reading and listening), administered in the school year 2018/19 to students in grade eight (lower secondary school) were examined. Statistical analyses were carried out both of a descriptive type, comparing the scores on the English tests, and of an inferential type, using regression analysis models. Results showed that the influence of the ESCS-s on the English tests varies according to the type of test (reading or listening) and the macro-regions. Moreover, ESCS-s, gender, and immigrant status impact performances. Overall, the results allow us to reflect on the complexity of the relationship between status and competence in the English language and on strategies to reduce educational inequalities in this area.

Il rapporto tra status economico, sociale e culturale e rendimento scolastico è un tema di ricerca molto dibattuto e di interesse internazionale. Lo scopo di questo lavoro è indagare l'effetto dell'ESCS sulla competenza della lingua inglese, confrontando le cinque macro-regioni italiane. Abbiamo, *inoltre, indagato se questa relazione è influenzata dal genere e dallo status di immigrato. Sono stati esaminati i risultati delle prove INVALSI di Inglese, somministrate nell'anno scolastico 2018/19 agli studenti della classe terza, secondaria di primo grado. Sono state effettuate analisi statistiche sia di tipo descrittivo, sia di tipo inferenziale (analisi di regressione). I risultati hanno mostrato che l'influenza di ESCS sulle prove di Inglese varia a seconda del tipo di prova e delle macro-regioni, e che anche genere e status di immigrato possono influire sulle prestazioni.*

1. Introduction

The relationship between economic, social and cultural status and school performance is a much-debated research topic of international interest. In Italy, socioeconomic differences are partly associated with territorial differences, as the social and economic distance between North and South represents a historical issue, which inevitably extends to the educational level. Results in the INVALSI tests constantly confirm this evidence (Di Sano and Balenzano, 2021). The territorial differences between the northern and southern regions are mainly attributable to contextual factors, such as family characteristics – the level of education and the profession of parents –, the socioeconomic peculiarities of the territory, the quality of resources, and school structures (Bratti *et al.*, 2007; Riccardi, Donno and Bagnarol, 2020; Russo *et al.*, 2020).

For many years, the analyses on the influence of social and territorial status on learning outcomes were conducted on the Italian and Mathematics tests only, as the English tests (reading and listening) were included in the INVALSI tests in the 2017/18 school year. The English tests are of particular interest both as it is a skill in which Italian students are lacking, and as they are influenced by factors that are partly different from the Italian and Mathematics tests.

For the English tests, as for the Italian ones, there is a gender difference in favor of females. The main difference, compared to the results of the Italian and Mathematics tests, is that in the English tests, the immigrant students (first and second generation) have the same or better results than the Italian students (INVALSI, 2019). In this study, we intend to investigate the influence of cultural socioeconomic status (ESCS) on performance in English tests (reading and listening) comparing the five macro-regions in which the INVALSI divides Italy. Furthermore, we aim to investigate the impact of the gender and origin of the students on this effect.

2. Understanding educational inequalities

2.1. A multidisciplinary perspective

The role of the birth family in influencing children's life chances and educational outcomes is a research topic considered across various disciplines. When examining factors related to student academic outcomes, much research has demonstrated the association between family socioeconomic status and children's school success. Regarding educational inequalities, in particular, scholars generally agree that the socioeconomic status of the family has a significant impact on student engagement and motivation in school learning. It influences children's participation in school activities, the availability of educational resources, the school environment, and the support received with homework (Chiang, 2018). Furthermore, we know that parents' beliefs reflect their status and influence their children's academic attitudes. However, the mechanisms by which family of origin resources influence children's academic achievement are not well understood (Kaiser and Pollmann-Schult, 2019). So, researchers need to better explain how some characteristics of the family of origin lead to unequal educational opportunities.

Sociologists studying intergenerational reproduction of educational inequalities focus their attention on the role of differences in socioeconomic resources; according to this perspective, educational outcomes are influenced by differences in the availability of financial, cultural, and social resources, and each of these resources influences educational attainment through distinct mechanisms (Bourdieu, 1986). However, the sociological perspective may overestimate the role of the socioeconomic resources of the family of origin, as scholars who adopt this approach do not consider the role of parents and student's cognitive ability, although it has been shown that parents transmit both their socioeconomic resources and genetically influenced skills to their children (Shultz *et al.*, 2017).

On the other hand, the psychological literature has pointed out that parents affect children's cognitive abilities both genetically and through the resources they offer; in turn, children's cognitive ability is an important predictor of educational attainment (Plomin *et al.*, 2016; Morris *et al.*, 2016). However, approaches that directly consider both parental socioeconomic resources and parental ability are unusual, as the contributions from the different perspectives are largely independent.

Another line of research investigated the role played by the parent-child relationship and motivation. In this context, Chen *et al.* (2018) showed that the parent-child relationship mediates the relationship between SES and reading skills; furthermore, this relationship is moderated by students' motivation to learn. In this sense, low-status parents can support the development of their children's learning skills in the context of a fruitful educational relationship. Furthermore, family and school, by promoting learning motivation in children, can counteract the potential negative effects of low socioeconomic status. Along the same lines, a systematic review by Devenish *et al.* (2017) investigating the possible pathways leading from low socioeconomic status to negative outcomes in adolescent development. Factors mediating the effect of risk factors associated with socioeconomic status include the following: parental depression, parental conflict, parenting practice, and adolescent resilience.

Other important aspects to consider in analyzing the relationship between Socioeconomic Status (SES) and school outcomes concern the influence of the neighborhood and the school climate. Ruiz *et al.* (2018) showed that the relationship between SES and educational achievement is mediated by the neighborhood (with reference to economic and safety aspects); at the same time, however, the school climate moderates this relationship. So, even if the low SES, in the context of a neighborhood characterized by high rates of poverty and violence, can favor lower school results, the presence of a positive school climate could counter these effects. Other research confirms that the school climate is of particular importance to adolescent mental health and learning and reduces the negative impact for schools targeting communities with high poverty rates (La Salle *et al.*, 2021; Berkowitz, 2021).

Moreover, while most scholars use family income and/or parental education as the most easily detectable proxy measure of the socioeconomic resources that students can use, these are often not reliable measures of the cultural opportunities offered to children. We think also that the relationship between family background and student outcomes likely differs across subjects and the opportunity to access additional resources for learning English (e.g., English courses, study trips abroad) is very important to stimulate the academic motivation of English students and improve educational outcomes. This hypothesis is consistent with the results of the Liu and Chiang (2019) study, according to which educational inequality is particularly significant for English tests, as students from advantaged family backgrounds have higher levels of learning motivation than students from disadvantaged family backgrounds even after controlling student-teacher interaction levels. In other words, students from different family backgrounds may be motivated differently to study English, possibly because learning English profoundly reflects class-based privileges. Parents, in fact, play an important role both in the choice of when their child begins to learn a foreign language and in the motivational processes of learning a foreign language (Szabó, Albert and Csizér, 2021).

Based on these results, various intervention paths have been developed to combat educational inequalities. Alongside the themes of the defense of human rights and multiculturalism, the perspective of social justice has increasingly developed in school psychology, integrating the questions of the defense of human rights and multiculturalism. The idea is that the school psychologist should pay particular attention to marginalized students tackling discrimination and combating educational inequalities. According to Shriberg *et al.* (2020), research on social justice, in the field of school psychology, converges on three main ideas: a) guaranteeing rights, and thorough compliance with the provisions of the Convention on the Rights of the Child; b) guaranteeing access, allowing all children to use the resources of society; c) ensuring respect, combating discrimination, and promoting participation and active involvement.

2.2. Education and Inequalities: sociological investigations

The world of education, as we have seen, moves on a plurality of concepts and on the principles of justice, so much to stir the interest of the sociology of education which has dedicated part of its studies precisely to inequalities of education.

Resuming the Bourdieusian conceptual apparatus, in fact, today we are faced with two lines of research and a change of perspective on the part of sociologists with respect to what are the causes that increase educational inequalities.

Researchers argue that many inequalities in children's educational progress are related to family context and attribute the root cause to social class and to the intergenerational and scholastic transmission of cognitive skills (Bonal and González, 2020). These disparities are already present in kindergarten: according to the Fifteenth UNICEF Report (2019), this gap persists throughout the child's educational path. Academic success comes from a series of behaviors that allow upper-class families to "claim" their class position. 15-year-olds, whose parents are in high-level work, are favored by the education system and are much more likely to pursue higher education, unlike working-class or low-profile children. The studies undertaken during an individual's educational journey will impact their future, influencing selection, assimilation, and confirmation of their social status.

A second line of research, on the other hand, also followed at the European level, has focused its attention on the sociopolitical context, arguing that a potentially important factor influencing equity in education is the level of educational and social segregation between schools. An evaluation in terms of scholastic equity is also precisely considered the social composition of the territory which is sometimes binding for the choice of school and class (Benadusi, 2021). The availability of different types of schools, the choice of school and its admission policies, the choice of educational path, and even the policies and practices related to the repetition of school years are intermediate variables between some structural characteristics of the educational systems and equity levels (Eurydice, 2020).

These inequalities, to date, are then exacerbated by the advent of the pandemic, which has jeopardized the role of schools as a means of guaranteeing social justice and equality (Drane *et al.*, 2021), increasing the gap in the conditions of educational activities (Bayrakdar and Guveli, 2020; Cullinane and Montacute, 2020).

2.3. Social Inequalities in the Italian Context

It is therefore clear that the school, as a social institution, plays a fundamental role in society's equal educational opportunities. This role also implies ensuring the transmission and acquisition of knowledge to all students, but it is seriously threatened by the multiple forms of segregation, disengagement, and expulsion that run through our educational system.

In cases where school differences account for a greater percentage of the variation between student achievement scores (or between their socioeconomic backgrounds), education systems are more segregated and less inclusive, at the school and social levels (OECD, 2019).

The analysis of educational inequalities between social classes in Italy, according to welfare models, has shown both the linear cuts suffered by the state school system in the last 15 years or so, and the expansion of the recruitment of services for early childhood. This depends on the political-economic character that establishes which and how many resources schools in the various regions must make available to students.

To date, the Italian education system still has some features of strong and marked inequality; this is without considering that the internal geographical divisions, the social and intergenerational disparities and the effects of the migratory background are strong and evident (Giancola and Salmieri, 2020). The latest research focuses precisely on the formation of classes which can be of fundamental importance: level groups amplify inequalities in results, while those that are equally heterogeneous, if managed with appropriate teaching methods, reduce them (Benadusi, 2021b). Added to this is that within the Italian school system there has been a tendency to reproduce gender differences in both content and form. This phenomenon, in fact, has always reported a difference in the academic success of boys and girls from different areas of Italy, demonstrating how girls, driven by greater motivation and interest, achieve better results than boys (Bianco, 2017).

2.4. Contextual and motivational factors in learning a foreign language

Although learning a second language seems a very complex activity influenced by social and psychological aspects, according to Gardner *et al.* (1985), motivation is one of the most influential factors in learning a new language. Besides the positive attitudes, such as desire and interest in learning a second language, motivation implies an effort to engage in regular activities. Csizér and Dorney (2005) conceptualized the generalized aspects of motivation to learn a second language identifying seven components: Integrativeness, Instrumentality, Vitality of the L2 Community, Attitudes toward the L2 Speakers/Community, Cultural Interest, Linguistic Self-Confidence, and Milieu. These dimensions were interrelated and combined into a structural model (Csizér and Dorney, 2005). The authors examined the influence of each domain on Language Choice and Efforts, concluding that Integrativeness may explain stronger the motivational behavior of students in the acquisition of a second language. According to Gardner (2001), Integrativeness may be defined as the emotional identification with another cultural group, which may reflect confidence with the language and the associated culture; students may desire to integrate and adapt themself to the culture and conform to the speakers (Csizér and Dorney, 2005). Although individual aptitudes are fundamental in defining the process of learning a second language, the extent to which students are open to a different culture, their willingness to adapt to it, and their attitude towards the second language could be determinants. By contrast, the individual's bias toward a different cultural group may shape the prejudice that can hinder the efficacy in the learning process of the language (as prejudices towards the culture may reduce the motivation towards the language). In this perspective, Integrativeness could reflect the willingness to engage with members of another lingo culture. Nicol and De France (2020) highlighted how the main indicators of prejudice (social dominance orientation; right-wing authoritarianism) are negatively correlated to Integrativeness, suggesting deeply examining the role of prejudice in the acquisition of a second language. According to the authors, it may be beneficial to promote successful learning of a second language to develop intercultural skills and encourage acceptance and tolerance.

In accordance with Mori and Gobel (2006), Integrativeness may also clarify gender differences in second language acquisition, explaining females' persistence in learning, homework completion, and high performance. A large body of research shows the discrepancy in the second language learning achievement between males and females (Clark, 1995; Bacon and Finnemann, 1992). Mori and Gobel (2006) highlighted female students have higher motivation and more positive attitudes toward foreign languages. Females show a bigger interest in culture, people, and the target language, revealing a greater desire to make native speakers friends, moreover, they are more interested in traveling and studying abroad compared to their counterparts (Mori and Gobel, 2006).

From the perspective of second language acquisition processing, the self-determination theory (SDT) represents a framework for describing the motivational learning dynamics (Alamer and Lee, 2019; Cho, 2021; Shelton-Strong, 2020). There are basic psychological needs (BPN) that are to be met for constant and deepened motivation: autonomy, competence, and relatedness (Ryan and Deci, 2017). In the perspective of second language acquisition, the need for autonomy refers to genuine interest and fulfillment while experiencing activities related to the subject. In this viewpoint, learners are allowed by the social context to choose the most relevant language tasks or games to keep engaged. The need for competence refers to learners' ability to manage their learning process feeling effective. They must have a clear structure and a set of objectives they are able to reach. The need for relatedness refers to the sense of belonging as part of a community. In this way, students feel connected to other people, cared for, and important (Alamer and Lee, 2019).

Let us now consider the role of socioeconomic-cultural status. What influence can it have on learning English as a second language?

According to Keumala *et al.* (2019), the sociocultural and educational background of students influences their motivation to learn English as a foreign language, although the effect depends on the type of culture (individualistic or collectivistic). In particular, since the second language is also associated with a different culture, students living in an individualistic society may see a new culture more positively than those living in a collectivistic society (Keumala *et al.*, 2019).

Furthermore, as Razi and Rahmat (2020) note, there are cultural barriers that hinder learning English as a second language which reduces the motivation to learn. A person living in his country of origin may have an interest in learning English in order to visit new places, and this motivation may be stronger in students of higher socioeconomic status, who have more opportunities to be able to go abroad, in some cases to complete a training course. In this sense, for Italian students, the socioeconomic-cultural status could be associated with greater competence in the English language.

Could this same argument also apply to a foreign person who comes to live in Italy? A line of research on the acquisition of English as a second language has investigated the social aspects of this competence, which involves cultural integration processes (Douglas Fir Group, 2016; Duff, 2019). Immigrants arriving in Italy are faced with a double task: having to learn a new language and having to integrate into a new social context. This double difficulty can lead to problems for them in the INVALSI tests of Italian and Mathematics. For the English tests, first and second-generation immigrants do not show any difficulties, and indeed in some cases, they are even better. A person who must emigrate abroad, for economic or humanitarian reasons, could have a strong motivation to learn English, which could be associated to a lesser extent with social status. In this sense, for immigrant students, the socioeconomic-cultural status may be less important as an influencing factor for the knowledge of English in reading and listening tests.

3. The research question: macro-regions comparison, social inequalities, and English learning achievement

The scarcity of financial resources has a stable impact on academic engagement, educational outcomes, and individual aspirations (Lohmann and Ferger, 2014). Historically, Italy is characterized by a territorial complexity, which is reflected in the fragmentation of the Italian school system, where the southern regions suffer from severe delays.

Starting from this premise, we hypothesize that schools in southern regions have less capacity to face and mitigate economic and sociocultural differences among students and that the drawback may have a higher impact on English language achievement.

The socioeconomic-cultural status of the student (ESCS-s) might have an influence on English test performance and this effect may vary according to the geographical areas, as also happens for the Italian and Mathematics tests (Di Sano and Balenzano, 2021) and may differ according to the type of tests (reading/listening) (Martini, 2020). Moreover, we can hypothesize an influence of gender and immigrant status of students. Highlighting the role of motivation in the complex process of learning a new language, we also assume females and immigrants might perform better in both reading and listening tests of English.

Therefore, this contribution aims to:

a) Compare the different groups of Italian regions (north-west, north-east, center, south, south and islands);
b) Investigate the influence of the ESCS-s indicator on the English reading (R) and listening (L) tests;
c) Relate the differences on the INVALSI tests of English to the origin and the gender of the students and compare the different sets of regions.

The choice of grade 8 (third grade of lower secondary school) as a school group is based on the following reasons:

a) Greater interpretability than those of upper secondary school, as there is the advantage of not having a differentiation by type of school (high schools/"licei", technical and vocational institutes);
b) Greater reliability of the data compared to those of primary school as Computer Based Tests (CBT) is less influenced by the phenomenon of cheating (INVALSI, 2019), instead of the paper-based tests used in primary education;
c) Greater territorial variability compared to primary school data: in primary school, the territorial gap between the geographical areas of the country does not reach statistical significance;
d) Educational equity among the national territory: the common education pathway ends with middle school ensuring a uniform base of fundamental skills to guarantee the equality of educational opportunities.

4. Method and analysis plan

4.1. Sample

Analyses are based on the sample data of grade 8 students who participated in the INVALSI English, reading (R), and listening (L) tests for the school year 2018/2019. The total number of students is 558,922 (Northwest: 146,509; Northeast: 104,066; Center: 106790; South: 114,899; South and Islands: 86,658). Weighted sample data were used to consider the representativeness of the sample. We have chosen to analyze the sample data for their higher quality, as they are collected in the presence of an external observer who monitors the correctness of the administration procedures.

4.2. Variables, procedure, and methodological choices

The research examined the results of the INVALSI Tests of the National Survey of English listening and reading tests related to the 2018/19 school year. The English test involves the assessment of two language comprehension skills: reading and listening. Writing and speaking skills are not assessed by the INVALSI test. As required by the CEFR, the level of competence that students must reach at the end of the first cycle of education (grade 8) is A2. For the reading test, a text of a maximum of 110 words is provided for level A1 and a text of a maximum of 220 words is provided for level A2. For the listening test, on the other hand, audio files are provided for both level A1 and level A2. These usually consist of a monologue or a dialogue between 2 or a maximum of 3 people lasting no more than two minutes, or a sequence of small monologues lasting a few seconds with different speakers. The audio file is always played twice. The types of questions for the English test are Multiple Choice Questions, Multiple Matching, and Short Answer Questions.

Specifically, the research examined the scores in the tests of English reading (WLE ERE) and listening (WLE ELI) according to the model of Rasch (1980), which allows putting on the same continuum for the difficulty of the items and the measured ability. These scores are expressed on a scale with a mean of 200 and a standard deviation of 40 and were used as dependent variables in the tested models.

The following factors were considered as independent variables:

- The student's macro-region (Northwest; Northeast; Center; South; South and Islands);
- The student's socioeconomic and cultural status index (ESCS-s) based on three indicators: a) the employment status of parents; b) the level of education of parents; c) the availability of a series of home resources, such as a quiet place to study, a personal desk for doing homework and an internet connection. This index is standardized and has a mean of zero and a standard deviation of 1 (Campodifiori *et al.*, 2010);
- The student's origin (1: native; 2: immigrant);
- The student's gender (1: male; 2: female).

Data were preliminarily weighted calculating the final weight of the student (or sample weight), which indicates how many pupils not involved in the sample survey are represented by the pupil participating in the surveys (Ricci, Falzetti and Falorsi, 2019).

Based on art. 7, c. 1 of Legislative Decree 62/2017 the INVALSI tests for eighth-grade students were CBT (computer-based) administered, taking place using computers connected to the online tests over a period (administration window) assigned to the school by INVALSI. Tasks consisted of exercises extracted from a repertoire (bank of items) and therefore vary from student to student, maintaining the same difficulty and structure for each form.

In order to ensure the representativeness of our findings and address the diversity within the population, we chose to estimate the regression models using population weights. Firstly, it allows us to account for the variations observed across the five macro-regions, ensuring that our results are more reflective of the entire population. By incorporating weights, we can obtain more accurate parameter estimates, thereby increasing the generalizability of our conclusions.

Moreover, population weights help to mitigate any potential bias resulting from disparities in participant distribution. By assigning appropriate weights, we can address the uneven representation of participants across different regions, resulting in more reliable estimates of regression parameters. This approach enhances the robustness of our analysis and provides a more comprehensive understanding of the relationship between the variables under investigation.

Additionally, considering population sizes within each macro-region through weights further contributes to the validity of our findings. By assigning proportional weights to observations from regions with more significant participant numbers, we ensure that these observations have a suitable influence on the analysis, maintaining the representativeness of our results.

Lastly, including population weights helps minimize the impact of outliers or extreme values that may distort the analysis. By accounting for the overall data distribution and assigning appropriate weights, we obtain more accurate estimates and reduce the potential bias introduced by extreme observations.

4.3. Models of analysis

Three types of analysis were applied. First, through descriptive analysis, the five groups of regions (northwest, northeast, center, south, and south and islands) were compared in the average scores for the English listening and reading tests. A second analysis compared the percentages of students belonging to each of the four quartiles to the three-competency level in both reading and listening English tests. Thirdly, a linear regression analysis was carried out to examine the impact of microregion on INVALSI English test scores. Furthermore, a multiple regression analysis was utilized to explore whether the variance attributed to macro-region could be affected by including the following variables within the statistical model: i) ESCS, ii) origin, and iii) gender. The regional groups were collectively incorporated into a unified regression analysis for listening and reading outcomes alongside ESCS, origin, and gender.

For the analysis, we used SPSS statistical software. The method ENTER was used as a procedure for the regression analysis.

Tab. 1 – Summary of the tested models

Models	*Predictor variables*
Model 1	Macro-regions (Northwest, Northeast, Center, South, South and Islands)
Model 2	Macro-regions + Student level ESCS
Model 3	Macro-regions + Student level ESCS + origin
Model 4	Macro-regions + Student level ESCS + origin + gender

5. Results

5.1. ESCS-s and tests results

The average scores on the English reading and English listening tests are compared to the students' groups of ESCS quartiles, considering each macro-region as shown in Tables 2 and 3.

Tab. 2 – Score on the English reading test averaged by ESCS quartiles

Reading (R)	*I Quartile*	*II Quartile*	*III Quartile*	*IV Quartile*
Northwest	195.6	208.3	214.7	224.6
Northeast	200.6	211.3	216.0	224.4
Center	192.7	203.8	211.6	221.2
South	177.5	193.1	203.5	216.4
South/Islands	171.2	186.7	195.3	211.0

Tab. 3 – Score on the English listening test averaged by ESCS quartiles

Listening (L)	*I Quartile*	*II Quartile*	*III Quartile*	*IV Quartile*
Northwest	197.4	208.6	215.1	223.2
Northeast	204.1	211.9	216.5	224.8
Center	194.1	204.1	207.9	219.3
South	175.8	189.8	196.4	208.7
South/Islands	171.0	184.4	191.3	205.5

The results show wide performance variations based on the quartiles of ESCS for both tests. Moreover, we observed that the disparity in English test scores between students in the highest quartile of ESCS (Economic, Social, and Cultural Status) and those in the lowest quartile is more prominent in the southern macro-regions.

5.2. Competency level and ESCS Quartiles

The grade 8 English test, for the 2018/19 school year, allows students to be classified into three distinct levels of competence, based on the Common European Framework of Reference for Languages (CEFR): Pre-A1, A1, and A2 both for listening and reading. The comparison between the percentages of students belonging to each of the four quartiles for the three competence levels in the English reading and English listening test, for every macro-region, are shown in the Tables below (4-8).

Tab. 4 – Average percentages of students belonging to each quartile of ESCS in the Northwest, comparing competency levels in reading and listening

		ESCS				
		I Quartile	*II Quartile*	*III Quartile*	*IV Quartile*	*Tot*
Pre-A1 (N 2143)	Reading	62.7	17.2	12.8	7.3	100.0
Pre-A1 (N 1200)	Listening	63.0	28.8	3.3	4.8	100.0
A1 (N 18523)	Reading	45.6	25.1	18.9	10.3	100.0
A1 (N 39450)	Listening	40.6	35.8	11.6	12.1	100.0
A2 (N117560)	Reading	22.6	24.2	26.2	27.0	100.0
A2 (N 98387)	Listening	19.8	34.8	14.7	24.5	100.0

Tab. 5 – Average percentages of students belonging to each quartile of ESCS in the Northeast, comparing competency levels in reading and listening

		ESCS				
		I Quartile	*II Quartile*	*III Quartile*	*IV Quartile*	*Tot*
Pre-A1 (N 986)	Reading	54.2	16.1	21.5	8.2	100.0
Pre-A1 (N 570)	Listening	58.8	15.4	17.7	8.1	100.0
A1 (N 11320)	Reading	39.9	26.8	20.0	13.4	100.0
A1 (N 23501)	Listening	33.8	38.0	12.0	16.1	100.0
A2 (N 86342)	Reading	20.3	24.1	26.6	28.9	100.0
A2 (N 75420)	Listening	19.3	35.5	15.8	29.4	100.0

Tab. 6 – Average percentages of students belonging to each quartile of ESCS in the Center, comparing competency levels in reading and listening

		ESCS				
		I Quartile	*II Quartile*	*III Quartile*	*IV Quartile*	*Tot*
Pre-A1 (N 2130)	Reading	57.1	26.3	10.9	5.7	100.0
Pre-A1 (N 1306)	Listening	46.0	29.8	11.3	12.9	100.0
A1 (N 16167)	Reading	37.1	25.3	23.1	14.5	100.0
A1 (N 33268)	Listening	35.0	34.3	14.8	15.9	100.0
A2 (N 81567)	Reading	20.5	22.0	25.2	32.2	100.0
A2 (N 66083)	Listening	18.0	30.8	15.8	35.4	100.0

Tab. 7 – Average percentages of students belonging to each quartile of ESCS in the South, comparing competency levels in reading and listening

		ESCS				
		I Quartile	*II Quartile*	*III Quartile*	*IV Quartile*	*Tot*
Pre-A1 (N 6411)	Reading	68.0	18.1	8.1	5.7	100.0
Pre-A1 (N 4927)	Listening	63.0	23.0	4.8	9.2	100.0
A1 (N 26310)	Reading	48.5	25.0	17.3	9.2	100.0
A1 (N 55244)	Listening	41.5	32.9	12.1	13.5	100.0
A2 (N 77542)	Reading	25.0	21.6	25.0	28.3	100.0
A2 (N 50895)	Listening	20.4	30.6	15.8	33.2	100.0

Tab. 8 – Average percentages of students belonging to each quartile of ESCS in the South and Islands, comparing competency levels in reading and listening

		ESCS				
		I Quartile	*II Quartile*	*III Quartile*	*IV Quartile*	*Tot*
Pre-A1 (N 7275)	Reading	59.2	23.1	12.9	4.8	100.0
Pre-A1 (N 4770)	Listening	56.2	28.5	10.5	4.8	100.0
A1 (N 23672)	Reading	47.3	24.7	18.8	9.2	100.0
A1 (N 45877)	Listening	42.0	33.6	11.9	12.5	100.0
A2 (N 51397)	Reading	24.3	24.0	25.0	26.7	100.0
A2 (N 32479)	Listening	19.8	31.8	16.8	31.5	100.0

The results show that in all five macro-regions, over 50% of students with a low proficiency level (Pre-A1) belong to ESCS quartile 1 (low status); except in one case (Pre-A1 listening of the Central macro-region), where this percentage is 46%. If, on the other hand, we consider the highest level of competence (A2), we see that, at least for the reading test, students in quartile 4 (high status) are present with a higher percentage than students belonging to the other three quartiles. For the listening test, the incidence of students in quartile 4 of ESCS is also higher in percentage among students of A2 level and lower in percentage among students of Pre-A1 level – but the differences are less pronounced than in the reading test. Generally, for both tests, we see a high percentage of students at the A2 level in the fourth quartile and a large percentage of students at Pre-A1 in the first quartile.

5.3. Results of the Regression Analysis

This study examines the regional disparities in English language proficiency, investigating both English listening and reading among Italian students, focusing on the influence of the socioeconomic indicator (ESCSs), origin, and gender. The Italian regions were categorized into five macro-regions based on territorial areas: northeast, northwest, central, south, and south and islands. We used linear regression analysis to explore the relationship between these predictors and English language proficiency (listening and reading).

Tables 9a and 9b show the results of a regression analysis conducted to examine the relationship between the listening (L) and reading (R) test competence (dependent variables) and the independent variables. The tables showcase four distinct models (Model 1, Model 2, Model 3, and Model 4) and offer a comprehensive exploration of the independent variables' effect on the listening and reading tests. Model 1 includes the independent variable macro-region (northwest, center, south, south and islands), while Model 2 adds ESCS-s as predictor variable; Model 3 further extends the analysis by including the predictor variable of origin; finally, Model 4 encompasses all the predictor variables: macro-regions, ESCS-s, origin, and gender. By examining the coefficients associated with each independent variable in the models, we gain insights into their impacts on the listening and reading tests. Additionally, the R-squared values indicate the proportion of weight in the listening and reading tests that the model explain.

The northeast macro-region served as the reference point; thus, the coefficients reflect the disadvantage of residing in other macro-regions compared to the reference one.

Tab. 9a – Regression analysis listening (L) test

	Model 1	*Model 2 + ESCS-s*	*Model 3 + origin*	*Model 4 + gender*
Constant	216.21**	214.8**	209.28**	198.42**
Northwest	-5.29**	-4.68**	-4.38**	-4.34**
Center	-8.32**	-8.64**	-8.3**	-8.27**
South	-23.06**	-21.31**	-20.31**	-20.27**
South and Islands	-28.48**	-25.91**	-24.82**	-24.80**
ESCSs		10.73**	11.20**	11.18**
Immigrant			4.25**	4.25**
Female				7.26**
R-squared	.099	.192	.196	.207

** Significance at level 0.01

Tab 9b – Regression analysis reading (R) test

	Model 1	*Model 2 + ESCS-s*	*Model 3 + origin*	*Model 4 + gender*
Constant	214.99**	213.25**	211.66**	199.71**
Northeast	-3.83**	-3.01**	-2.92**	-2.87**
Center	-5.43**	-5.79**	-5.70**	-5.66**
South	-16.6**	-14.48**	-14.19**	-14.13**
South and Islands	-23.1**	-20.11**	-19.79**	-19.77**
ESCSs		12.54**	12.67**	12.65**
Immigrant			1.23*	1.22*
Female				8**
R-squared	.055	.169	.169	.181

** Significance at level 0.01; * Significance at level < 0.05

The linear regression analysis was performed to examine the weight of the predictors. The results indicate the estimated coefficients for each predictor variable. These coefficients represent the magnitude and direction of the relationship between each predictor and the outcome variable.

Results of the linear regression models indicate a statistically significant predictive relationship among all variables.

Moreover, results reveal notable variations among regions: northwest, center, south and south and islands. Among these groups, the northeast region performs highest in both reading (R) and listening (L) tests.

Model 1 includes the scores for listening and reading in the INVALSI English test as dependent variables. It is observed that students from the

North attain slightly higher performances. Moving geographically from the North to the South and Islands, the coefficients explaining the variability progressively decrease. The regression tables (Tabs. 9a and 9b) clearly illustrate the distinction between Northern Italy and Southern Italy. Model 2 incorporates the socioeconomic status (ESCSs) indicator as an additional variable, with a greater predictor effect observed in reading (12.54) compared to listening (10.73). Additional variables were incorporated into the analysis to comprehensively investigate the factors influencing the students' score. Model 3 introduces the variable origin of the student as a potential determinant to explore any noticeable variations in scores among macro-regions. On the other hand, Model 4 incorporates the variable gender to assess whether it plays a significant role in explaining the students' score, in order to gain further insights into the complex interplay between various factors and the students' performances. In Model 3, adding the variable origin of the student, there is a difference in the coefficients between the reading test (1.23) and the listening test (4.25). The last Model 4, adding the variable gender, does not explain its effect on the students' score.

Generally, when considering the variable origin (referring to immigrants), the multiple regression analysis indicates that immigrants tend to achieve higher scores on the listening tests than Italian students. Additionally, including the variable gender reveals a notable pattern: irrespective of the regional categorization, female students consistently score higher than their male counterparts in the reading and listening tests.

6. Conclusions

The main aim of this study was to explore the relationship between social, economic, and cultural status (ESCSs) and performance levels in the INVALSI grade 8 English reading (R) and listening (L) tests for the school year 2018/19 within the context of the five macro-regions. Additionally, the study sought to investigate any potential associations between gender, origin, and test scores. Moreover, the study explored the potential influences between gender, origin, and test scores, aiming to assess variations in these aspects across the different macro-regions. By considering these regional differences, we sought to gain a comprehensive understanding of the interplay between ESCSs, test performance, and demographic factors across distinct geographical areas.

In general, as expected, northern macro-regions performed higher in both reading and listening English tests. Moreover, students of higher economic-social-cultural status obtained better results than lower-status students.

The combination of historical disparities in educational opportunities and socioeconomic factors has resulted in distinct regional differences in educational outcomes across the Italian Peninsula. The northern macro-regions of Italy, encompassing the Italian northeast, northwest, and central regions, have consistently exhibited higher levels of educational attainment and stronger socioeconomic indicators. These factors are believed to contribute to the above test performances observed in these regions. In contrast, the southern regions, including the south and islands, face significant social and economic development obstacles, which are reflected in comparatively lower average test scores. Remarkably, differences in status have a more significant effect on English test performances in the southern regions compared to the northern regions. Results from the first analysis confirm the Italian Peninsula's territorial complexity, highlighting the persistent difficulty of southern Italy facing internal social and educational disparities.

In the second analysis, instead of weighted scores (based on Rasch's analysis), we considered the three distinct competency levels in the English language (pre-A1; A1; A2). Results show the same trend for both reading and listening. At the end of lower secondary school (grade 8), the expected achievement is the CEFR A2 level. None of the macro-regions reached this expected result significantly. Nevertheless, the acquisition of level A2 is more frequent in the north than in the south. As for the relationship between ESCS and English proficiency level, results show wide performance variations based on the quartiles of ESCS for both tests. The percentage of students with low ESCS decreases with increasing competency level; vice versa the percentage of students with high ESCS increases with increasing competency level. Moreover, we observed that the disparity in English test Levels of Competence between students in the highest quartile of ESCS (Economic, Social, and Cultural Status) and those in the lowest quartile is more prominent in the southern macro-regions. This disparity may be due to socioeconomic factors, educational infrastructure, regional language influence, cultural attitudes, and migration demographics. Considering the importance of English proficiency for international interactions and economic trade, these findings raise concerns as they can potentially impact the social and economic progress and growth of the southern regions.

The third analysis used multiple regression to examine the predictors of performance differences on the two English tests, considering the influence of macro-region, socioeconomic status (ESCS), gender, and origin. The results confirmed that differences in performances vary across the distinct geographic areas under investigation (macro-regions, namely northeast, northwest, center, south, south and islands). Moreover, according to the analysis, the socioeconomic-cultural status of the student (ESCSs) is an essential factor explaining some of the variability in the INVALSI tests of English.

The addition of the gender and origin variables in the regression models does not reduce the variability explained by ESCS, confirming that the three factors operate independently across all the macro-regions for both listening and reading tests. Regarding the origin of the students, the analysis revealed that immigrants (first and second generation) achieved higher scores compared to their Italian counterparts, mainly in the listening test. This difference could be attributed to the distinct nature of the two tests: the reading test appears to be more closely associated with school experiences. In contrast, the listening test reflects everyday experiences in non-school contexts. It is worth noting that the limited presence of native English-speaking teachers in Italian schools may also play a role in this discrepancy.

In line with the literature discussed, females outperform males in reading and listening tests, regardless of the macro-region under consideration. This performance difference can be attributed to various factors such as cognitive development, language processing abilities, reading habits, educational factors, and societal expectations.

Socioeconomic and cultural status has a stronger impact than origin and gender on the English test performances of students.

Alongside the greater motivation linked to social aspects, there is also a cognitive effect linked to exposure to a second language, which can contribute to developing critical cognitive skills which do not emerge in Italian and Mathematics. To understand these differences, it seems essential to consider the multiplicity of cognitive, motivational, and cultural influences that contribute to performance in English learning tests.

In other words, using the English language only in the learning process is less motivating than its use in the real-life context. It might also be applied to the knowledge of English by foreigners who come to Italy when the English language becomes part of a process of social and cultural emancipation.

The populations of students in our schools are characterized by considerable heterogeneity in cognition, culture, ethnicity, and social status, and it is becoming increasingly important to guarantee everyone the same opportunities for learning and school adaptation. A possible reason international students in Italy do not struggle with English can be motivational.

For immigrants, knowledge of English may have been essential in reaching the place of destination. In this sense, the interest in mobility between countries may be associated with a greater interest in foreign languages, mainly English. Furthermore, the choice of mobility can also be associated with a perspective of greater cultural openness and willingness to have exchanges with people from different cultures.

These results imply that promoting English knowledge does not simply mean identifying and using effective teaching strategies but promoting a solid need for cultural openness that can motivate people who do not need to learn English to live and work.

The findings of this study should be interpreted in the context of several limitations. Firstly, the analysis was conducted at the macro-regional level, which may not capture the specific effects of socioeconomic and cultural status (ESCS) on individual regions or different areas. Comparing single regions or specific areas would provide a more accurate assessment of the effect of ESCS on English learning outcomes. Another limitation pertains to the measure used to assess ESCS, which is a composite indicator encompassing various financial, social, and cultural resources available to students. This composite measure does not allow for identifying specific factors within a student's background that may significantly impact success in one subject over another. Additionally, the data analyzed in this study correspond to the 2018/19 school year. Therefore, further investigation is needed to understand why the post-pandemic decrease in learning achievements in Italian and Mathematics was inconsistent with the English tests. Future research could explore territorial disparities by examining individual regions or specific areas and also differentiate the contributions of different components of ESCS to performance on the English tests. Comparing results before and after the pandemic would further enrich our understanding of the effects of external factors on English test performance.

References

Alamer A., Lee J. (2019), "A motivational process model explaining L2 Saudi students' achievement of English", *System*, 87, 102133.

Bacon S.M., Finnemann M.D. (1992), "Sex differences in self-reported beliefs about foreign-language learning and authentic oral and written input", *Language Learning*, 42 (4), pp. 471-495.

Bayrakdar S., Guveli A. (2020), *Inequalities in home learning and schools' provision of distance teaching during school closure of COVID-19 lockdown in the UK (No. 2020-09)*, ISER Working Paper Series.

Benadusi L. (2021), "La giustizia sociale", *Scuola democratica*, 12 (2), pp. 93-105.

Benadusi L. (2021b), "Giustizia come equità: il dibattito teorico", in L. Benadusi, O. Giancola, *Equità e merito nella scuola*, FrancoAngeli, Milano.

Berkowitz R. (2021), "School climate and the socioeconomic literacy achievement gap: Multilevel analysis of compensation, mediation, and moderation models", *Children and Youth Services Review*, 130, 106238.
Bonal X., González S. (2020), "The impact of lockdown on the learning gap: family and school divisions in times of crisis", *International Review of Education*, 66 (5), pp. 635-655.
Bourdieu P. (1986), "Habitus, code et codification", *Actes de la recherche en sciences sociales*, 64 (1), pp. 40-44.
Bratti M., Checchi D., Filippin A. (2007), "Geographical differences in italian students'Mathematical competencies: evidence from PISA 2003", *Giornale degli Economisti e Annali di Economia*, 66 (3), pp. 299-333.
Campodifiori E., Figura E., Papini M., Ricci R. (2010), Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia, INVALSI, Roma.
Chen Q., Kong Y., Gao W., Mo L. (2018), "Effects of Socioeconomic Status, Parent-child Relationship, and Learning Motivation on Reading Ability", *Frontiers in Psychology*, 9: 1297.
Chiang Y.L. (2018), "When things don't go as planned: contingencies, cultural capital, and parental involvement for elite university admission in China", *Comparative Education Review*, 62 (4), pp. 503-521.
Cho Y.A. (2021), "Learner Need Satisfaction and Effort in an Online English Language Learning Context", 언어학 연구, 61, pp. 207-228.
Clark A. (1995), "Boys into modern languages: An investigation of the discrepancy in attitudes and performance between boys and girls in modern languages", *Gender and Education*, 7 (3), pp. 315-326.
Csizér K., Dörnyei Z. (2005), "The internal structure of language learning motivation and its relationship with language choice and learning effort", *The Modern Language Journal*, 89 (1), pp. 19-36.
Cullinane C., Montacute R. (2020), C*OVID-19 and Social Mobility Impact Brief# 1: School Shutdown*, retrieved on January 25, 2024, from https://www.suttontrust.com/wp-content/uploads/2021/01/School-Shutdown-Covid-19.pdf.
Devenish B., Hooley M., Mellor D. (2017), "The pathways between socioeconomic status and adolescent outcomes: A systematic review", *American Journal of Community Psychology*, 59 (1-2), pp. 219-238.
Di Sano S., Balenzano C. (in press), *The influence of socioeconomic-cultural background on academic results in the INVALSI tests of Italian and Mathematics in two southern regions: Puglia and Abruzzo*.
Douglas Fir Group (2016), "A transdisciplinary framework for SLA in a multilingual world", *The Modern Language Journal*, 100 (S1), pp. 19-47.
Drane C.F., Vernon L., O'Shea S. (2021), "Vulnerable learners in the age of COVID-19: A scoping review", *The Australian Educational Researcher*, 48 (4), pp. 585-604.
Duff P.A. (2019), "Social dimensions and processes in second language acquisition: Multilingual socialization in transnational contexts", *The Modern Language Journal*, 103, pp. 6-22.

European Education and Culture Executive Agency, Eurydice (2020), *L'equità nell'istruzione scolastica in Europa: strutture, politiche e rendimento degli studenti*, Publications Office, retrieved on December 12, 2023, from https://data.europa.eu/doi/10.2797/720431.

Gardner R.C. (2001), "Integrative motivation and second language acquisition", *Motivation and second language acquisition*, 2 3(1), pp. 1-19.

Gardner R.C., Lalonde R.N., Moorcroft R. (1985), "The role of attitudes and motivation in second language learning: Correlational and experimental considerations", *Language Learning*, 35 (2), pp. 207-227.

Giancola O., Salmieri L. (2020), *Family Background, School-Track and Macro-Area: the Complex Chains of Education Inequalities in Italy*, Working papers series, Dipartimento di scienze sociali ed economiche (No. 4/20).

INVALSI (2019), *Rapporto nazionale. Rapporto prove INVALSI 2019*, retrieved on December 12, 2022, from https://INVALSIareaprove.cineca.it/docs/2019/Rapporto_prove_INVALSI_2019.pdf.

Kaiser T., Li J., Pollmann-Schult M. (2019), "The reproduction of educational inequalities-do parenting and child behavioural problems matter?", *Acta Sociologica*, 62 (4), pp. 420-439.

Keumala M., Samad N.M.A., Samad I.A., Rachmawaty N. (2019), "The Influence of Socio Cultural and Educational Background on EFL Learners' Motivation", *Indonesian TESOL Journal*, 1 (1), pp. 67-77.

La Salle T.P., Rocha-Neves J., Jimerson S., Di Sano S., Martinsone B., Majercakova Albertova S., Gajdošová E., Baye A., Deltour C., Martinelli V., Raykov M., Hatzichristou C., Palikara O., Szabó É., Arlauskaite Z., Athanasiou D., Brown-Earle O., Casale G., Lampropoulou A., Mikhailova A., Pinskaya M., Zvyagintsev R. (2021), "A multinational study exploring adolescent perception of school climate and mental health", *School Psychology*, May, 36 (3), pp. 155-166.

Liu R., Chiang Y.L. (2019), "Who is more motivated to learn? The roles of family background and teacher-student interaction in motivating student learning", *The Journal of Chinese Sociology*, 6 (1), pp. 1-17.

Lohmann H., Ferger F. (2014), *Educational poverty in a comparative perspective: Theoretical and empirical implications*, retrieved on January 25, 2024, from https://d-nb.info/1046953575/34.

Martini A. (2020), "Istruzione e diseguaglianze sociali nel Veneto", *Economia e società regionale*, 8 (1), pp. 60-78.

Mori S., Gobel P. (2006), "Motivation and gender in the Japanese EFL classroom", *System*, 34 (2), pp. 194-210.

Morris T., Dorling D., Davey Smith G. (2016), "How well can we predict educational outcomes? Examining the roles of cognitive ability and social position in educational attainment", *Contemporary Social Science*, 11 (2-3), pp. 154-168.

Nicol A.A., De France K. (2020), "Prejudice, Integrativeness, and Motivation to Learn a Second Language: A Mediation Analysis", *Psychological Reports*, 123 (2), pp. 420-434.

OECD (2019), *PISA 2018 Results*, Vol. II: *Where All Students Can Succeed*, PISA, OECD Publishing, Paris, retrieved on December, 12, 2023, from https://doi.org/10.1787/b5fd1b8f-en.
Plomin R., DeFries J.C., Knopik V.S., Neiderhiser J.M. (2016), "Top 10 Replicated Findings From Behavioral Genetics", *Perspectives on Psychological Science*, 11 (1), pp. 3-23.
Rasch G. (1993), *Probabilistic models for some intelligence and attainment tests*, MESA Press, Chicago.
Razi N.I.M., Rahmat N.H. (2020), "Barriers and motivation for learning english: a case study", *European Journal of English Language Teaching*, 6 (1), pp. 40-51.
Riccardi V., Donno S., Bagnarol C. (2020), "Disuguaglianze territoriali e povertà educativa nell'emergenza Covid-19: uno studio sulle famiglie italiane e la necessità di "fare scuola ma non a scuola", *Q-Times Webmagazine*, 12 (4), pp. 379-395.
Ricci R., Falzetti P., Falorsi P.D. (2019), *Le metodologie di campionamento e scomposizione della devianza nelle rilevazioni nazionali dell'INVALSI. Le rilevazioni degli apprendimenti AS 2018-2019*, FrancoAngeli, Milano.
Ruiz L.D., McMahon S.D., Jason L.A. (2018), "The role of neighborhood context and school climate in school-level academic achievement", *American Journal of Community Psychology*, 61 (3-4), pp. 296-309.
Ryan R.M., Deci E.L. (2017), *Self-determination theory: Basic psychological needs in motivation, development, and wellness*, Guilford Publications, New York.
Shelton-Strong S.J. (2020), "Advising in language learning and the support of learners' basic psychological needs: A self-determination theory perspective", *Language Teaching Research*, 1362168820912355.
Scriber D., Brooks K., Oca J.M.D. (2020), "Child rights, social justice, and professional ethics", in *International Handbook on Child Rights and School Psychology*, Springer, Cham, pp. 37-48.
Schulz W., Schunck R., Diewald M., Johnson W. (2017), "Pathways of intergenerational transmission of advantages during adolescence: social background, cognitive ability, and educational attainment", *Journal of Youth and Adolescence*, 46, pp. 2194-2214.
Szabó F., Albert Á., Csizér K. (2021), "The Effects of Family Background on the Processes of Foreign Language Learning in Hungary", *Central European Journal of Educational Research*, 3 (2), pp. 87-97.

3. A quantitative model for gender gap in G8 standardized Mathematics tests in Italian schools

by Riccardo Orlando, Ottavio G. Rizzo

The gender gap in Mathematics, i.e. the different performances of male and female students, is a well-known and well-documented phenomenon. Testing from OCSE-PISA, in particular, highlights how the gap in Italy is much larger than the international average. The didactic component of this gap has been investigated in the literature through one of two broad strategies: either large-scale, statistical analysis of test results, or item-level analysis of very few selected items with the theory of the didactic contract.

This work is an explorative analysis which aims to determine if it is possible to define specific item categories in which the gender gap is particularly notable, and to interpret these categories with the theory of the didactic contract.

Il divario di genere in Matematica, ovvero la differenza di performance tra studenti maschi e femmine, è un fenomeno ben noto e ben documentato. In particolare, le prove di OCSE-PISA evidenziano il gap italiano, molto maggiore della media internazionale. La componente didattica di questo divario è stata investigata in letteratura mediante una di due principali strategie: l'analisi statistica di un grande numero di prove, oppure un'analisi di alcuni item selezionati basata sulla teoria del contratto didattico.

Questo lavoro è un'analisi esplorativa con l'obiettivo di determinare se sia possibile definire specifiche categorie di item nelle quali il divario di genere è particolarmente notevole, e interpretare queste categorie con la teoria del contratto didattico.

1. Introduction

The gender gap in academic performance is a well-studied topic both at a national and international level (Leder and Forgasz, 2008), in particular thanks to the results of large scale international testings showing that on average, women fare better in language tests while men fare better in Mathematics tests (OECD, 2016).

We know that "[n]o significant differences between boys' and girls' Mathematics achievement [is] found before boys and girls [enter] elementary school or during early elementary years" (Fennema, 1974, mainly referring to the US context) while the gap "is large and significant in the middle school years and beyond" (Fryer and Levitt, 2010). Contini *et al.* (2017) show that in Italian context the gap appears in primary school and widens between grade 5 and 10. We also know that the size of the gap varies considerably between different educational systems (OECD, 2015, 2016), and this suggests that biological or physiological differences (Gallagher *et al.*, 2000) could not be the reason, or at the very least not the only reason.

Overall emancipation of women, as measured by the World Economic Forum Gender Gap Index (Guiso *et al.*, 2008) explains partially the gender gap in Mathematics as a result of social and cultural factors. Indeed, the World Economic Forum Gender Gap Index ranks Italy as 72nd in the world (Hausmann, Tyson and Zahidi, 2009), while Italy presents one of the highest gender gaps in Mathematics in the OCSE-PISA standardized tests (Contini *et al.*, 2017).

On the other hand, once Math anxiety and Math self-beliefs are taken in account, the Mathematics gender gap disappears (OECD, 2015); and we know that men and women utilize different strategies in problem solving (Giberti, 2019; Gallagher *et al.*, 2000; Fennema and Carpenter, 1998); hence, results could vary according to which strategies are activated by a given problem.

In this work, we aim to investigate the relationship between the characteristics of items in standardized testing and gender gap, by constructing a model that highlights item characteristics that produce different results in male and female students.

2. Theoretical Framework

This model assumes the existence of certain categories of items that independently cause discrimination. Each category is associated with a discrimination score.

Each item belongs to zero or more categories, and we expect the item discrimination to be the sum of the discrimination scores of the categories to which it belongs.

That is, given n categories, let **c** be the vector of their discrimination scores; given an item, let **m** be the vector that marks the categories to which it belongs: $\mathbf{m}_i$ is 1 if the item belongs to category *i*, 0 otherwise.

Therefore, an item's total discrimination is given by the scalar product $\mathbf{c}\cdot\mathbf{m}$.

Considering now *N* items, and their discrimination vector **d**, their classification is given by a matrix *M* and we expect $\mathbf{d} = M\mathbf{c}$.

This model formulation treats the item discrimination as unknown, and the category discrimination as known. Of course, in practice the opposite is the case and we obtain c with a least-squares method.

3. Methodology

To obtain each item's discrimination, we first compute the uniform Differential Item Functioning score, using as reference and focal groups male and female students respectively (Meyer, 2014).

This yields a value E on a multiplicative scale, from 0 to ∞, with a score of 1 indicating no discrimination. Therefore, we transform this score to D = -100 log(E), in order to obtain values on an additive scale such that positive values indicate discrimination in favor of female students, and a value of 0 indicates no discrimination.

The figure below shows item discrimination values computed for four INVALSI tests, and their approximate distribution.

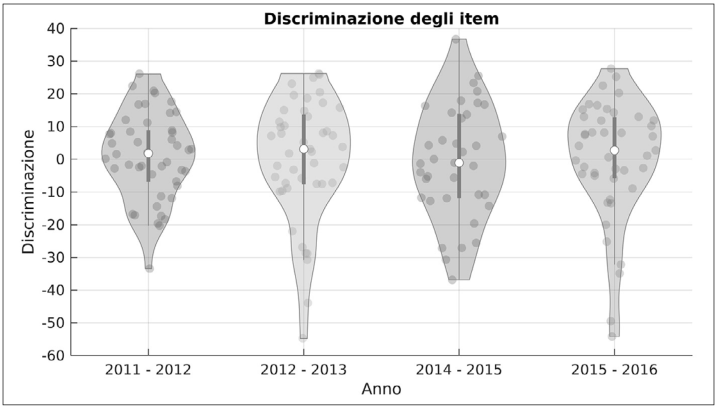

Fig. 1 – Item discrimination distribution for four INVALSI tests

The categories are constructed "a posteriori", according to our analysis of *half* of the available data, namely the tests from school years 2011/12, 2012/13, 2014/15 and 2015/16. We refer to this data as the construction set, and the remaining data (from years 2009/10, 2010/11, 2013/14 and 2016/17) as the validation set.

We compute the E score described above for each item in the construction set, and select the items with highest score (in absolute value).

We then examine these items for common features, with respect to the tested skills or the presentation, and use these commonalities to define a set of possibly discriminating categories.

In particular, these are the main features that we use to construct the categories:

Topic: the Mathematical topic of the question, with reference to the pre-existing INVALSI classification known as "ambiti di contenuto". These group grade 8 items in four "content domains", which are:

– Numbers ("Numeri");
– Relationships and functions ("Relazioni e funzioni");
– Space and Geometric Figures ("Spazio e figure");
– Data and Forecasts ("Dati e previsioni").

We also consider the specific Mathematical skills required to answer the question correctly.

Item type: we consider the type of the item, such as open or multiple choice, as well as the type of the answer: numeric, text, ...

Information density and accessibility: we consider the language and reading comprehension skills required to understand the question, as well as other comprehension skills (such as estimation, or reading a plot).

After drafting the categories, we classify all items of the validation set. This process highlights the definitions that need clarification, as well as those that match too many items or too few.

This allows us to clarify the definitions, and discard the categories that don't match enough items.

Finally, we classify every item of both the construction set and the validation set, denoting all the categories to which each item belongs.

This process results in the following categories:

– **Algebra**: The item asks to deduce or manipulate an algebraic expression;
– **Arithmetics – distractor**: An arithmetic item, such that the simplest solution strategy does not yield the correct result;
– **Asymmetric distractors**: A multiple choice item with a numeric answer, with a wide range of possible options and such that the correct answer is an extreme of the range;

- **Estimate**: The item asks to estimate a measure or an amount given an image or a plot, or to estimate the result of an arithmetic operation;
- **Explain your reasoning**: The item asks to justify the answer – simply giving the correct answer is not enough; the student must provide a correct reasoning. This category includes arithmetic items, where a student must give or choose from a list the expression used to obtain the result;
- **Extrapolation**: The item asks to extrapolate a pattern from given information;
- **Geometry – compute**: An analytic geometry item, or an Euclidean geometry item that asks to compute a value without a proof;
- **Geometry – draw**: The item asks to draw or complete a drawing of a geometric figure;
- **Implicit hypothesis**: The item assumes a hypothesis implicitly, without mentioning it in the text;
- **Multiple choice item**: The item has at most four possible answers, or is clearly divided in sub-items that meet this definition;
- **Nonlinear relationship**: The item includes two quantities that are connected in a non-linear relationship;
- **Numeric answer**: The item answer is a single number, eventually including a unit of measure, or the item is clearly divided in sub-items that meet this definiton;
- **Plot**: The item asks to draw, read or complete a plot or chart;
- **Probability – intuition**: The item is a probaility question that does not require computation;
- **Redundant information**: The item text contains much redundant or unnecessary information;
- **Standard arithmetics**: An arithmetic item, that can be solved correctly using only standard procedures;
- **Base**: All items belong to this category.

The *Base* category exists to allow for non-didactic sources of discrimination: we expect those sources to contribute equally to all items' discrimination scores.

In order to assess the model's performance, for each category we consider the model score on the validation set, and we consider as robust the categories that meet the following criteria:

- They contain at least 3 items of the validation set;
- They contain at least 3 items of the construciton set;
- They have a discrimination score of at least 5 (in absolute value).

4. Results

The following table shows the results for each category.

– **Nc** is the number of items in the construction set that belong to that category;
– **Nv** is the number of items in the validation set that belong to that category;
– **Average** is the average discrimination of those items;
– **Model** is the discrimination score computed by the model.

We obtain five robust categories, in bold in the table.

Tab. 1 – Results by cathegory (robust cathegories in bold)

Cathegory	*Nc*	*Nv*	*Average*	*Model*
Algebra	**17**	**18**	**7.7**	**15.1**
Arithmetics – distractor	8	2	17.4	25.8
Standard arithmetics	43	29	-0.2	1.5
Asymmetric distractors	**4**	**4**	**-4.1**	**-7.9**
Extrapolation	**14**	**31**	**-6.8**	**-8.1**
Geometry – compute	25	18	0.9	1.2
Geometry – draw	**5**	**9**	**17.5**	**21.8**
Explain your reasoning	**22**	**40**	**3.9**	**7.8**
Plot	37	36	-2.2	1.1
Redundant information	19	16	-0.1	1.6
Implicit hypothesis	4	0	-	0
Multiple choice item	95	107	0.1	2.9
Probability – intuition	5	8	-2.4	-4.3
Nonlinear relationship	2	5	-18.5	-21.9
Numeric answer	71	56	-3.2	0.6
Estimate	15	13	-2.3	-0.8
Base	171	171	0.2	-4.9

We consider some variations on this model, which did not yield interesting results.

– Including item difficulty as a parameter did not increase the model effectiveness;
– Computing the category discrimination score on the construction set, then computing the expected discrimination on the validation set, did not change significantly the model effectiveness.

5. Conclusions

We find that the categories *Algebra*, *Geometry – draw* and *Explain your reasoning* have positive discrimination, and therefore favor female students.

This result is not surprising, given their greater language skills and sensitivity to the didactic contract. In particular, this matches with the effect known as the *need for formal justification* (Bolondi *et al.*, 2018).

The two robust categories with negative discrimination are *Asymmetric distractors* and *Extrapolate*. These categories favor male students.

In the second category especially, the need for formal justification imposes the necessity of using only the values explicitly mentioned in the text.

In general, we observe that robust categories match up with known effects from the theory of the didactic contract. Therefore, further work may wish to construct new categories based on this framework, rather than on direct item observation.

This model can be applied with no modifications to grade 5 tests, while its application to grade 2 tests will require the construction of new categories.

Grade 10 and 13 tests, however, will require careful handling of student self-selection as Italy has a common curriculum only up to grade 8: high schools (grade 9 to 13) offer different tracks, with a different amount of time spent on Mathematics and many tracks presenting a significant gender imbalance.

Finally, with the advent of computer-based testing, the Mathematical tools will have to adapt to a sparser coverage of item answers.

References

Bolondi G., Ferretti F., Giberti C. (2018), "Didactic Contract as a Key to Interpreting Gender Differences in Maths", *EPCS Journal,* 18, pp. 415-435.

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

Fennema E. (1974), "Mathematics Learning and the Sexes: A Review", *Journal for Research in Mathematics Education*, 5 (3), pp. 126-139.

Fennema E., Carpenter T.P. (1998), "New Perspectives on Gender Differences in Mathematics: An Introduction", *Educational Researcher*, 27 (5), pp. 4-5.

Fryer R.G., Levitt S.D. (2010), "An Empirical Analysis of the Gender Gap in Mathematics", *American Economic Journal: Applied Economics*, 2 (2), pp. 210-240.

Gallagher A.M., De Lisi R., Holst P.C., McGillicuddy-De Lisi A.V., Morely M., Cahalan C. (2000), "Gender Differences in Advanced Mathematical Problem Solving", *Journal of Experimental Child Psychology*, 75 (3), pp. 165-190.

Giberti C. (2019), “Gender differences in Mathematics: from the international literature to the Italian situation”, *Didattica della matematica. Dalla ricerca alle pratiche d’aula*, 5, pp. 44-69.
Guiso L., Monte F., Sapienza P., Zingales L. (2008), “Culture, gender, and math”, *Science*, 320, pp. 1164-1165.
Hausmann R., Tyson L.D., Zahidi S. (2009), *The Global Gender Gap Report: 2009*, World Economic Forum, Geneva.
Leder G., Forgasz H. (2008), “Mathematics education: new perspectives on gender”, *ZDM*, 40 (4), pp. 513-518.
Meyer J.P. (2014), *Applied Measurement with JMetrik*, Routledge, New York.
OECD (2015), *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*, OECD Publishing, Paris.
OECD (2016), *PISA 2015 Results*, Vol. I: *Excellence and Equity in Education*, OECD Publishing, Paris.

4. The Italian Regional Divide in Education: the role of School Starting Age

by Giorgio Monti

This work wants to understand the effect of students' age on INVALSI test scores. The Italian Schooling System allows parents of children born between January 1st and April 30th to enroll them to primary school in the year when they turn 5 (instead of 6). Although this rule is the same for all the Italian territory, there is a strong geographical heterogeneity in the choice. In regions in the South the 50% of student born in the initial four months of the year are enrolled earlier, against the 10% in the North of Italy. This generates a difference in average age at the moment of the test between students in the South and in the North, since the test is administered according to the grade. Several studies show how, especially in primary school, age is a main driver of scores at standardized test and also how school entry age can affect the development of non-cognitive skills with consequences also at later stages.

In this paper i want to highlight cognitive how the difference in average test scores between North and South Italy is partly due to differences in age between the two pools of students and to analyze what are the factors that bring parents to choose the early enrollment to primary school and to understand why it is more common in the South Italy, with a particular focus on the concept of parenting style introduced in the economic literature by Doepke and Zilibotti (2017, 2019).

Questo lavoro si pone l'obiettivo di capire quanta incidenza abbia l'età degli studenti al momento del test (e quindi l'età di iscrizione alla scuola primaria) sui risultati ai test INVALSI. Il sistema scolastico italiano permette ai genitori dei bambini nati tra il 1° gennaio e il 30 aprile di anticipare l'iscrizione dei propri figli alla scuola primaria all'anno in cui compiono 5 anni (anziché 6). Pur essendo una regola valida su tutto il territorio nazionale, c'è una forte eterogeneità geografica nella scelta di questa pratica. Nelle regioni del Sud Italia, infatti, ben il 50% dei bambini nati nel primo quadrimestre dell'anno vengono iscritti in anticipo, contro il 10% del Nord Italia.

Questo crea una differenza di età al momento del test, che viene somministrato in base al grado dello studente. La letteratura mostra come, specialmente alla scuola primaria, l'età sia una delle determinanti principali del punteggio ai test standardizzati e come l'età di ingresso a scuola abbia degli effetti sullo sviluppo delle capacità non cognitive, che possono perdurare anche fino alla maggiore età.

In questo lavoro si vuole evidenziare come le differenze nei punteggi medi ai test Nord e Sud siano in parte dovute alla differenza di età degli studenti che svolgono il test e vuole analizzare quali siano i fattori che spingano le famiglie a optare per l'anticipo scolastico e a capire perché al sud esso sia così maggiormente diffuso, con un'attenzione particolare al concetto di stile genitoriale introdotto nella letteratura economica da Doepke e Zilibotti (2017, 2019).

1. Introduction

Enrollment to primary school is based on the date of birth. In most cases, students born in the same calendar year are enrolled to the first grade at the same moment. For example, in England, students born between September 1st and August 31st of the following year start together, whereas in Spain the reference dates are the 1st of January and the 31st of December. This system is aimed to generate the minimum possible differences in age among students in the same class. Differences in outcomes, both cognitive and non-cognitive skills, between students in the same grade, but with different age, is called Relative Age Effect. 1.

However, some school systems allow parents to decide about the timing of the enrollment. This is the case of Italy, where children born between January and April may be enrolled to primary school at the age of 5 or at the age of 6, whereas all the other children must be enrolled the year they turn 6. I will discuss the system more in detail later in the paper.

As noticed by several articles in the literature, it is impossible to disentangle school starting age effect from age at the test effect in empirical application like the one in this work. Moreover, in the Italian setting, another issue arises: the selection in age. In fact, as already mentioned, parents of children born between January and April can decide when to enroll the child to primary school. This generates a strong endogeneity in school starting age. Parents will indeed enroll their child earlier because of their beliefs about her ability, because of social interactions or for their expectation about the child's future. This amplifies the problem for identification since the difference in score depends on both difference in age and difference in ability and parental background.

In this work I will focus on the results in standardized test score in Italian and Mathematics of students in primary schools. Italy presents a significant heterogeneity in terms of several economic and social outputs, and this is also true for results at standardized tests in both Mathematics and Literacy, with students from the North performing constantly better than their peers from the South in both National and International Standardized Test (INVALSI, PISA, TIMSS and PIRLS).

There is another difference between North and South Italy which have not been explored in depth so far: the age of enrollment at primary school. In fact, even if the Italian education system is almost completely controlled by the central government and does not present differences in the enrollment rules across the territory, people from different regions follow these rules in different ways. The flexibility left to parents about the timing of enrollment results indeed in a significant difference in the average age of enrollment across macro-area of the country, with parents in the North enrolling their children to primary school older than in the South.

Given this difference in age of enrollment, and therefore in age at the test, among geographic macro-areas, one of the aims of this work is to measure how much of the gap in scores is due to difference in average age of students.

In the last section I will try to understand the reasons why parents from the North have different preferences than those in the South. I will show that early enrollment is correlated with less permissive parents. Moreover, since schools have some room in deciding the rules of enrollment, and which children have priority over others, it was important to understand whether the heterogeneity lies in parents' preferences or in schools' rules. From interviews made to school managers in Bologna (North Italy) and in Palermo (South Italy) it seems that most of heterogeneity is in the demand, since in Bologna they declared that almost no parents ask for early enrollment, whereas the opposite is true in Palermo. In addition, school managers in Palermo declared that it is very difficult to refuse the early enrollment mainly for two reasons: parents may appeal to administrative justice, and secondly if they refused all requests, they will have much fewer enrollments with severe consequences on the formation of classes.

2. Related Literature

The North-South divide in Italy has been widely studied from several perspectives. In the country it is called the "Questione Meridionale" (the Southern Question) and has been in the political debate for decades. The gap in education has been object of several studies, mainly Italian, and an extensive review has been made by Asso *et al.* (2015). Moreover, the Italian Ministry of Education publishes every year a report on the status of the Italian education system, showing the trend in differences among macro-areas in the countries, evidencing how students from the South perform worse in almost every output analyzed, from international test scores (e.g. PISA, TIMSS and PIRLS) to INVALSI scores (INVALSI Report 2012-2017). However, these studies are mainly descriptive, showing correlations and aggregate data. None of this report studies empirically the effect of age, nor investigates the effect of the differences in early enrollment across the Italian territory.

The effect of age on education has been instead studied for years, by both economists and educational scientists.

School starting age (SSA) effects have been investigated from different perspectives and in different countries. The main problem in the quantitative analysis of such effects is that it is impossible to disentangle the effect of school starting age from the effect of the age of the child the day of the test.

Several articles have been written on the effects of the SSA on different outcomes, from educational attainment to non-cognitive skills. Black *et al.* (2011) used Norwegian data and was able to separate the effect of SSA from test age effects by using IQ scores taken by students at about 18 years old outside school. They argue that the major effect is given by the age at the test, but also found that children who started school older are less likely to have poor mental health at 18 and teenage pregnancy. Fredriksson and Ockert (2013) used a wide dataset from Sweden and used the school entry cut-off as instrument to find that students who started school older have higher educational attainment and that postponing tracking until age of 16 reduces this effect. In their seminal work, Bedard and Duhey (2006) showed, using data from OECD countries, how younger children in a given cohort score significantly worse than their older peers and are less likely to attend the university in US and Canada, arguing that the effects of the school starting age are very persistent. On the other hand, Elder and Lubotsky (2009) argue that the effect of SSA on academic performance declines as the children grow up. Muhlenweg *et al.* (2012) focused instead on effects on the development of non-cognitive skills, and, using data from Germany, they found that children who begin primary school older are less often hyperactive and more adaptable to change. Cornelissen and Dustmann (2019) studied the effect of receiving additional schooling before age five on both cognitive and non-cognitive abilities, exploiting variation of age at school entry in England. They found that the effect on cognitive ability disappears by age 11, but that non-cognitive ability is still affected at later stage. Parents who enroll their children earlier are in fact reducing the exposition of the pupil to pre-school programs, which have important effects on skill development, as stressed also by Cunha *et al.* (2006), Cunha and Heckman (2007), and Heckman (2008). The only recent study on Italian students has been made by Ponzo and Scoppa (2014), who used PIRLS, TIMSS and PISA data on Italian students to see how the SSA affects results: students enrolled earlier perform better at tests and are more likely to choose academic high schools rather than vocational schools. Their analysis is focused on a smaller dataset than the one I will use, and used the expected age as instrumental variable for actual age of the pupil, in order to quantify the effect of SSA. They used data from cohorts who started primary school with the old system, when there was small flexibility given to parents on the time of enrollment. Finally, Fenoll *et al.* (2018), studied the selection in early enrollment among Italian students born between January and April, showing how early entrants would have scored much better than regular students of the same age if they were enrolled one year later. I will partly use their methodology, but to look into regional differences and to do robustness checks on the main empirical strategy. To study the relation between early enrollment and parenting style I looked at the work of Doepke and Zilibotti (2017) who categorizes parents according to the classification of Baumrind (1967), shows how parenting styles are correlated with several economic outcomes such as inequality, social mobility and returns to education and presents a theoretical model to explain why parents choose a particular parenting style.

3. The Italian Enrollment System to Primary School

Italian children start the I grade of primary school in September of the year when they turn 6. Only children born from January to April can be enrolled one year before if their parents ask so. The possibility of early enrollment was introduced by the "Moratti reform", named after the minister of education who signed it in 2003. Before this reform the cut-off date was the 31st of December so that only children who turned 6 in the same year of the beginning of classes could start primary school. In other words, before 2003, if the school started in September of year t, it was mandatory for parents to enroll their child to school if she turned 6 in the same year t, and cannot enroll them if he turned 6 in year t+1. It remains the prohibition, except for healthy issue or language comprehension of the child due to his origins, of the so called "red-shirting", the late enrollment that is becoming popular in other education systems, especially in the United States and Australia. The reform represented an exogenous variation in the enrollment system. However, the possibility of moving up the enrollment of children existed even before the reform, even if it was more costly. In fact, before 2003, and this is still true for children born after the 30th of April, parents can ask to enroll their children directly to the second grade of primary school even if they turn 6 in the same year, instead of 7 as regular students who completed the first grade. In order to do so, children have to pass an exam in September, just before the school year starts. In practice, this is slightly different from early enrollment since it means a "skip" of the first grade, and it is conditional on a test. The reform hence made it easier for children born from January to April to be among the youngest in the class. The writer of this reform justified this choice saying that "parents should have more flexibility in choosing when to enroll their child to primary school". The problem seems to be that no teachers' evaluation is mandatory, and all depends on parents' beliefs about their children readiness and home-made cost-benefit analysis.

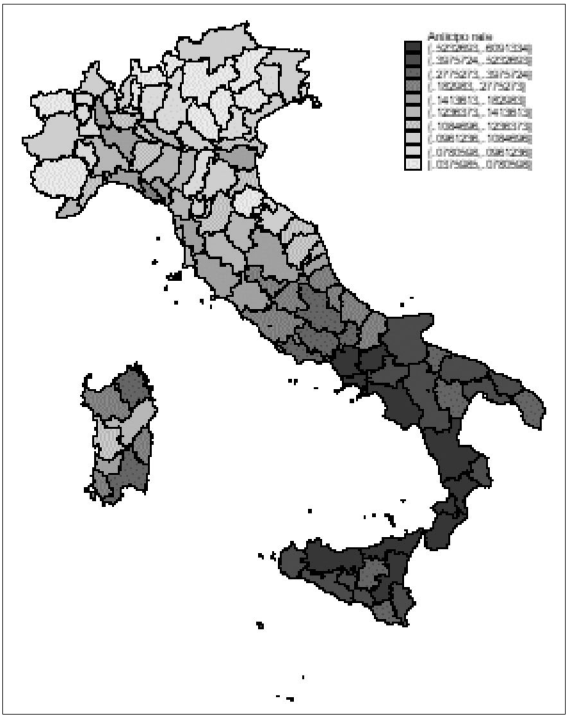

Fig. 1 – Early Entrants rate, by province

Figure 1 shows the heterogeneity of the percentage of early entrants across Italian territory. Darker provinces have a higher percentage of students born between January and April who are early entrants, lighter ones have it lower. The heterogeneity is very strong and follows a clear pattern: provinces in the last decile are all in South Italy and have from 63% to 74% of early entrants, whereas those in the first decile are all in the Center-North and have only 0.5% to 1%.

4. The INVALSI test and Dataset

The INVALSI test is a written standardized exam that every pupil has to do in her II, V, VIII and X grade. It was introduced in the school year 2007/08 and is divided in two parts, one testing the ability in Mathematics while the other the ability in reading and grammar. In the main empirical analysis, I will use data from the II and V grade from 2012 to 2017, which are likely to be the cleanest, since grade repetition is almost impossible during primary school. This data comes from student all enrolled after the reform of the enrollment system. Moreover, I will also use scores of pupils from a selected sample of classes where external monitors were randomly assigned. Score manipulation is indeed a problem, especially in the South (Bertoni *et al.*, 2013; Angrist *et al.*, 2017) and to limit it, the ministry of education selects each year a random sample of class where to send external monitors. In addition, the INVALSI dataset has also a variable for an index of score manipulation, measured by a statistical model that looks for abnormal high average scores, low within-class variability, and implausible missing data patterns. However, cheating seems not to be related to the age of the students and then is not likely to bias results.

INVALSI Data also provides information about family background (index for socioeconomic status, parents education and employment, citizenship), gender and date of birth.

5. Descriptive Statistics

Table 1 reports summary statistics for the whole sample and for the subsample of monitored classes for both II and V grade. The first row shows that the scores decrease dramatically when looking at data from monitored classes. This is an evidence of how severe is the issue of score manipulation, especially in the South. Moreover, we can see that cheating seems to be a bigger issue in II grade than in V. Looking at differences between monitored and not monitored class we can see that it is larger in II grade than in V grade and in the South than in the North. Table 1 also reports the difference in several outputs among Italian macro-areas. Mothers without a job are more than the 46% in the South and 25% in the North, but also the percentage of mothers with a diploma is higher in the North, and the same is true for the percentage of children who attended nursery schools, and for the ESCS, an index proxying the socioeconomic status of the family (available only for V graders). The statistics for background variables do not change significantly when moving to the subsample of monitored classes, confirming the goodness of the randomization made by the INVALSI.

Tab. 1 – Summary Statistics

Panel A: II Grade								
	Whole Sample				*Monitored Class*			
Variable	*Italy*	*North*	*Center*	*South*	*Italy*	*North*	*Center*	*South*
Std. Score Ita	0.33	0.22	0.36	0.47	0.00	0.03	0.08	-0.06
Mother Dip	0.71	0.70	0.80	0.66	0.71	0.7	0.79	0.68
Father Dip	0.64	0.62	0.73	0.61	0.64	0.62	0.71	0.62
Mother Grad	0.16	0.13	0.23	0.17	0.15	0.12	0.21	0.16
Father Grad	0.17	0.14	0.25	0.17	0.16	0.13	0.22	0.17
Mother Unemp.	0.32	0.25	0.24	0.46	0.32	0.24	0.24	0.44
Father Unemp.	0.05	0.03	0.04	0.08	0.05	0.04	0.04	0.06
Attended KG	0.90	0.92	0.92	0.87	0.91	0.93	0.93	0.87
Attended Nursery	0.37	0.39	0.45	0.29	0.37	0.39	0.46	0.30
Class Size	20.50	20.80	20.9	19.8	20.8	20.8	21.2	20.4
N	2,812,686	1,313,621	530,643	968,422	150,451	59,336	29,900	61,215
Panel B: V Grade								
	Whole Sample				*Monitored Class*			
Variable	*Italy*	*North*	*Center*	*South*	*Italy*	*North*	*Center*	*South*
Std. Score Ita	0.21	0.20	0.26	0.20	0.01	0.08	0.10	-0.10
Mother Dip	0.67	0.66	0.77	0.62	0.68	0.67	0.76	0.64
Father Dip	0.60	0.59	0.70	0.57	0.61	0.6	0.69	0.59
Mother Grad	0.14	0.12	0.20	0.14	0.13	0.11	0.19	0.14
Father Grad	0.15	0.13	0.21	0.15	0.15	0.12	0.20	0.15
Mother Unemp.	0.34	0.26	0.26	0.49	0.34	0.24	0.25	0.46
Father Unemp.	0.05	0.03	0.03	0.08	0.05	0.03	0.03	0.07
ESCS	0.06	0.15	0.20	-0.13	0.08	0.18	0.19	-0.06
Attended KG	0.89	0.92	0.90	0.84	0.89	0.92	0.92	0.84
Attended Nursery	0.32	0.34	0.40	0.27	0.33	0.35	0.41	0.28
Class Size	20.40	20.60	20.90	19.9	20.8	20.8	21.2	20.80
N	2,751,867	1,245,023	506,793	1,000,051	148,846	57,546	29,166	62,134

6. Check for Birth Seasonality

In the whole empirical strategy I use in this work, I always assume that the month of birth is exogenous with respect to other variables that may affect educational outcomes. Bound, Jaeger and Baker (1995) found indeed for the United States that quarter of birth is correlated with other background variables and argued that the date of birth may be an inappropriate instrument in many frameworks. They review the evidence that quarter of birth is correlated with school attendance rate, likelihood that a student will be assessed as having behavioral difficulties and also with mental and physical health. To check if Italian students show the same correlation between month of birth and other variables, I check the distribution of many background variables across months of birth. INVALSI provides many background variables for students in V grade. For other grades, data are less comprehensive.

Tables 2, 3 and 4 show the results for regressions of the dummy for Eligibility for early enrollment (to be born between January and April) on several background variables. Since I will use the variation in early enrollment between North and South, I run these regressions also for the two subsamples of North and South Italy.

Tab. 2 – Check for Birth Seasonality, Italy

Variables	*(1) Mother Grad.*	*(2) Father Grad.*	*(3) Mother HS*	*(4) Father HS*	*(5) Mother Unemp.*	*(6) Father Unemp.*	*(7) ESCS*
Eligible	0.000	0.001	-0.001**	0.001	-0.002***	-0.001**	0.009***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Constant	0.109***	0.096***	0.648***	0.599***	0.039***	0.050***	0.146***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Observations	3,235,536	3,235,536	3,235,536	3,235,536	3,235,536	3,235,536	3,147,187

Robust standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1

Tab. 3 – Check for Birth Seasonality, South Italy

Variables	*(1) Mother Grad.*	*(2) Father Grad.*	*(3) Mother HS*	*(4) Father HS*	*(5) Mother Unemp.*	*(6) Father Unemp.*	*(7) ESCS*
Eligible	0.004***	0.003***	0.004***	0.004***	-0.003***	-0.001	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Constant	0.088***	0.078***	0.068***	0.572***	0.481***	0.075***	-0.057***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Observations	1,169,900	1,169,900	1,169,900	1,169,900	1,169,900	1,169,900	1,120,321

Robust standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1

Tab. 4 – Check for Birth Seasonality, North Italy

Variables	*(1) Mother Grad.*	*(2) Father Grad.*	*(3) Mother HS*	*(4) Father HS*	*(5) Mother Unemp.*	*(6) Father Unemp.*	*(7) ESCS*
Eligible	-0.003***	-0.001*	-0.005***	-0.001	-0.002**	-0.000	0.013***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.002)
Constant	0.122***	0.109***	0.637***	0.583***	0.258***	0.039***	0.251***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.002)
Observations	1,466,267	1,466,267	1,466,267	1,466,267	1,466,267	1,466,267	1,443,218

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Results reported in the tables seem to assure that birth seasonality is not a big issue in this framework. Even if some coefficients many coefficients are statistically significant from zero, and differ across macro areas, their size is negligible, being always less than 0.01. Since all the covariates are dummies (except for the socioeconomic status index, ESCS), this means that differences between eligibles and not eligibles are never bigger than 1%. The statistical significance, in this case, is probably due to the very large sample size (more than 3 million observations for the whole sample and more than 1 million in the subsample).

7. Descriptive statistics for Early Entrants

Figure 2 shows the percentage of Early Entrants for each month from January to April and for each Italian Macro-area. This ratio is clearly decreasing when moving from January to April, suggesting that the month of birth is a determinant variable in the choice. This is understandable: parents of children born in April have more concerns than parents of those born in January in enrolling the child earlier given that they are 3 months younger and are more likely to be perceived as not ready for primary school.

As already mentioned, the fraction of students who were enrolled earlier is much higher in the South: almost the 80% of born students born in January were enrolled earlier, against the 21% in the North. The percentage for born in April goes down to 30% in the South and to 3.5% in the North.

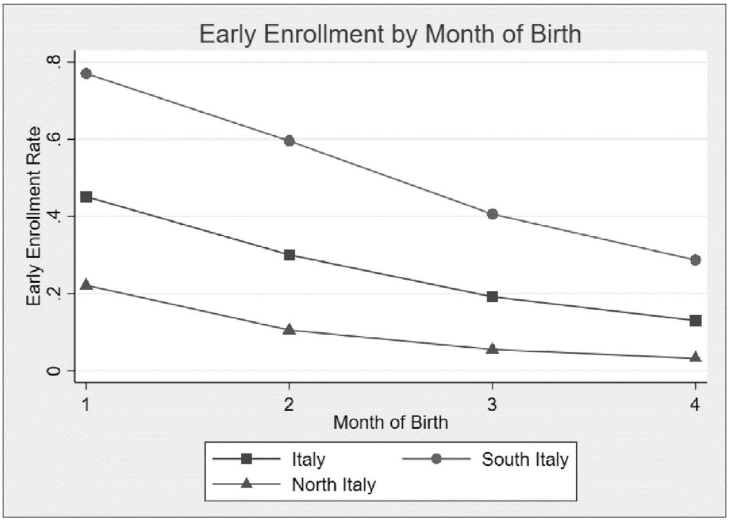

Fig. 2 – Early Entrants by Month of Birth

Given that the child is born before the 30th of April, the decision of enrolling her one year earlier to primary school is almost completely up to the parents. As reported by table 5 and studied by Fenoll *et al.* (2018), there is a strong selection in the choice. This is consistent with the hypothesis that the decision depends on how they perceive the ability and the readiness of the child. The table shows indeed significant differences in all the variables considered. Firstly, we can notice that early enrollment is more common among females than among males. This is consistent with the fact that girls mature earlier than boys (Bierman *et al.*, 2009; Son *et al.*, 2013). Moreover, early entrants have on average more educated parents, and a higher index of socioeconomic status (ESCS). In the end, the last row reports that the percentage of immigrants among early enrollers is higher than that among regular in the North, and the opposite is true in the South. The percentage of immigrant students in the North is much higher than in the South, and if immigrants were less inclined to early enrollment, that could have biased the results, since immigrant students perform lower on average. The finding that immigrants are more likely to be early enrolled in the North is surprising but reassures that estimates will not be amplified by the North-South heterogeneity in the percentage of immigrant students. However, in the main analysis I only use data from Italian student.

Tab. 5 – Summary Statistics by enrollment status (only pupils born before 30/04)

	Italy			*North*			*South*		
Variables	*Early*	*Regular*	*Diff.*	*Early*	*Regular*	*Diff.*	*Early*	*Regular*	*Diff.*
Female	0.54	0.47	0.07***	0.58	0.48	0.10***	0.53	0.45	0.08***
Mother Emp.	0.54	0.63	-0.09***	0.66	0.71	-0.04***	0.47	0.38	0.10***
Father Emp.	0.89	0.91	-0.02***	0.92	0.93	-0.01***	0.88	0.86	0.02***
Mother HS	0.75	0.69	0.06***	0.76	0.69	0.07***	0.73	0.61	0.12***
Father HS	0.70	0.61	0.09***	0.61	0.71	0.10***	0.68	0.55	0.13***
Mother Grad.	0.20	0.15	0.05***	0.26	0.17	0.09***	0.18	0.09	0.08***
Father Grad.	0.16	0.11	0.05***	0.21	0.13	0.09***	0.14	0.07	0.07***
ESCS	0.15	0.04	0.11***	0.35	0.14	0.21***	0.02	-0.27	0.29***
Attended Nursery Sc.	0.32	0.33	-0.01***	0.38	0.35	0.04***	0.28	0.25	0.03***
Attended KG	0.85	0.90	-0.04***	0.90	0.92	-0.02***	0.83	0.85	-0.02***
Immigrant	0.07	0.11	-0.04***	0.18	0.13	0.05***	0.04	0.06	-0.02***

Emp.: Eployed; HS: High School; ESCS è spiegato nel testo; KG: Kindergarden
*** p-value < 0.01

To have a better sense of the determinants of early enrollment, I run a multivariate OLS regression of a dummy for early entrants on several background variables. I will use only students eligible for early enrollment (born in the first quarter of the year), and I will present also results for subsamples of South and North Italian students. Table 6 reports the results of these regressions. Results confirm what shown in table 5. All coefficients have the expected signs and are statistically significant. Looking at North-South heterogeneity it is possible to notice that coefficients for parents' education are much larger in the South than in the North, especially looking at mother's education. However, coefficients for parents' employment status have different signs: in the South unemployed mothers are less likely to enroll earlier their children, and the opposite is true in the North. The same is true for unemployed fathers, but the difference is less large. The reason of this difference is not analyzed in this work, however, it should not drive our results, since parents unemployment is associated with lower scores at standardized test.

Tab. 6 – OLS Regressions of Early Enrollment on Background Characteristics

Variables	*(1) Italy*	*(2) South Italy*	*(3) North Italy*
Mother Grad.	0.101***(0.002)	0.135*** (0.004)	0.078*** (0.002)
Father Grad.	0.055*** (0.002)	0.062*** (0.003)	0.049*** (0.002)
Mother HS	0.028*** (0.001)	0.047*** (0.002)	0.015*** (0.001)
Father HS	0.025*** (0.001)	0.029*** (0.002)	0.020*** (0.001)
ESCS	0.007*** (0.001)	0.028*** (0.001)	-0.008*** (0.001)
Mother Unemp.	-0.006*** (0.001)	-0.023*** (0.002)	0.017*** (0.001)
Father Unemp.	-0.006*** (0.002)	-0.005 (0.003)	0.023*** (0.003)
Feb.	-0.156*** (0.001)	-0.180*** (0.002)	-0.121*** (0.002)
March	-0.276*** (0.002)	-0.387*** (0.003)	-0.175*** (0.002)
April	-0.337*** (0.002)	-0.503*** (0.003)	-0.199*** (0.002)
Constant	0.429*** (0.002)	0.733*** (0.003)	0.204*** (0.002)
Observations	996,327	350,463	460,141
R-Squared	0.127	0.200	0.075
Nr of Schools	17,073	6,433	7,511

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Given this strong selection and the consequent differences between regulars and early entrants, to identify the effect of moving the enrollment of one year is a challenging task. In the next sections, I will propose a methodology to quantify how much of the gap in scores between North and South is due to the difference in average age, and robustness checks to validate the results.

8. North *vs* South Italy Analysis

Descriptive statistics presented in the previous sections show both the gap in scores and the gap in the number of early entrants between North and South Italy. Moreover, they evidence how early entrants come, on average, from families with a higher socioeconomic background. As noticed in the section about the literature, several studies have shown the effect of age on scores, and Fenoll and coauthors show how this effect is linear when there is no choice about the timing of enrollment, with younger students performing worse. Given the selection in the early entrance, would this effect be lower for early starters? First of all we can look at Figure 3 and Figure 4.

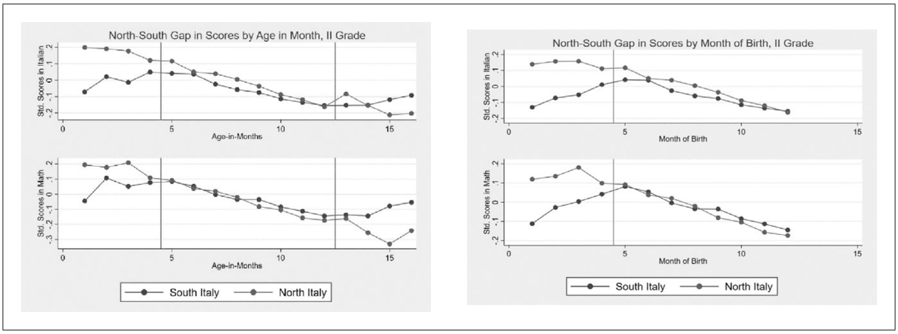

Fig. 3 – Average scores by age and by month of birth, II grade

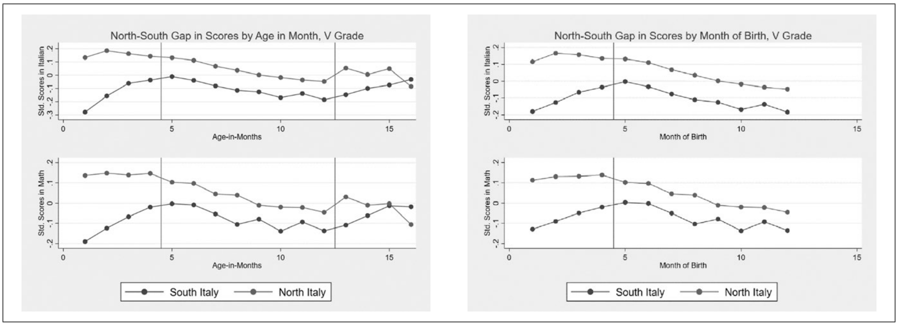

Fig. 4 – Average scores by age and month of birth

The graphs are obtained by pooling together data from all the school years available. The graphs in the top panel shows the average score for each month in South and North Italy. The variable on the horizontal axis represents the age-in-months, with 1 being students born in January who are regularly enrolled and 13 students born in January who are early starters. The area between the two vertical red lines includes the months of birth for which parents cannot choose about the enrollment. It is evident that selection is in place: for those months of birth not eligible for early enrollment, the relation between age and scores looks linear and negative. On the other hand, the graph shows how this trend change when looking at born between January and April. The graph on the lower panel instead pools together regular and early entrants, showing the scores on the month of birth, independently of the year of birth. Hence, those on the right of the vertical red line are born in the same year, whereas those on the left may be either regulars or early enrollers.

The main message coming from these graphs is that the North-South gap in scores increases dramatically when looking at students born between January and April. Moreover, in the second grade, there is almost no gap in scores between Northern and Southern students born between May and December. The main hypothesis I want to test is that this is due to the difference in age, being the rate of early entrants much higher in the South.

To measure how much of the difference in scores is due to the difference in age, I used a difference-in-difference strategy, comparing the North-South difference in scores between not eligible and that between eligible to early enrollment. To do so I have to control that the effect of age on scores of not eligible is the same for South and North, in other words, I will assume that the relationship between month of birth and scores would be the same also for months between January and April if early enrollment was not allowed.

I then run 3 different regressions, using only data from students not eligible to early enrollment (born between May and December), for South and North separately and then pooling data together and adding an interaction term (South*Month). Tables 7 and 8 show the results from regression made by using data pooled from all school years in the dataset.

Tables 7 and 8 show that the coefficients in North and South for months of birth on scores in Italian INVALSI test at the II grade are slightly different for both subjects, whereas they are not for scores in V grade. Overall, in V grade the coefficient is smaller than in II, and this is consistent with other results from the literature that show how the effect of age at test and age at the enrollment on scores decreases with the age of the student. The coefficient for II grade is indeed around 0.035 s.d. whereas in V grade it is about 0.025 s.d. This means that on average, to be born one month later decreases the score of about 0.035 s.d. in grade II and of 0.025 in grade V. This is true for all school years, and all the coefficients remain quite stable over time. Moreover, results are similar also across subjects.

Tab. 7 – Coefficients for linear effect of age on INVALSI Scores in Italian, North vs South Italy

	Panel A: II Grade			
Variables	*Italy*	*South-Italy*	*North-Italy*	*Interaction*
Month of birth	-0.035*** (0.002)	-0.030*** (0.003)	-0.037*** (0.003)	-0.037*** (0.003)
South				-0.096*** (0.034)
Interaction				0.007* (0.004)
Constant	0.254*** (0.017)	0.150*** (0.027)	0.293*** (0.026)	0.271*** (0.026)
Observations	89,617	36,768	35,151	71,919
R-squared	0.005	0.005	0.007	0.005
	Panel B: V Grade			
Variables	*Italy*	*South-Italy*	*North-Italy*	*Interaction*
Month of birth	-0.024*** (0.002)	-0.022*** (0.003)	-0.027*** (0.003)	-0.027*** (0.003)
South				-0.182*** (0.035)
Interaction				0.005 (0.004)
Constant	0.170*** (0.017)	0.08 (0.028)	0.236*** (0.026)	0.248*** (0.025)
Observations	88,338	37,015	34,013	71,028
R-squared	0.002	0.002	0.004	0.007
School FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1

Tab. 8 – Coefficients for linear effect of age on INVALSI Scores in Math, North vs South Italy

	Panel A: II Grade			
Variables	*Italy*	*South-Italy*	*North-Italy*	*Interaction*
Month of birth	-0.035*** (0.002)	-0.031*** (0.003)	-0.038*** (0.003)	-0.038*** (0.003)
South				-0.053 (0.035)
Interaction				0.008** (0.004)
Constant	0.254*** (0.017)	0.218*** (0.029)	0.253*** (0.026)	0.261*** (0.025)
Observations	86,933	35,399	34,318	69,717
R-squared	0.005	0.005	0.008	0.005
	Panel B: V Grade			
Variables	*Italy*	*South-Italy*	*North-Italy*	*Interaction*
Month of birth	-0.021*** (0.002)	-0.018*** (0.003)	-0.022*** (0.003)	-0.022*** (0.003)
South				-0.140*** (0.035)
Interaction				0.004 (0.004)
Constant	0.147*** (0.017)	0.111 (0.028)	0.163*** (0.026)	0.205*** (0.026)
Observations	85,729	35,684	33,243	68,927
R-squared	0.002	0.003	0.004	0.004
School FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1

9. The Difference-in-Difference Approach

To measure empirically how much of the North-South difference in scores is due to differences in early enrollment, I used pooled data from every school year 2012-2017. I estimated two different equations, using as dependent variables age in month and score at the INVALSI Italian and Math test for II and V graders:

– $Ageinmonth_i = \alpha + \beta South_i + yEligible_i + \delta \mathrm{South}_i * Eligible_i + \varepsilon_i$

– $Score_i = \eta + \lambda South_i + \theta Eligible_i + \rho South_i * Eligible_i + \varepsilon_i$

The two equations can be seen as differences-in-differences where the coefficients of interests are δ (for equation 1) and for equation 2. δ would how the North-South difference in average age differs across eligibles and not eligibles. The coefficient ρ would do the same for average scores.

The coefficient from the linear regressions would estimate the causal effect of the early entrance if the difference between North and South Italy in early enrollment rate was exogenous, and this is clearly not the case. However, the coefficients can provide a quantitative estimate of the portion of the gap that is explained by the different early enrollment rates, without having the ambitions of understanding the reasons behind this heterogeneity.

From the first linear regressions, I would expect that the coefficient *β* is not significantly different from 0 since it represents the difference in the average age of not eligible between students in the North and in the South. On the other hand coefficient *γ* is negative by construction: the variable *Eligible* is indeed a dummy equal to 1 if the student is born between January and April and 0 otherwise, whereas the variable *AgeinMonth* is built such that for regular students it has value of 1 if born in January and 12 if born in December and has value from 13 (born in January) to 16 (born in April) for early starters. The coefficient δ for the interaction terms should be instead positive: given the higher number of early enrollers in the south, this coefficient measure how much the gap in age between North and South increases when looking at eligibles.

The second regression is the same as the first but with INVALSI Score as the dependent variable. Then, the coefficient *λ* represents the gap in scores between not eligible in the South and in the North. The coefficient *θ* is instead the gap in scores between eligible and not eligibles in the North Italy, and *ρ* is the main coefficient of interest, measuring how much the North-South gap in scores increase when moving from not-eligibles to eligible students.

Tables 9 and 10 presents the results for II and V grade respectively.

Tab. 9 – Diff-in-Diff. II Grade, 2012-2017

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Variables	*Age-in-Months*	*Age-in-Months*	*Score ITA*	*Score ITA*	*Score MAT*	*Score MAT*
South	1.487*** (0.021)	0.044*** (0.016)	-0.092*** (0.006)	-0.039*** (0.007)	-0.044*** (0.006)	0.011 (0.007)
Eligible		-4.781*** (0.028)		0.164*** (0.009)		0.180*** (0.009)
South* Eligible		4.489*** (0.051)		-0.164*** (0.012)		-0.170*** (0.013)
Constant	6.966*** (0.025)	8.463*** (0.021)	0.018** (0.007)	-0.033*** (0.008)	0.003 (0.007)	-0.054*** (0.007)
Obs.	120,058	120,058	120,369	119,886	116,713	116,250
R-squared	0.041	0.223	0.002	0.005	0.000	0.004
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tab. 10 – Diff-in-Diff. V Grade, 2012-2017

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Variables	*Age-in-Months*	*Age-in-Months*	*Score ITA*	*Score ITA*	*Score MAT*	*Score MAT*
South	1.394*** (0.021)	0.003 (0.016)	-0.171*** (0.006)	-0.141*** (0.007)	-0.132*** (0.006)	-0.102*** (0.007)
Eligible		-4.724*** (0.028)		0.110*** (0.009)		0.104*** (0.009)
South* Eligible		4.332*** (0.051)		-0.105*** (0.012)		-0.099*** (0.013)
Constant	6.893*** (0.025)	8.394*** (0.022)	0.068*** (0.007)	0.034*** (0.008)	0.056*** (0.007)	0.023*** (0.008)
Obs	119.081	119,081	119.678	119,081	116,119	115,539
R-squared	0.036	0.213	0.007	0.009	0.004	0.006
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Columns 1, 3 and 5 of tables 9 and 10 present the overall North-South gap in age and in scores. The coefficient for age says that on average, II and V graders from the south are respectively 1.5 and 1.4 months younger than those from the North, score 0.09 and 0.17 standard deviation less at the Italian INVALSI test and 0.04 and 0.13 s.d. less in the Mathematics test. Columns 2 show instead the results for the equation of age defined above: as expected the coefficient for the variable *South* is very close to 0 (the coefficient for II graders is positive and significant but its magnitude is negligible), and that γ = -4.7 in both grades, meaning that in the North, Eligible students are on average 4.7 months older that not eligible. The coefficient δ for the interaction term is instead positive: it says that looking only at students born between January and April, the average age in the South II grade cohort is 4.5 (4.3 in V grade) months less than in the North. This difference simply reflects the fact that in the South early enrollment is far more common. Finally, columns 4 and 6 present the results of main interest. From the coefficient λ for the dummy variable South, the regression estimates for scores in Italian are $\widehat{\lambda}$ = -0.039 for II Graders and $\widehat{\lambda}$ = -0.141 for V Graders. For scores in Mathematics, they are $\widehat{\lambda}$ = -0.011 (not statistically different from zero) and $\widehat{\lambda}$ = -0.102 for II and V graders respectively. These values reflect the North-South gap in scores for not eligibles. The results for the Diff-in-Diff estimators are very impressive, especially for II graders. In fact, for scores in Italian in II grade it is $\hat{\rho}$ = -0.164 and it measures how much the North-South difference in average score increases when looking at eligible students. This means that the total gap in scores for students born between January and April is: $\widehat{\lambda} + \hat{\rho}$ = -0.203. This result is dramatic: if we look at column 3 of table 9, we see that the overall gap in scores between the two macro-region is 0.092, but when we look at not eligible it decreases to only 0.039. This means that almost 60%, calculated as (0.092-0.039)/0.092) of the gap is due to the difference in age among eligible students: in other words, to the difference in early enrollment. Looking at the results in Mathematics, the difference in scores explained by the difference in age is even larger: among not eligible there is no significant difference in scores. This means that all the difference in average test scores in Mathematics between North and South Italy comes from the difference in average age. When looking at the results for V grade, we see that the overall gap increase, and that the effect of the difference in early enrollment is less intense. This is consistent with the fact that the effect of age in school scores decrease with the student growing up. The estimate for for scores in Italian is – 0.105 and $\widehat{\lambda} + \hat{\rho}$ = -0.246. Given that the overall North-South difference in scores is 0.171 and the one for not eligibles is $\widehat{\lambda}$ = -0.141, this means that still in V grade, almost the 20% of the gap in scores between Northern and Southern students is due to difference in the enrollment. For scores in Mathematics, the percentage of the gap explained by the difference in age is larger also in V grade, being around 23%: (-0.132-0.102)/(-0.132).

If we look better at the magnitude of the Diff-in-Diff estimator, we can notice that it is very close to the linear coefficient estimated by running the regression of score on age-in-month for not eligibles. In fact, from that regression, the estimate says that on average, both in the North and in the

South, being one month younger leads to a decrease in Italian INVALSI Score of around 0.035 s.d. in II grade and of around 0.025 s.d. in V grade. Given that the increase in the difference in age (δ), is equal to 4.5 months and 4.3 months in the II and V grade respectively, we can see that multiplying this coefficient to the corresponding estimate for the linear effect of age on scores, we get 0.157 for II grade and 0.108 for the V grade, which are very similar to the estimates of the diff-in-diff coefficients which are $\hat{\rho}$ = -0.164 for the grade II and $\hat{\rho}$ = -0.107 for grade V. The very same results hold also for scores in Mathematics.

However, this estimate can be biased, and deeper checks are required. For example, we are assuming that not eligible are not affected by the number of early entrants. In the South, in fact, not eligibles would be relatively older within a class than in the North, since they will have more early entrants as classmates. To be older in the distribution of age in the class can have different effects: it can lower your results because the teacher has to "slow down" and to flatten the learning curve in order to help younger students. On the other hand, to be older can also have some positive effects on cognitive ability. Even if, at a first glance, I would expect that the negative effect of having younger peers prevails, and then that the estimates from the diff-in-diff analysis can be seen as a lower bound for the true effect, in the next sections I will propose some robustness checks that will improve the reliability of the results.

10. Robustness Checks[1]

10.1. Month by Month Analysis

In this section, I will use the same diff-in-diff strategy but using as eligible group students born in one of the first four months of the year. In other words, I will run the same regression as before four times, one for each month, dropping observation of born in other months that are also eligible. Hence, I want to estimate 4 different equations, one for each month of birth of eligible (January-April):

$$Score_i = \eta + \lambda South_i + \theta Jan_i + \rho Jan_i * Eligible_i + \varepsilon_i$$

[1] Additional Robustness Checks are not presented due to space limit and are available on request by email to giomonti92@gmail.com.

Tab. 11 – Diff-in-Diff, Month by Month Analysis

Panel A: II Grade

VARIABLES	(1) Jan Age-in-Months	(2) Jan ITA	(3) Jan MATH	(4) Feb Age-in-Months	(5) Feb ITA	(6) Feb MATH	(7) Mar Age-in-Months	(8) Mar ITA	(9) Mar MATH	(10) Apr Age-in-Months	(11) Apr ITA	(12) Apr MATH
South	0.046***	-0.040***	0.010	0.047***	-0.040***	0.010	0.044***	-0.039***	0.010	0.045***	-0.040***	0.010
	(2.923)	(-5.707)	(1.439)	(2.988)	(-5.732)	(1.432)	(2.765)	(-5.628)	(1.469)	(2.843)	(-5.726)	(1.396)
Eligible	-4.910***	0.161***	0.165***	-5.261***	0.179***	0.182***	-4.835***	0.180***	0.227***	-4.115***	0.134***	0.145***
	(-70.207)	(10.714)	(11.243)	(-96.159)	(11.216)	(11.890)	(-122.005)	(11.778)	(15.176)	(-131.455)	(8.598)	(9.527)
South*Eligible	6.103***	-0.231***	-0.242***	5.199***	-0.190***	-0.174***	3.802***	-0.172***	-0.189***	2.849***	-0.062***	-0.069***
	(58.381)	(-10.830)	(-11.072)	(49.280)	(-8.391)	(-7.545)	(41.460)	(-7.976)	(-8.584)	(33.696)	(-2.860)	(-3.067)
Constant	8.514***	-0.038***	-0.058***	8.526***	-0.041***	-0.060***	8.471***	-0.035***	-0.055***	8.483***	-0.035***	-0.054***
	(412.206)	(-4.744)	(-7.231)	(420.704)	(-5.062)	(-7.358)	(430.359)	(-4.389)	(-6.827)	(440.533)	(-4.304)	(-6.681)
Observations	91,978	92,332	89,469	91,015	91,371	88,553	91,912	92,258	89,463	91,342	91,698	88,877
R-squared	0.157	0.003	0.002	0.158	0.002	0.002	0.152	0.002	0.003	0.123	0.002	0.001

Panel B: V Grade

VARIABLES	Jan Age-in-Months	Jan ITA	Jan MATH	Feb Age-in-Months	Feb ITA	Feb MATH	Mar Age-in-Months	Mar ITA	Mar MATH	Apr Age-in-Months	Apr ITA	Apr MATH
South	0.003	-0.137***	-0.099***	0.006	-0.137***	-0.100***	0.005	-0.137***	-0.099***	0.004	-0.137***	-0.099***
	(0.188)	(-19.573)	(-13.935)	(0.359)	(-19.567)	(-13.971)	(0.298)	(-19.560)	(-13.927)	(0.258)	(-19.568)	(-13.926)
Eligible	-4.817***	0.083***	0.089***	-5.134***	0.133***	0.106***	-4.892***	0.124***	0.108***	-4.056***	0.102***	0.114***
	(-66.279)	(5.590)	(5.918)	(-88.949)	(8.492)	(6.542)	(-129.423)	(8.465)	(7.132)	(-122.988)	(6.757)	(7.365)
South*Eligible	6.417***	-0.160***	-0.144***	5.183***	-0.157***	-0.123***	3.457***	-0.087***	-0.084***	2.286***	-0.035	-0.059***
	(61.969)	(-7.502)	(-6.684)	(48.409)	(-6.948)	(-5.271)	(38.728)	(-4.098)	(-3.810)	(28.890)	(-1.619)	(-2.690)
Constant	8.364***	0.029***	0.021**	8.386***	0.030***	0.019**	8.393***	0.031***	0.023***	8.372***	0.028***	0.021**
	(392.369)	(3.540)	(2.561)	(398.453)	(3.618)	(2.291)	(409.171)	(3.843)	(2.761)	(420.882)	(3.431)	(2.561)
Observations	91,084	91,681	88,950	90,089	90,686	87,984	91,022	91,619	88,858	90,741	91,338	88,632
R-squared	0.154	0.006	0.004	0.147	0.007	0.004	0.159	0.006	0.003	0.129	0.006	0.004
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust t-statistics in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Where Jan is a dummy equal to 1 if student i was born in January and 0 otherwise, and observation of student born from February to April are dropped. I do the same for each month until April and see if the effect for each month is related to the change in the percentage of early starters. We would expect higher estimates for January than for April since the difference in number of early entrants is much bigger for born in the first month of the year.

Tables 11 reports the estimation for each month of eligibility and for the diff-in-diff estimation for both age-in-months and score in Italian and Mathematics.

As expected, the table shows that the diff-in-diff estimator for test scores decrease together with the one for age-in-month, when moving from January to April. It is interesting to notice that for each month, the diff-in-diff estimator for scores is close to the product of the linear effect of being one month younger and the North-South difference in age-in-months. This relationship holds for both II graders and V graders. One of the main questions this paper wants to investigate is indeed whether to be enrolled earlier gives a disadvantage to the pupils which is bigger (or lower) than the normal linear effect of age on scores which affects also not eligible students. The month-by-month analysis gives further evidence that this is not the case.

11. Possible determinants of Early Enrollment: Parenting Styles

So far I have investigated the effects of different parental choice on school enrollment on the gap in standardized scores between north and South Italy. However, nothing has been said about the reasons behind this heterogeneity. It seems that this behavior represents almost a norm among parents of children born between January and April. Difference in parental characteristics explains this choice only partially, and do not say nothing about the difference between the two areas. The main driver of the early enrollment seems indeed to be found in different cultural and behavioral characteristics between North and South parents. Here I will look a one characteristic: parenting style. The recent work by Doepke and Zilibotti (2017), categorizes parenting styles according to the classification proposed by developmental psychology (see Baumrind, 1967) and uses specific items from the World Value Survey (WVS) to divide parents among three different parenting styles: authoritative, authoritarian and permissive. In their work they show through a cross-country analysis, that the diffusion of a style is correlated with several economic outcomes such as inequality, social mobility and return to school. Where inequality is higher, parents tend to be more authoritarian and less permissive, the same happens when social mobility is low and returns to education are high. Parenting style is also correlated with fiscal policies, with more redistributive countries having more permissive parents. Diffusion of a particular parenting style in a country is also reflected by how the school system is designed. In countries with a majority of permissive parents, like Scandinavians pupils aren't graded until middle school (12-13 y.o.) and primary school starts when children are 7. Generally speaking, more permissive parents prefer that they children spend more time playing before starting primary school and do not feel competition with children of other parents. Here I will look at correlation between parenting style and the choice of early enrollment. In the Wave 5 of the World Value Survey, for each respondent is also indicated the region of residence within each country. Following the same categorization as Doepke and Zilibotti, I found that the 32% of parents in the North are permissive, whereas in the South they represent only the 19%. On the other hand, authoritative parents are the 23% (17% among parents who completed High School) in the North and the 35% in the South. This is already a first evidence on how parents in the two areas have different parenting styles. However, the World Value Survey used here is limited to one wave and to a sample size of one thousand people. To investigate further the role of parenting style in the choice of early enrollment I used the data from the PISA questionnaire for parents of 15 years old students who takes the test. The sample for Italy is much larger than the WVS but do not contain specific question about child qualities they consider important. As a result, is not possible to classify parents according to parenting styles. However, the data tells us how many activities they do with the child and with which frequency. The question is asked referring to the current period and to the time when the child starts primary school. The dataset itself contains two variables constructed using answers to these questions. These variables are standardized scores of parental support when the student is 15 years old (current parental support) and parental support at the beginning of primary school. Figure 7 e 8 show the difference in parental support among Italian regions and between parents of early entrants and of regular children. Figure 5 shows the average parental support at 1st grade and at the age of 15 for each Italian region. Only parents of students born between May and December are considered in this graph, since early entrance may affect the parental support in the first grade, as younger students could need more involvement by parents. We can see that parental support and the percentage of early entrants are positively correlated. Figure 8 instead, shows the same indexes for each macro area and for each school starting age. We can see that parental support is much higher in the South and for early enrollers.

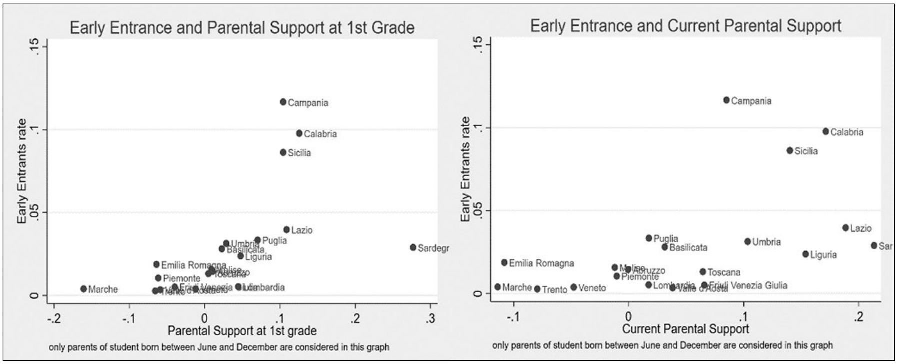

Fig. 5 – Early Entrance rate and parental support (regional level)

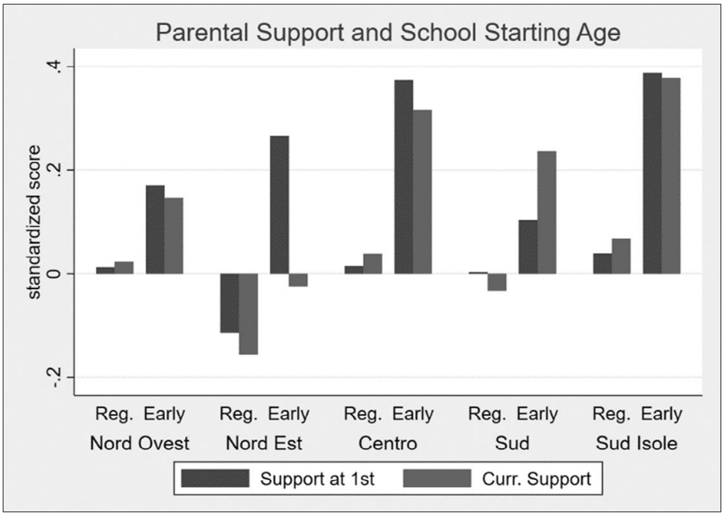

Fig. 6 – Early Entrance rate and parental support by macro-area and School Starting Age

12. Early Entrants Rate and other economic outcomes

As reported by Doepke and Zilibotti, permissive parenting style is more diffused where inequality and returns to education are low and social mobility is high. As possible reasons for these correlations, they argue that parents are less worried about their children future, and consequently less competitive, in those countries where the educational system puts less emphasis on grading students, access to college is less competitive and having a lower education is less penalizing. On the other hand, when returns to education are higher, parents put more attention on results and tend to that "Helicopter parenting" style introduced by Ginott (1969) that consists in being extremely involved in children's school life. As already mentioned, regions in the North of Italy presents lower inequality, higher social mobility and lower returns to education. Figure 7 shows the correlation between the share of early entrants and Social Mobility measured as in Guell *et al.* (2018) and the ratio inequality as measured by D'Onofrio and Giordani (2019). Social Mobility is measured by Informational Content of Surnames (ICS) which is an index strongly and negatively correlated with social mobility, inequality is instead measured by the ratio p75/p25 and results do not change much using other measures such as p90/p10 and the Gini Index.

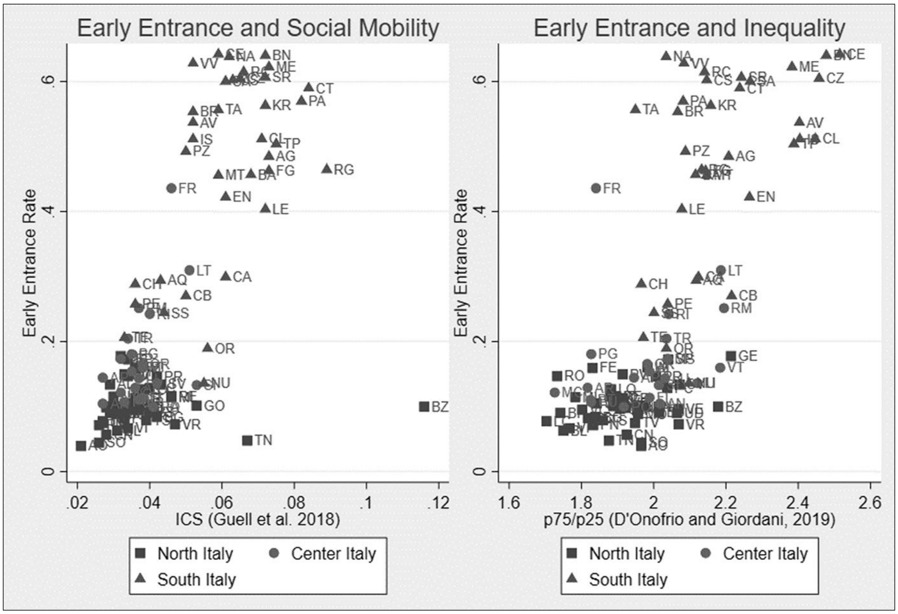

Fig. 7 – Early Entrance Rate, Social Mobility and Inequality at the province level

The scatter plots show the correlation between the percentage of early entrants and both social mobility and inequality (a higher ICS index reflects lower social mobility). The correlation remains even within macro areas. This finding is consistent with the hypothesis early entrance is correlated with a less permissive parenting style, and may reflects the fact that parents who choose early enrollment do so because they believe that one additional year of kindergarten (learning by playing, with children free to move in the class) is less useful for their children, compared to primary school (traditional lecture, children normally staying sit). Moreover, in the South of Italy, the effect of schooling on labor force participation and on employment are much higher than in the North (Brunello *et al.*, 2001; Ciccone *et al.*, 2006), and also this is consistent with the hypothesis that parenting style is correlated with early enrollment.

13. Conclusions

Statistics on academic results of Italian students are presented every year by the Ministry of Education. A lot of attention is given on the divide between Northern and Southern students. This work shows that a significant percentage of this gap is due to the difference in early enrollment. Even if the INVALSI reports recognize that in the South early enrollment is more common, this study is the first one to provide empirical evidence on the effects on scores of this heterogeneity in school starting age. Moreover, I provide evidence that even if more skilled pupils are selected into early enrollment, they do not recover better than the average. This could mean that a lot of potential from high skilled children is wasted by the choice of the parents of sending them too early to primary school. The geographic heterogeneity in early enrollment could not be explained simply by heterogeneity in parents' characteristics. Moreover, the difference is in the demand for early enrollment: school rules are very similar across regions and I do not find systematic difference in the supply of kindergartens. Consequently, it is likely that two different equilibria arise because of some kind of social norm, or because parents in the south have different beliefs and parenting style. In the last part of the paper I showed that early enrollment is positively correlated with a more authoritative and less permissive parenting style and that parents who choice the early enrollment are more involved in the school life of the children. Indeed, parents are less permissive in the South and this can be related to other economic outcomes such as social mobility, returns to education and inequality. Early enrollment is also correlated with lower social mobility and higher inequality and returns to education. However, further investigation is needed, for example about the role of social interaction in the choice of early enrollment, which is likely to be determinant.

The second important result of this paper lies in the analysis of the reform of the primary school enrollment system. The change in the rules made it easier for parents to enroll children earlier, as a result, the North-South gap in the average age of the same grade cohort increased. This exogenous variation generates an increase in the regional divide in scores at standardized test, creating a dramatic jump in the number of early starters, especially in the South. After the reform, even more high skilled pupils in the South have been sent to primary school earlier, potentially harming their skill formation.

Policymakers should take into account this phenomenon when studying academic achievement differentials within the country and this work suggests that the reform had a negative effect on students' performance. Not only it generated more within-class variation in age in the South, making the teachers' work harder, but also it lowers the academic performance of students.

The main conclusion of this paper is then that early enrollment is a negative practice and that parents should not have flexibility in the choice of school starting age.

References

Angrist J.D., Battistin E., Vuri D. (2017), "In a small moment: Class size and moral hazard in the Italian Mezzogiorno", *American Economic Journal: Applied Economics*, 9 (4), pp. 216-249.

Asso P., Azzolina L., Pavolini E. (2015), *L'istruzione difficile. I divari nelle competenze fra Nord e Sud*, Donzelli, Roma.

Baumrind D. (1967), *Child care practices anteceding three patterns of preschool behavior, Genetic Psychology Monographs*, 75 (1), pp. 4388.

Bedard K., Dhuey E. (2006), "The persistence of early childhood maturity: International evidence of long-run age effects", *Quarterly Journal of Economics*, 121 (4), pp. 1437-1472.

Bertoni M., Brunello G., Rocco L. (2013), "When the cat is near, the mice won't play: The effect of external examiners in Italian schools", *Journal of Public Economics*, 104, pp. 65-77.

Bierman K.L., Torres M.M., Domitrovich C., Welsh J.A. (2009), "Behavioral and cognitive readiness for school: Cross-domain associations for children attending head start", *Social Development*, 18 (2), pp. 305-323.

Black S.E., Devereux P.J., Salvanes K.G. (2011), "Too young to leave the nest? The effects of school starting age", *Review of Economics and Statistics*, 93 (2), pp 455-467.

Bound J., Jaeger D.A., Baker R.M. (1995), "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak", *Journal of the American Statistical Association*, 90 (430), pp. 443-450.

Brunello G., Comi S., Lucifora C. (1999), "The Returns to Education in Italy: A New Look at the Evidence", *IZA Discussion Paper*.

Ciccone A., Cingano F., Cipollone P. (2006), "The Private and Social Return to Schooling in Italy", *Bank of Italy Economic Research Paper*, No. 569.
Cornelissen T., Dustmann C. (2019), "Early school exposure, test scores, and non-cognitive outcomes", *American Economic Journal: Economic Policy*, 11 (2), pp. 35-63.
Cunha F., Heckman J. (2007), "The technology of skill formation", *American Economic Review*, 97 (2), pp. 31-47.
Doepke M., Zilibotti F. (2017). "Parenting With Style: Altruism and Paternalism in Intergenerational Preference Transmission", *Econometrica*, 85 (5), pp. 1331-1371.
D'Onofrio A., Giordani P.E. (2019), "Infrastructures and income inequality: The case of Italian provinces", *Politica Economica*, 35 (1), pp 27-54.
Elder T.E., Lubotsky D.H. (2009), "Kindergarten Entrance Age and Children's Achievement", *Journal of Human Resources*, 44 (3), pp. 641-683.
Fenoll A., Campaniello N., Monzon I. (2018), *Parental Love Is Not Blind: Identifying Selection into Early School Start*, Working Paper 286, Red Nacional de Investigadores en Economía (RedNIE).
Fredriksson P., Björn Öckert B. (2014), "Life-cycle effects of age at school start", *Economic Journal*, 124 (579), pp. 977-1004.
Güell M., Pellizzari M., Pica G., Rodríguez Mora J. V. (2018), "Correlating social mobility and economic outcomes", *The Economic Journal*, 128 (612), F353-F403.
Ginott H.G. (1969), Between Parent and Teenager, Scribner, New York.
Heckman J.J. (2008), "Schools, skills, and synapses", *Economic Inquiry*, 46 (3), pp 289-324.
Mühlenweg A., Blomeyer D., Stichnoth H., Laucht M. (2012), "Effects of age at school entry (ASE) on the development of non-cognitive skills: Evidence from psychometric data", *Economics of Education Review*, 31 (3), pp. 68-76.
Ponzo M., Scoppa V. (2014), "The long-lasting effects of school entry age: Evidence from Italian students", *Journal of Policy Modeling*, 36 (3), pp. 578-599.
Son S.E., Lee K., Sung M. (2013), "Links Between Preschoolers' Behavioral Regulation and School Readiness Skills: The Role of Child Gender", *Early Education and Development*, 24 (4), pp. 468-490.

5. *Sense of belonging at school and students with special educational needs: evidence from TIMSS 2019*

by Francesco Annunziata, Elisa Caponera, Laura Palmerio

In the last decades, different studies evidenced that a strong sense of belonging plays a key role in inclusion at school: the students with higher sense of belonging feel accepted, included, encouraged and supported within the classroom context. Furthermore, for children with special education needs (SEN), feeling part of the same group along with peers without special education needs plays a key role in their development.

The objective of this study was to investigate SEN students' sense of belonging with the one of their companions without SEN. Data from subsample of Italian students participating in TIMSS (Trend in international Mathematics and Science Study) 2019 – fourth grade- were used. An analysis of variance was conducted. The results show a difference in the sense of belonging between two different categories of SEN students: students certified with functional, cognitive, behavioral, or emotional disabilities and students with specific learning disorders (SLD): the former report a lower sense of belonging than SLD students, who instead feel to be an integral part of the school.

In order to integrate these students in the school context, it seems necessary for the school community to do further work to make them feel more involved and participate.

Further studies on the sense of belonging and more in general on well-being of students with special educational needs using large-scale data are needed to provide information to the school system.

Negli ultimi decenni, diversi studi hanno evidenziato il ruolo che gioca il senso di appartenenza nell'inclusione scolastica: gli studenti con un maggiore senso di appartenenza si sentono accettati, inclusi, incoraggiati e sostenuti nel contesto della classe, sono più motivati nell'apprendimento, partecipano più volentieri alle attività scolastiche.

In particolare, per i bambini con bisogni educativi speciali sentirsi rispettati e trattati in modo simile ai coetanei che non hanno bisogni educativi speciali potrebbe favorire lo sviluppo di un forte senso di appartenenza al gruppo.

Obiettivo del presente studio è indagare il senso di appartenenza scolastica degli studenti con bisogni educativi speciali, confrontandolo con quello degli studenti senza bisogni educativi speciali.

Sono stati utilizzati i dati di un sottocampione di 196 studenti partecipanti a TIMSS (Trend in international Mathematics and Science Study) 2019 – quarto grado. È stata condotta un'analisi della varianza, che ha evidenziato come gli studenti DIS hanno un senso di appartenenza statisticamente inferiore rispetto agli studenti con DSA (-0,24 vs 0,35). Al fine di integrare questi studenti nel contesto scolastico sembra necessario un ulteriore lavoro da parte della comunità scolastica per farli sentire più coinvolti e partecipi.

Ulteriori studi sul senso di appartenenza e più in generale sul benessere degli studenti con bisogni educativi speciali utilizzando dati su larga scala sono necessari per fornire informazioni utili alle istituzioni scolastiche.

1. Introduction

In the past, the concept of disability was associated with the person unable to manage their own independence; in need of protection and help to make up for their inabilities. The United Nations General Assembly in December 2006 enshrined a new concept of disability by focusing on upholding human dignity and rights to ensure full and effective participation in all social environments (Mahar *et al.*, 2013).

In this line, also the school policy recognizes the right of students with disabilities to be part of the school community. Regarding Italy, one basic goal of our school system is to promote the school inclusion, based on the assumption that all students, regardless of their functional or physical disabilities, have the right to develop their potential and to experience individual and social growth.

For students with disabilities, schools are fundamental in guaranteeing full integration. A recent report by ISTAT (2020) evidenced that peers play a key role, in terms of both relationships and learning, in the process of school inclusion. The development of supportive relationships can be an additional resource in the process of inclusion, and for this reason, it would be useful if all the educational activities of the student with disabilities were to be carried out in class together with his peers. In schools of the first cycle (grades 1 to 8), students with disabilities who have limitations in their autonomy spend most of their time in the classroom (27.6 hours per week) and carry out educational activities outside the classroom only for a residual number of hours (2.6 hours per week). However, if the student has more severe limitations, the number of teaching hours spent outside of the classroom increases considerably (6.4 hours per week)[1].

Unfortunately, because of the COVID-19 pandemic, the relational aspect and the role of peers has been greatly reduced in general and for these children in particular. The educational change focused on emergency responses to the crisis, and the vulnerable learners are more at risk, also because they are more affected by the digital divide. In this sense different actions have been implemented by countries to attempt to overcome these difficulties (European Agency for Special Needs and Inclusive Education, 2021).

Providing an environment in which children feel accepted is essential for the development of a sense of belonging in the school context and hence for inclusion: trying to understand how students with disabilities are integrated in class and feel part of the school becomes crucial.

2. History of school inclusion in Italy

The Italian Constitution provides for the guarantee of schooling for all (Article 34) and prescribes the fulfilment of the mandatory duty of solidarity (Article 2). It also states that it is "the task of the Republic to remove obstacles that limit the freedom and equality of citizens, in order to ensure the full development of the human person" (Art. 3).

Italy has always been at the forefront, promoting the inclusion of all students. The current set of practices that ensure the inclusion of all students in Italian schools is the result of a long process that began in the 70s and is continuously in progress to allow the full integration of students. In fact, Italian legislation, widely considered among the most advanced both in Europe and worldwide, has long abandoned its original welfare approach to embrace the concepts of inclusion and social participation (Karagiannis, Stainback and Stainback, 1996; Kauffman, 1999; Kavale and Forness, 2000; Kanter, Damiani and Ferri, 2014; Carnovali, 2017).

With the start of the inclusion process, the school radically changes its perspective. In this new approach, the focus is no longer only on the most intellectually gifted students, but on everyone. In fact, social, economic, and

[1] For detailed description see https://www.istat.it/it/files/2020/02/Alunni-con-disabilita-2018-19.pdf.

learning disadvantage is considered a condition to be overcome through the use of specific resources.

Although oriented towards an emancipation from the regime of exclusion and segregation of pupils with disabilities, in some countries – including Italy and France – at the beginning of the 1970s, there were still “adaptation classes”, where children with difficulty would have to stay for a short time before being integrated into regular classes. Despite the persisting limits set by the old concept of learning, Italy managed to take the first step, allowing the entrance of children with less severe disabilities into mainstream classes (Ciambrone, 2017).

This was made possible, in 1971, thanks to Law 118 that established the right to education in the mainstream schools, except in cases where the students are affected by serious – either intellectual or physical – disabilities that make very difficult to learn or to stay in these mainstream classes. To support the feasibility of the inclusion process, several facilities are provided: free transport to and from school, access to school through appropriate measures to overcome and eliminate architectural barriers, assistance during school hours of the most serious disabled (Pavone, 2012).

The application of Law 118/71 led to a heated debate between those who were in favor of full inclusion of all students and those who favored school inclusion only for students with less serious problems.

In 1975, the Ministerial Commission chaired by Senator Franca Falcucci presented a report on the process of inclusion in Italian schools. The document lays out fundamental principles that allow this moment to be defined as the beginning of the school integration phase: collegiality, the key role of the family, the integrated management of services, and teacher training.

In 1977, with the enactment of Law No. 517, we witnessed the final abolition of special classes, which gave full implementation to the principle of inclusive education, a concept that goes beyond that of simple inclusion in “normal” classes. The Law No. 517 aimed to achieve not only formal, but also substantial, equality through the provision of programs implemented by all teachers and the introduction of the role of the special education teacher. (Cecchini and McCleary, 1985; Daniels and Hogg, 1991; Abbring and Meijer, 1994; Vitello, 1994; Berrigan and Taylor, 1997; Paparella, 2010; D’Alessio, 2011; Arconzo, 2013; Troilo, 2013; Carnovali, 2017).

Special education teachers are professionals capable of carrying out remedial action with students with severe disabilities and of alternating rehabilitative and specialized interventions with students with average difficulties.

The sentence of the Constitutional Court n. 215 in 1987 recognized the full right of students with disabilities to attend upper secondary school stating that "being able to participate in the educational process together with teachers and peers with no disabilities is a significant factor in socialization, recovery and overcoming exclusion". Finally, in 1988, the Ministerial Circular No 153 reiterated that "it is unlawful to educate the disabled pupils by making them leave their class, except in cases where a defined period of activity outside the class is expressly provided by the Individualized Education Plan and agreed between specialized teacher and curricular teachers".

In 1992, Law 104 guaranteed full respect for human dignity and rights for freedom and autonomy of people with disabilities, and promoted their full integration in the family, school, work, and society. Full school inclusion is finally achieved.

In relation to the potential and initial learning levels of the student (Law No. 104/1992), the Individualized Educational Plan is composed of four dimensions (Socialization and Interaction; Communication and Language; Autonomy and Orientation; Cognitive, Neuropsychological, and of Learning) and for each of the dimensions the expected objectives and outcomes and the educational and methodological actions – organized in activities, strategies and tools – are to be identified.

In 2000, Law 328 defined the "integrated system of interventions and social services" aiming at promoting social interventions, assistance and social and health care interventions that guarantee concrete help to individuals and families in difficulty, and in 2003, Law 53 defined the essential levels of provision in education and training.

Disabilities show up in different forms, some mild, others severe, and for each of them it is necessary to identify mechanisms to remove obstacles according to the type of disability. Law 170/2010 recognized dyslexia, dysgraphia, dysorthographia and dyscalculia as specific learning disorders (SLD) that occur in the presence of adequate cognitive skills, in the absence of neurological disorders and sensory deficits, but can be a major limitation for some daily-life activities. The Ministerial Directive of 27.12.2012, widened the range of problems included in the category of special education needs (SEN), extending it to the area of socioeconomic disadvantage, language, culture and difficulties arising from the lack of knowledge of Italian culture and language because they belong to different cultures. The integrated system of interventions and social services defined in Law 328/2000 can be applied to migrants as well.

SEN children manifest special learning needs, which may be permanent or temporary and caused by different factors: social and cultural disadvantage, specific learning disorders and/or specific developmental disorders, and difficulties arising from not knowing Italian culture and language (CM n. 8/2013). It is possible to differentiate three categories of SEN:

– Certified disabilities (Law 104/92 art. 3, paragraphs 1 and 3): sight-impaired, hearing-impaired, and psychophysical disables.
– Specific developmental disorders (Law 170/2010): SLD, Attention-deficit/hyperactivity disorder (ADHD), Oppositional Defiant Disorder, borderline cognitive impairment, school phobia, conduct disorder, ecc.
– Disadvantage (DM of 27.12.12 and CM n.8 of 6.03. 2013): socioeconomic, linguistic-cultural, behavioral/relational distress, etc.

More recently, in 2010, Law 170, the following implementing decree (DM 5669 of 12/07/2011), and related Guidelines introduced the mandatory annual drafting of a Individualized Educational Plan for students with SLD, indicating the compensatory tools and dispensatory measures adopted.

Decree n.66/2017, updated by Decree 96/2019, represents the most recent step on the path towards the realization of school inclusion.

In this decree, the Individualized Educational Plan is defined as a document that encompasses the individualized planning for each student with disabilities to ensure school inclusion. The recent Interministerial Decree No. 182 of 2020 has defined the procedures for the assignment of the support measures provided by Legislative Decree No. 66 of 2017 and the models of the Individualized Education Plan.

The approach to inclusion in Italian schools is to minimize the number of SEN students in each classroom. In fact, in composing the classes, "the principal shall ensure a fair distribution of SEN students among the various classes and, in case of the presence of more than two SEN students in a class, the class shall be formed with no more than 20 students". Specifically, in the 2018/2019 school year the average number of SEN children per class with at least one SEN pupil is 1.40. In primary schools we have 1.37 SEN pupils per class.

3. Inclusion Policy in Europe

This approach does not seem obvious; even today, many countries, including some European ones (e.g., Germany, Belgium, and the Netherlands), have not adopted the participatory approach; indeed, in these contexts, the education of students with disabilities is taught in special schools or special classes (Ainscow and Haile-Giorgis, 1998; Vislie, 2010; Powell, 2011; Carnovali, 2017).

Even though the current tendency in Europe is to develop a policy towards inclusion of SEN pupils into mainstream schools, and to support teachers with additional staff, tools, and in-service training, the policies adopted in this regard vary across countries. More in detail, while the inclusion of pupils with mild disabilities in the common classes is widespread, this is not the case for the inclusion of persons with medium to severe intellectual disabilities.

Table 1 shows the percentage of SEN students that are in mainstream classes across some European countries. In general, it is possible to evidence three different approaches (European Agency for Development in Special Needs Education, 2020):

– One-track approach. Countries focused on the mainstream school (e.g., Spain, Greece, Italy);
– Multitrack approach. In these countries there is a variety of services between the two systems (i.e. mainstream and special education needs systems) (e.g., France, the United Kingdom);
– Two-track approach. There are two distinct education systems. Pupils with SEN are usually placed in special schools or special classes (e.g. Switzerland).

Tab. 1 – Inclusion in some European countries – Primary school

Country	*School system*	*School years*	*% of SEN students in mainstream classes*	*% SEN students*
Italy	One-track	2019/20	> 99	3.7
England (UK)	Multitrack	2019/20	45	2.9
France	Multitrack	2019/20	65	4.0
Germany	Multitrack	2019/20	52	6.3
Greece	One-track	2019/20	86	5,5
Netherlands	Multitrack	2019/20	0	2.4
Spain	Multitrack	2019/20	78	3.4
Switzerland	Two-tracks	2019/20	54	4.0

Source: European Agency for Development in Special Needs Education (2020)

4. Sense of belonging

Students with SEN generally report high levels of rejection, perceive themselves to be unaccepted, and, compared to peers, perceive more feelings of loneliness and poor friendships with one another. Moreover, students with SEN have a lower level of belonging than their non-SEN peers (e.g., Cullinane, 2020).

In literature, there are different definitions of sense of belonging (e.g, Allen and Bowles, 2012; Allen *et al.*, 2016; Slaten *et al.*, 2016), one of the most established ones was Goodenow and Grady's (1993), who define school belonging as "the extent to which students feel personally accepted, respected, included, and supported by others in the school social environment" (p. 80).

In the school setting, a sense of belonging is a factor related to increased motivation and achievement. Furrer and Skinner (2003) stated that a sense of belonging in the school setting influences performance through the "energy function", stimulating enthusiasm, interest, and willingness to participate in school activities. In support of this theory, McGraw (2008) found that a lack of the familiar sense of school belonging is associated with depressive symptoms.

In a study conducted by Banks *et al.* (2018), relationships between SEN and non-SEN students in Ireland were examined based on the quantity and quality of peer relationships. The findings from this study showed that, after controlling for a series of variables, students with particular disadvantages were more likely to have fewer friends and experience negative (peer) relationships than their peers. In the same direction, similar research, e.g., Frostad and Pijl (2007) in Norwegian schools, and Kostel *et al.* (2010), in the Netherlands, found that, on average, SEN students have fewer friends and belong less often to a group of friends. However, these studies acknowledge that the number of friends does not reflect the quality of friendship or the level of acceptance of students with SEN within a mainstream setting (Banks *et al.*, 2018).

The qualitative approach focuses primarily on students' experience of feeling part of a social network with peers. Several research studies find that children with SEN in mainstream classrooms are less accepted mainly in function of the type of SEN: students with autism spectrum disorders, communication, or behavioral problems have a much higher risk of social isolation (Bossaert *et al.*, 2015; Cook, Ogden and Winstone, 2017; Locke *et al.*, 2010; Pijl, Frostad and Flem, 2008; Banks *et al.*, 2018).

Friendships in the school setting represent a fundamental building block in the construction of children's self-image. The poor qualitative-quantitative set-up of friendships in the school setting can lead to a negative impact on children's social and emotional well-being with consequences for motivation, engagement, and academic performance (Perdue, Manzeske and Estell, 2009), thus leading to a very high risk of social exclusion in the future, even as adults (Banks *et al.*, 2018).

The sense of belonging is also one aspect that can contribute to explain the well-being at school and good quality friendships.

In the last decades, different studies evidenced that a strong sense of belonging in students promotes greater engagement and active willingness to participate in school activities and plays a key role in inclusion and school:

the students with higher sense of belonging feel accepted, included, encouraged and supported within the classroom context. These students are more motivated in learning, participate more willingly in school activities, have higher study expectations, better social relationships, and a better academic performance (Goodenow, 1993; Osterman, 2000; Dimitrellou and Hurry, 2019). In general, belonging can be seen as a basic human need that, when positive, helps an individual function effectively and feel motivated to be part of a community (Baumeister and Leary, 1995).

When the need for belonging is not met, there can be significant consequences, including an impact on intellectual achievement, hindering the realization of the learning potential. For students with marginalized identities, an intensification of factors that create challenges puts them at risk for disengagement, and their sense of belonging in school is more likely to be compromised.

Pendergast *et al.* (2018), in fact, through a thematic content analysis based on a series of interviews and focus groups conducted with 25 students (ages 12-16), 25 of their teachers, and 39 school leaders in Australia, found five themes related to sense of belonging at school: 1) Relationships at school, e.g., students can realize their potential, building relationships among students, teachers and parents, and the school community; 2) School climate factors, e.g., attendance, primary to high school transition, and school culture; 3) Pedagogical practices, e.g., assessments and grades, academic engagement, personalized learning, and curriculum differentiation; 4) Specific programs and activities, e.g., interventions and programs designed to build students' capacity to make connections and build relationships; these were often in collaboration with other agencies; 5) Other issues, e.g., family, mental health, trauma, and poverty, that impacted a student's sense of belonging to school (Pendergast *et al.*, 2018).

Using focus groups, Foley *et al.* (2012) examined narratives about the obstacles and coping skills of children with a disability. One of the most important themes that emerges is that they want to feel like they belong. In fact, they reported that for many children, the perception of social exclusion is more troublesome than the physical restrictions associated with their disability. Several psychologists have confirmed that the need to belong is a fundamental social motive and carries negative consequences when threatened (Daley, Phipps and Branscombe, 2018). Therefore, positive interactions with peers are crucial to these children because allow them to feel respected and accepted as part of a group (Rose and Shevlin, 2017; Crouch, Keys and McMahon, 2014).

Based on the previous literature, the main aim of this study was to investigate the sense of belonging of SEN students in comparison with the one of their companions without SEN, before the COVID-19 pandemic.

5. Method

5.1. Participants and data

To investigate the students' sense of belonging, data from TIMSS (Trend in international Mathematics and Science Study) 2019 – fourth grade – were used.

TIMSS (Trends in International Mathematics and Science Study) is a long-standing international assessment of Mathematics and science at the fourth and eighth grades promoted by the International Association for the Evaluation of Educational Achievement (IEA). It has been collecting trend data every four years since 1995. TIMSS measures trends in Mathematics and science student achievement, and studies the differences in national education systems in almost 60 countries in order to help improve teaching and learning worldwide.

TIMSS uses a two-stage sampling design. In the first stage, schools are selected. In the second stage, within each sampled school, one or two classes are randomly selected. All students in the sampled class participate in the survey.

The following instruments are used in TIMSS:

- Test on student proficiency in science and Mathematics, including both multiple-choice and open-ended questions. The analyses of students' responses to these questions allow not only to establish possible relationships between the different skills investigated, but also to identify any differences in student performance in a diachronic dimension;
- Student Questionnaire, including background variables related to the student socioeconomic and cultural status, to the students' attitudes towards Mathematics and Science, to the students' familiarity with information and communication technologies, and to the students' sense of belonging at school;
- Teacher Questionnaire, addressed to the teachers of the sampled students, collects information about the school and classes, the background of the Mathematics and Science teachers, their attitude towards the subjects taught, the way they teach, teaching practices;
- School Questionnaire, aimed at school principals, who are asked to provide information regarding the school context and climate, available resources, classes and teachers, and the involvement of students' parents in school activities;
- National Curriculum Questionnaire: a group of experts at the national level respond to a questionnaire that collects information about the Math and science curriculum used in each country;

– Encyclopedia: additional information is provided by each country regarding the social and demographic context of the country, the organization and structure of the education system, and the resources available for education.

In TIMSS, SEN students are divided in two different groups in function of their disabilities:

– Students certified with functional, cognitive, behavioral, or emotional disabilities (from now on DIS students). They are normally excluded from taking the test. However, for inclusion reasons, these students were given the opportunity to take the test and complete the context questionnaire, but the results were not considered in the national and international report;
– Students with specific learning disorders (SLD students). They take the test using the same aids they usually use in school activities and their data are included in the national and international datasets.

The overall sample of TIMSS fourth grade was 6,855. Of these, 30 SLD students who took the test using aids were considered. In addition, we included in this analysis 69 students who took the test but were excluded from the international database because of their special education needs. To compare the results of these students with the students who took the test regularly, and to control for the bias due to comparing students who attended different schools and/or classes, we decided to select a random sub-sample of 97 of all the students who attended the same class as the DIS students. Hence, the overall student sample considered in the analysis was of 196: 97 students who took the test regularly; 30 SLD students; and 69 DIS students.

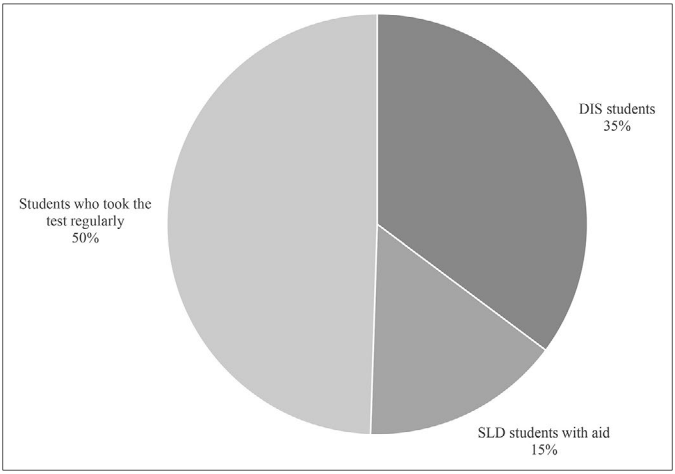

Fig. 1 – Description of the sample

6. Measures

Sense of belonging. Four variables found in the student questionnaire were considered for the analysis (for a detailed description see, Martin, von Davier and Mullis, 2020): "I like being in school"; "I feel safe when I am at school"; "I feel that I belong to this school"; "I am proud to attend this school". Students had to indicate their agreement/disagreement on a 5-point Likert scale ranging from "Disagree a lot" to "Agree a lot". The mean was zero, and the standard deviation was 1.

7. Results

7.1. Descriptive statistics

Tab. 2 – Percentage of students' agreement with the different statements on sense of belonging at school

	I feel safe when I am at school	*I feel that I belong to this school*	*I am proud to attend this school*	*I like being in school*
Students took the test regularly	81.4	78.1	95.9	89.7
DIS students	77.3	87.9	86.6	80.6
SLD students	70.0	100.0	86.7	80.0
Overall sample	80.8	92.7	92.8	74.7

In all four items considered students reported a high sense of belonging.

The percentage of students' agreement is greater than or equal to 70% in all four items considered.

More than half of the students say they agree or completely agree with each of the four items considered.

Based on these four variables, we carried out maximum likelihood exploratory factor analysis on the overall sample: the data showed a single-factor structure explaining 53% of the variance and with an adequate Cronbach's alpha (.74).

The factor loadings of the four items are high, ranging from 0.50 to 0.80. The regression scores obtained by the factor analysis are standardized using T-scores, where a score of 50 represents the mean. A difference of 10 from the mean indicates a difference of one standard deviation.

7.2. ANOVA

To better investigate whether there were statistically significant differences, a one-way analysis of variance (ANOVA) was conducted, with the type of students (three levels: DIS students; SLD students; Students who took the test regularly) as between factor and students' sense of belonging as dependent variable.

The purpose was to test whether the level of sense of belonging of students who regularly took the test was higher than DIS students and SLD students.

The results of the ANOVA (see Fig. 2) show a difference in the sense of belonging between students in the different categories ($F_{[217;\,2]} = 3.961$; $p < 0.02$).

Specifically, a post-hoc analysis was performed using the Bonferroni method, where the observed significance level is adjusted for the fact that multiple comparisons are being made.

Tab. 3 – Tests of between-Subjects Effects

Source	*Type III Sum of Squares*	*df*	*Mean Square*	*F*	*Sig.*
Corrected model	8.768	2	4.384	3.961	.021
Intercept	1.036	1	1.036	.936	.335
Type of students	8.768	2	4.384	3.961	.021
Error	208.101	188	1.107		
Total	217.213	191			
Corrected Total	216.870	190			

Tab. 4 – Pairwise Comparisons (Bonferroni Method)

(I) Type of students	*(J) Type of students*	*Mean Difference (I-J)*	*Std. Err.*	*Sig.*	*95% Confidence Interval for Difference*	
					Lower Bound	*Upper Bound*
Students took the test regularly (1)	2	.37	.17	.08	-.03	.78
	3 (SLD)	-.21	.22	1.00	-.74	.32
DIS (2)	1	-.37	.17	.08	-.78	.03
	3 (SLD)	**-.59**	**.23**	**.04**	**-1.15**	**-.03**

Coefficient in bold is significant at $p < 0.05$

The results show that DIS students had a significantly lower sense of belonging than SLD students (-.24 *vs* .35).

In contrast, the differences between students who took the test regularly and each of the two groups of SEN students were not significant (see Tab. 4).

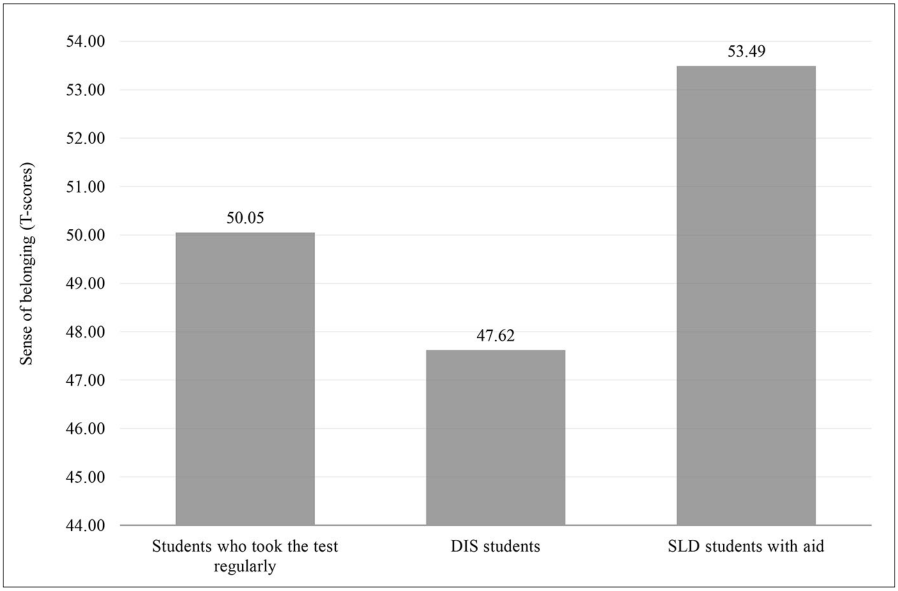

Fig. 2 – Students' sense of belonging at school

8. Discussion and Conclusions

In this study, the results concerning the inclusion of SEN students seem to indicate two different patterns: DIS students feel less integrated in school than SLD students who instead show a higher sense of belonging to their school. As highlighted in the literature, there is a strong link between students' sense of belonging, school inclusion, academic performance, and, more generally, well-being at school (Goodenow, 1993; Osterman, 2000). Hence, the increased sense of belonging of SLD students is an encouraging finding, highlighting how these students feel as an integral part of the school.

A different situation emerged for students certified with functional, cognitive, behavioral, or emotional disabilities. The DIS students considered in this study have severe functional, cognitive, behavioral or emotional disabilities (certified) disabilities, which may make it difficult for these students to fully experience the school context. In addition, pupils with behavioral difficulties may lack social skills, which negatively affect their ability to build and maintain satisfactory social relationships (Frostad and Pijl, 2007).

The results obtained in this study highlighted that DIS students experience the school, social, and relational context differently than their peers, presenting lower levels of belonging. As confirmed in the literature, this can have a negative impact on educational and social development.

In addition, for individuals with severe special education needs, adapting their behavior in order to feel a sense of belonging (Juvonen, 2006) might mean hiding the unique characteristics of a person's special needs (Pesonen, 2016).

Another possible hypothesis is that SLD students benefit from an intervention that is specifically targeted to their difficulties and that succeeds in involving the entire school community (students and teachers in particular) leading to true integration of the students. Whereas with DIS students it is more difficult for the school to take full ownership of a project that succeeds in involving the school or at least the class.

School well-being and inclusion of students with disadvantages are fundamental factors in student development and cannot be underestimated; they require constant monitoring and support. Pesonen (2016) also argues that a strong sense of belonging is fostered by a multidisciplinary collaboration between teachers that helps create a school climate of acceptance and support. Based on this consideration, in addition to moving beyond common classes to facilitate the integration process, the results of our study suggest the need for additional interventions targeted to the specific needs of DIS students, such as increased consultation with students about their needs and the supports they require, correlated by educational activities involving all students and tailored to the specific needs of DIS students.

A limitation of the present study is that it is a correlational study and therefore provides a snapshot of a particular point in time. Further in-depth studies should be done to confirm the present findings, including qualitative studies such as examining the school Self Evaluation Reports (RAV), to try to understand the factors that may improve the sense of belonging of DIS students within the school context. Also, a future study would benefit from a larger sample that would allow for a comparison across school grades, with a focus on teachers' and students' views.

Despite these limitations, we hope that the results of this study can be a pioneer of further studies on the sense of belonging and more in general on well-being of students with special educational needs using large-scale data, that usually are targeted to students with no particular disadvantages.

References

Abbring I., Meijer C.J.W. (1994), "Italy", in C.J.W. Meijer, S.J. Pijl, S. Hegarty (*eds.*), *New Perspectives in Special Education*, Routledge, London, pp. 9-24.

Ainscow M., Haile-Giorgis M. (1998), *The Education of Children with Special Needs: Barriers and Opportunities in Central and Eastern Europe*, UNICEF, Firenze.

Allen K.A., Bowles B. (2012), "Belonging as a guiding principle in the education of adolescents", *Australian Journal of Educational & Developmental Psychology*, 12, pp. 108-119.

Allen K., Kern M.L., Vella-Brodrick D., Hattie J., Waters L. (2016), "What schools need to know about fostering school belonging: A meta-analysis", *Educational Psychology Review*, 30, pp. 1-34.

Arconzo G. (2013), "La normativa a tutela delle persone con disabilità nella giurisprudenza della Corte costituzionale" [The legislation for the protection of people with disabilities in the case law of the Constitutional Court], in M. D'Amico, G. Arconzo (a cura di), *Università e persone con disabilità. Percorsi di ricerca applicati all'inclusione a vent'anni dalla legge n. 104 del 1992*, FrancoAngeli, Milano, pp. 17-32.

Banks J., McCoy S., Frawley D. (2018), "One of the gang? Peer relations among students with special educational needs in Irish mainstream primary schools", *European Journal of Special Needs Education*, 33 (3), pp. 396-411.

Baumeister R.F., Leary M.R. (1995), "The need to belong: desire for interpersonal attachments as a fundamental human motivation", *Psychological Bulletin*, 117 (3), pp. 497-529.

Berrigan C., Taylor D. (1997), "Everyone belongs: School inclusion and social relationships in Italy", *TASH Newsletter*, pp. 21-22.

Bossaert G., Colpin H., Pijl S.J., Petry K. (2015), "Quality of Reciprocated Friendships of Students with Special Educational Needs in Mainstream Seventh Grade", *Exceptionality*, 23, pp. 54-72.

Carnovali S. (2017), "The Right to Inclusive Education of Persons with Disabilities in Italy. Reflections and Perspectives", *Athens Journal of Education*, 4 (4), pp. 315-326.

Cecchini M., McCleary I.D. (1985), "Preschool handicapped in Italy: A research-based developmental model", *Journal of the Division for Early Childhood*, 9, pp. 254-265.

Ciambrone R. (2017), "Evoluzione dell'inclusione scolastica in Italia e in Europa a 40 anni dalla legge 517. Problemi e prospettive", *L'integrazione scolastica e sociale*, 16 (4), pp. 390-402.

Cook A., Ogden J., Winstone N. (2017), "Friendship Motivations, Challenges and the Role of Masking for Girls with Autism in Contrasting School Settings", *European Journal of Special Needs Education*, 33 (3), pp. 302-315.

Crouch R., Keys C.B., McMahon S.D. (2014), "Student-teacher relationships matter for school inclusion: School belonging, disability, and school transitions", *Journal of Prevention & Intervention in the Community*, 42 (1), pp. 20-30.

Cullinane M. (2020), “An Exploration of the Sense of Belonging of Students with Special Educational Needs”, *Journal of Inclusive Education in Ireland*, 33 (1), pp. 2-12.

D’Alessio S. (2011), *Inclusive Education in Italy: A critical analysis of the policy of “integrazione scolastica”*, Sense Publishers, Rotterdam.

Daley A., Phipps S., Branscombe N. R. (2018), The social complexities of disability: discrimination, belonging and life satisfaction among Canadian youth, *SSM-population Health*, 5, pp. 55-63.

Daniels H., Hogg B. (1991), “An intercultural comparison of the quality of life of children and youth with handicaps in Denmark, Italy, the United Kindom, and Germany”, *Educational and Child Psychology*, 8, pp. 74-83.

Dimitrellou E., Hurry J. (2019), “School belonging among young adolescents with SEMH and MLD: the link with their social relations and school inclusivity”, *European Journal of Special Needs Education*, 34 (3), pp. 312-326.

European Agency for Special Needs and Inclusive Education (2021), *The Impact of COVID-19 on Inclusive Education at the European Level: Literature Review* (C. Popescu, ed.), Odense, Denmark.

Foley K.R., Blackmore A.M., Girdler S., O’Donnell M., Glauert R., Llewellyn G., Leonard H. (2012), “To feel belonged: The voices of children and youth with disabilities on the meaning of wellbeing”, *Child Indicators Research*, 5 (2), pp. 375-391.

Frostad P., Pijl S.J. (2007), “Does being friendly help in making friends? The relation between the social position and social skills of pupils with special needs in mainstream education”, *European Journal of Special Needs Education*, 22 (1), pp. 15-30.

Furrer C., Skinner E. (2003), “Sense of relatedness as a factor in children’s academic engagement and performance”, *Journal of Educational Psychology*, 95 (1), pp. 148-162.

Goodenow C. (1993), “The psychological sense of school membership among adolescents: Scale development and educational correlates”, *Psychology in the Schools*, 30 (1), pp. 79-90.

ISTAT (2020), L*’inclusione scolastica degli alunni con disabilità. Anno scolastico 2018-2019*, Roma.

Juvonen J. (2006), “Sense of Belonging, Social Bonds, and School Functioning”, in P.A. Alexander, P.H. Winne (*eds.*), *Handbook of educational psychology,* Lawrence Erlbaum Associates Publishers, Mahwah, 2nd ed., pp. 655-674.

Kanter A.S., Damiani M.L., Ferri B.A. (2014), “The Right to Inclusive Education under International Law: Following Italy’s Lead”, *Journal of international Special Needs Education*, 17, pp. 21-32.

Karagiannis A., Stainback W., Stainback S. (1996), “Rationale for inclusive schooling”, in S. Stainback, W. Stainback (*eds.*), *Inclusion: a guide for educators*, Brookes, Baltimore, pp. 3-16.

Kauffman J.M. (1999), “Commentary: Today’s special education and its messages for tomorrow”, *The Journal of Special Education*, 32, pp. 244-254.

Kavale K.A., Forness S.R. (2000), “History, rhetoric and reality: Analysis of the inclusion debate”, *Remedial and Special Education*, 21, pp. 279-296.
Koster M., Pijl S.J., Nakken H., Van Houten E. (2010), “Social Participation of Students with Special Needs in Regular Primary Education in the Netherlands”, *International Journal of Disability, Development and Education*, 57 (1), pp. 59-75.
Locke J., Eric H., Ishijima E., Kasari C., London N. (2010), “Loneliness, Friendship Quality and the Social Networks of Adolescents with High-functioning Autism in an Inclusive School Setting”, *Journal of Research in Special Educational Needs*, 10 (2), pp. 74-81.
Mahar A.L., Cobigo V., Stuart H. (2013), *Conceptualizing belonging. Disability and rehabilitation*, 35 (12), pp. 1026-1032.
Mcgraw K., Moore S., Fuller A., Bates G. (2008), “Family, peer and school connectedness in final year secondary school students”, *Australian Psychologist*, 43 (1), pp. 27-37.
Osterman K. F. (2000), “Students’ need for belonging in the school community”, *Review of Educational Research*, 70 (3), pp. 323-367.
Paparella E. (2010), “Il diritto all’integrazione e all’istruzione scolastica dei soggetti diversamente abili: una fondamentale declinazione del diritto allo studio nella prassi amministrativa e nelle recenti pratiche governative” [The right to integration and school education of disabled people: a fundamental declination of the right to study in administrative practice and in recent government practices], in *Studi in onore di Vincenzo Atripaldi*, Jovene, Napoli, pp. 961-987.
Pendergast D., Allen J., McGregor G., Ronksley-Pavia M. (2018), “Engaging marginalized, at-risk middle-level students: A focus on the importance of a sense of belonging at school”, *Education Sciences*, 8 (3), pp. 1-19.
Perdue N.H., Manzeske D.P., Estell D.B. (2009), “Early Predictors of School Engagement: Exploring the Role of Peer Relationships”, *Psychology in the Schools*, 46 (10), pp. 1084-1097.
Pesonen H. (2016), *Sense of belonging for students with intensive special education needs: An exploration of students’ belonging and teachers’ role in implementing support*, Research report 380, University of Helsinki, Helsinki.
Pijl S.J., Frostad P., Flem A. (2008), “The Social Position of Pupils with Special Needs in Regular Schools”, *Scandinavian Journal of Educational Research*, 52 (4), pp. 387-405.
Powell J.J.W. (2011), *Barriers to Inclusion. Special Education in the United States and Germany*, Routledge, New York.
Rose R., Shevlin M. (2017), “A sense of belonging: Childrens’ views of acceptance in ‘inclusive’ mainstream schools”, *International Journal of Whole Schooling*, 13 (1), pp. 65-80.
Slaten C.D., Ferguson J.K., Allen K.A., Brodrick D.V., Waters L. (2016), “School belonging: A review of the history, current trends, and future directions”, *The Educational and Developmental Psychologist*, 33 (1), pp. 1-15.
Troilo S. (2013), “I ‘nuovi’ diritti sociali: la parabola dell’integrazione scolastica dei disabili” [The “new” social rights: the parable of the integration of disabled

people], in *I diritti sociali dal riconoscimento alla garanzia: il ruolo della giurisprudenza: atti del Convegno di Trapani 8-9 giugno 2012*, *Quaderni del Gruppo di Pisa*, Editoriale Scientifica, Napoli, pp. 535-544.
Vislie L. (2010), "From integration to inclusion: focusing global trends and changes in the western European societies", *European Journal of Special Needs Education*, 1 (18), pp. 17-35.
Vitello S.J. (1994), "Special education integration: The Arezzo approach", *International Journal of Disability Development and Education*, 41, pp. 61-70.

6. *The digital divide: a challenge for the schools*

by Rita Marzoli, Ornella Papa, Lorenzo Mancini

During the COVID-19 pandemic, digital skills became crucial for schooling. However, Italian students had not yet adequately developed these skills, which has contributed to raising the risk of educational poverty and student inequalities (OECD, 2020; Fraillon *et al.*, 2020). The present study refers to data from the Italian ICILS 2018, which was collected by INVALSI in collaboration with the IEA. The sample comprises 146 schools and 2,810 students with an average age of 13.26 years. The study investigates the relationship between Computer Information Literacy (CIL) results and students' socioeconomic background, while also considering the digital resources of the schools. The research hypothesis is that schools effectively counter the digital divide in the presence of specific resources and adequate learning practices. The results of the study confirm lower average scores of students attending schools located in disadvantaged contexts and with unfavourable characteristics for learning CIL, such as reduced access to the Internet and less pedagogical support for the use of ICT. Similarly, students from low socioeconomic backgrounds perform better when they learn digital and informative skills at school, in an educational environment with conditions that allow for it. These results highlight the role of schools in fighting the digital divide, particularly in geographical regions with more disadvantaged catchment areas.

Durante la pandemia di Covid19, le competenze digitali sono diventate cruciali per le scuole, sebbene gli studenti italiani non avessero ancora sviluppato adeguatamente queste competenze (Fraillon et al.*, 2020). Tale carenza ha contribuito ad aumentare il rischio di povertà educativa e di disuguaglianze tra gli studenti (OECD, 2020). Il presente studio fa riferimento ai dati italiani ICILS 2018 raccolti dall'INVALSI in collaborazione con la IEA. Il campione è* *composto da 146 scuole e 2810 studenti con un'età media di 13, 26 anni. L'oggetto indagato è la relazione tra i risultati in Computer Information Literacy e il background socioeconomico degli studenti, considerando le risorse delle scuole in ambito digitale. L'ipotesi di ricerca è che le scuole contrastino efficacemente il divario digitale in presenza di risorse specifiche e pratiche adeguate di apprendimento. I risultati dello studio confermano punteggi medi più bassi per gli studenti frequentanti scuole situate in contesti svantaggiati e con caratteristiche sfavorevoli all'apprendimento del CIL, come il ridotto accesso a internet e il minore supporto pedagogico all'uso delle TIC. Analogamente gli studenti con basso background socioeconomico ottengono punteggi migliori quando le competenze digitali e informative vengono apprese a scuola, in un ambiente educativo dotato delle condizioni che lo consentono. Questi risultati evidenziano il ruolo delle scuole nella lotta al digital divide, in particolare nelle regioni geografiche con bacini di utenza più svantaggiata.*

1. Introduction

During the COVID-19 pandemic, digital skills became indispensable in all areas, including schools. However, the necessary shift to distance learning was not without difficulties. An interesting study conducted in Italy at that time revealed the poor preparation of teachers and families for distance learning, the lack of infrastructure and technological equipment, and the problems related to the management of learning environments and student involvement (Lucisano, 2021). Data from international surveys conducted by the OECD and the IEA in 2018 suggest that Italian students were not ready for this change. In the TALIS[1] 2018 survey, 43% of school principals in Italy reported insufficient access to the Internet, compared to the OECD average of 19%. In addition, 31% of Italian school principals reported a lack of adequate digital technologies for education, compared to the OECD average of 25%. Data from the PISA[2] 2018 survey provide additional evidence,

[1] The Teaching and Learning International Survey is an international OECD survey that examines salient aspects of the professional activity of teachers and school leaders: pedagogical orientations, teaching practices and interaction within the school with colleagues. The survey, which takes place every five years and involves almost 50 countries and economies, including Italy.

[2] The Programme for International Student Assessment is an international survey, promoted by the OECD, which measures the skills of 15-year-old students. During administration, students also answer a questionnaire, providing information about themselves and their attitude towards learning. In addition, school principals fill out a questionnaire about their schools. The survey, which takes place every three years, involves over 80 countries, including Italy.

where school principals reported that only 46% of students in Italian schools had an effective platform available for online learning support, a lower percentage than the OECD average of 54% (OECD, 2020a).

Further elements are considered in the paper "Learning remotely when schools close: How well are students and schools prepared? Insights from PISA" where the OECD concluded that "most education systems need to pay close attention to ensure that technology does not amplify existing inequalities in access and quality of learning further" (OECD, 2020b).

Thus, it is appropriate to examine the national level concerning digital skills and the risk of educational inequality. To this end, we consider Italian data from ICILS 2018, a large-scale assessment that focuses on the digital and informative skills of eighth-grade students. In the ICILS framework, CIL is defined as "an individual's ability to use computers to investigate, create, and communicate in order to participate effectively at home, at school, in the workplace, and in society" (Fraillon *et al.*, 2019). The recent acceleration of digitalization makes this competence even more crucial, and the lack of it may lead to social marginalization.

IEA ICILS 2018 reported overall unsatisfactory results, particularly in Italy compared to other participating European countries. Additionally, the significant association of CIL with students' socioeconomic background (Fraillon *et al.*, 2020) reveals significant differences even within countries, with lower performance by disadvantaged students. This study aims to define the nature and extent of the digital divide in the national context, investigating the differences among students, schools, and geographic territories. At the same time, the study aims to investigate the role of schools in counteracting the digital divide in the presence of adequate resources and support for ICT learning.

In the next section, we provide an overview of the concept and definitions of the digital divide. In section 3, we summarize the ICILS 2018 framework and key findings at the international level. In the last section, we present our study on the digital divide by analysing the Italian ICILS 2018 data.

2. The digital divide

The digital divide has many roots and landscapes, and for this reason, it has no single definition. The Internet is probably the first thing that comes to mind when we talk about the digital divide. Behind the Internet, there is a divide between those who have access to information and those who do not. The so-called information-haves go online to surf the net for information, entertainment, and other services. They are better educated and live a higher quality of life. On the contrary, the information have-nots live in the poorest regions of the world and are mostly illiterate. The digital divide mirrors traditional social stratifications in terms of Socioeconomic Status (SES), income, and education. The world is not fair, and the digital divide is an expression of this unfairness (OECD, 2000).

Hohlfeld *et al.* (2008) define the digital divide as a social inequity between individuals regarding access to ICT, frequency of technology use, and the ability to use ICT for different purposes. The concept of the digital divide was first associated with technological disparity, but the digital literacy gap expands the inequity into a social divide. The digital divide is both technical and educational. Access to ICT is the first step to overcome it, but the fundamental difference is educational: the focus is on digital literacy. Merely having a personal computer and access to the Internet is not sufficient in the presence of a lack of know-how. Illiteracy is a huge contributor to the global digital divide, and it is connected to issues of equity and quality of life. In this perspective, investments in education are crucial, and schools play an important role in ensuring equity of access to ICT and thereby act as an equalizing role (OECD, 2000).

Having a computer is not sufficient to close the digital divide. Furthermore, owning a computer with internet access is only a starting point, from which only an information-literate person can benefit. The real divide relates to literacy, and the literacy divide relates to social background. In this sense, the digital divide is much more complicated than an internet connectivity issue (Blau, 2002).

In 2020, schools underwent one of the most intense and sudden changes, moving from fully face-to-face to distance learning, and then to blended learning in the first months of the 2020/21 school year. In this new didactic mode, having a connection and a PC at one's disposal is not only a fundamental predictive factor for adequate skills development but also a requirement for access to education. In addition to the availability of IT devices, distance learning has highlighted the difficulties in the digital skills of the Italian population, which has one of the worst situations in Europe (ISTAT, 2021).

In the transition from paper-and-pencil assessments into digital environments, students' ICT capabilities become relevant as they may affect students' ability to read and write (Ercikan, Asil and Grover, 2018).

In this direction, it is worth mentioning the publication of the Joint Research Network (JRC) of the European Commission concerning the assessment of digital competences (Vuorikari, 2016).

The DigCompEdu framework is directed towards educators at all levels of education, from early childhood to higher and adult education, includ-

ing general and vocational training, special needs education, and non-formal learning contexts. The DigComp Framework has 5 dimensions:
- Competence areas identified as part of digital competence;
- Competence descriptors and titles that are pertinent to each area;
- Proficiency levels for each competence;
- Knowledge, skills, and attitudes applicable to each competence;
- Examples of use, on the applicability of the competence to different purposes.

Eight proficiency levels for each competence have been defined through learning outcomes (using action verbs, following Bloom's taxonomy) and inspired by the structure and vocabulary of the European Qualification Framework (EQF).

Digital competence is crucial, and this framework is of great importance in creating a common understanding in this regard. Additionally, it would be a useful instrument for assessment and evaluation purposes for learning and teaching in schools. It should be noted that DigComp is an exhaustive and comprehensive tool for citizens. Several countries have proposed their own frameworks derived from DigComp to develop a common framework for primary and secondary education. The European Commission funded research project CRISS has proposed a digital competence framework for students (DCFS). Specifically, the study was guided by the following two research questions: "(1) What knowledge, skills, and attitudes do primary and secondary students in Europe need to become digitally competent? (2) How can this competence be assessed?".

For wider engagement, also at the national level, it would be interesting to follow the developments of this project whose next stage will be to roll out a large-scale pilot across Europe (Guitert, 2020).

3. IEA ICILS 2018

ICILS is a computer-based assessment survey conducted by IEA to investigate how students at the eighth grade "are developing Computer and Information Literacy (CIL) to support their ability to participate in the digital age" (Fraillon *et al.*, 2013). The school-related aim of this large-scale assessment is to promote school programs adapted to the transformations of the digital age.

Since 2013, the study has been administered every 5 years; in 2018, the second cycle enrolled 12 countries, including Italy, and 2 benchmarking entities. ICILS 2018 involved 2,226 schools, for a total of 46,651 students and

26,530 teachers. The schools were sampled by a probability proportional to size method, and 20 students and 15 teachers were then randomly drawn from each sampled school.

The administration of ICILS 2018 generally took place in the latter part of the school year, and the target age of the students was 13.5 years or slightly more. In Italy, the administration took place at the beginning of the school year, so the average age of Italian students was 13.26 years, and they attended the eighth grade for a shorter period than the other students. For this reason, according to IEA, CIL results are considered not fully comparable with those of other participating countries; certainly, this issue does not interfere with an in-depth look at the results within the country.

ICILS 2018 included the optional component Computational Thinking (CT) in addition to the main component CIL, but Italy has joined only for the CIL, on which this contribution is focused. According to the ICILS 2018 framework, the CIL is divided into four areas of skills and knowledge: understanding the use of the computer, collecting information, producing information, digital communication (Fraillon *et al.*, 2019). The CIL assessment consisted of five 30-minute modules related to real-world topics. Each module included a sequence of 5-8 short tasks (that took students less than 1 minute) and was relevant to the contextual knowledge that supported work on the single large task, i.e., developing an information product (such as a presentation, poster, website, or social media post). Each student was assigned 2 modules, based on a randomized balanced design. Context Questionnaires were also administered to the students about their family background and their use of ICT, as well as to the school coordinator and teachers to obtain information on the ICT resources and CIL teaching in their school (Fraillon *et al.*, 2020).

For an overview of ICILS 2018 results, the CIL average of all participants is shown below (Fig. 1).

Denmark is the best-performing country (553), followed by Moscow (549) and Korea (542). On a European level, Finland (531), Germany (518), and Portugal (516) also achieved average scores significantly higher than the international average (496). France (499), Luxembourg (482), and Italy (461) followed. The CIL average score of Italian students is significantly lower than the international one. To delve deeper into the outcomes from the perspective of the digital divide, it is interesting to examine the percentage of students who reached the CIL minimum level. These underperforming students failed to understand and perform even the most basic ICT operations. Overall, 43% of students stopped at the lower CIL levels (Minimum level and Level 1); the EU average percentage of underperforming students is 39%, but there are significant differences among European countries (Fig. 2).

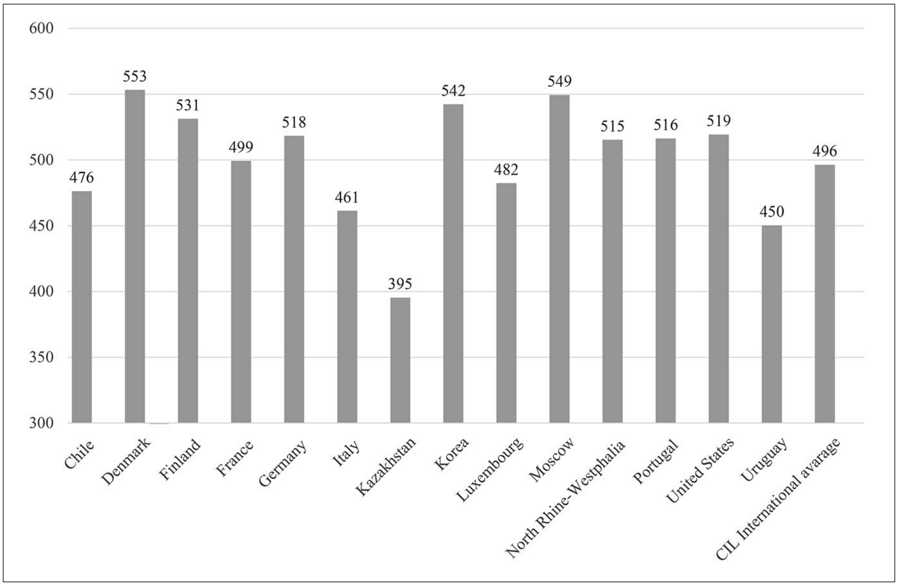

Fig 1 – ICILS 2018 average scores

Source: IEA ICILS (2018)

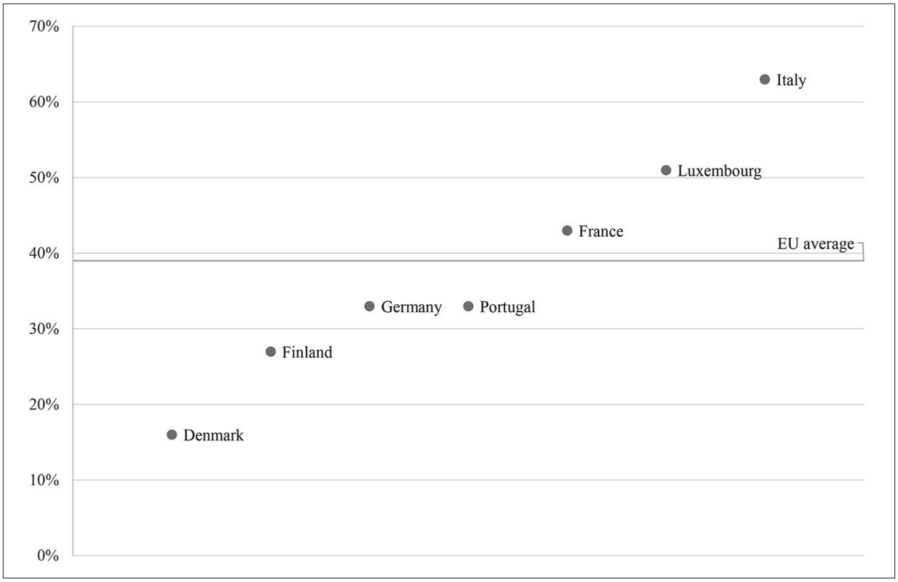

Fig. 2 – Percentages of under-performing students in EU countries

Source: IEA ICILS (2018)

These underperforming students failed to understand and perform even the most basic ICT operations. Overall, 43% of students stopped at the lower CIL levels (Minimum level and Level 1). The EU average percentage of underperforming students is 39%, but there are significant differences among European countries (Fig. 2).

In Italy, the percentage of underperforming students reached 63%, the highest among participating European countries, in line with the lowest average CIL score. For this reason, specific recommendations have been addressed to Italy by the European Commission in Education and Training Monitor 2020 (European Commission, 2020).

Considering ICILS data important and critical overall, the European Commission also focused on this issue in "The 2018 International Computer and Information Literacy Study (ICILS): Main findings and implications for education policies in Europe" (European Commission, 2019). In addition, the target "to have less than 15% underperforming students in CIL by 2030" has been introduced as an EU target for students' digital competence. This target is part of the broader "Digital Education Action Plan 2021-2027 Resetting Education and Training for the Digital Age". The plan offers a long-term strategic vision for high-quality, inclusive, and accessible European digital education; the stated objectives are:

Addresses the challenges and opportunities of the COVID-19 pandemic, which has led to the unprecedented use of technology for education and training purposes.

Seeks stronger cooperation at the EU level on digital education and underscores the importance of working together across sectors to bring education into the digital age.

Presents opportunities, including improved quality and quantity of teaching concerning digital technologies, support for the digitalization of teaching methods and pedagogies, and the provision of infrastructure required for inclusive and resilient remote learning.

This new Digital Action Plan is an operational response to the ICILS results that debunked the myth of digital natives. In fact, ICILS 2018 revealed that students do not have sophisticated digital skills if they are not taught. From the analyses of international data, the IEA determined that the educational systems with the best CIL scores have a high percentage of students in schools where they frequently learn and practice ICT tasks. These schools have more professional ICT learning support, in addition to infrastructure and learning materials related to this area. Therefore, providing schools with computer equipment is necessary but not enough to improve CIL. The two ICILS 2018 highlights regarding the digital divide are:

– A gap in CIL associated with socioeconomic backgrounds of students;
– A greater variability in CIL scores within countries than between countries.

These findings are investigated at the national level in the study described in the following section.

4. The digital divide study

The object of the present study was the relationship between Italian students' CIL achievement and socioeconomic background of students and schools. To better characterize the nature and extent of digital divide, both schools' and students' characteristics in the digital domain have been considered i.e., the availability of devices and the internet access at home and school. The hypothesis explored was that schools can counteract the digital divide, if provided with adequate resources and effective teaching practices for learning of digital and informative skills.

This study refers to Italian data, collected by INVALSI[3] as part of the IEA ICILS 2018 survey. The ICILS 2018 Italian sample consisted of 2,810 students attending the eighth grade in 146 schools distributed on the national territory as shown in the table 1.

Tab. 1 – Number of schools by Macro-area[4]

	Frequency	*Percentage*
North	64	43,8
Centre	28	19,2
South	54	37,0
Tot.	146	100,0

In detail, the data analysed come from:
– CIL test for students;
– Student questionnaire, including questions about the socioeconomic background, the use of information technology and the attitude to computer use;

[3] National Institute for the Evaluation of the Educational System of Education and Training.

[4] The Macro-areas are composed of the following Italian regions:
– Nord: Lombardia, Veneto, Emilia-Romagna, Piemonte, Liguria, Friuli-Venezia Giulia, Trentino-Alto Adige, Valle d'Aosta;
– Center: Marche, Lazio, Toscana and Umbria;
– South: Abruzzo, Basilicata, Calabria, Campania, Molise, Puglia, Sicilia and Sardegna.

– Coordinator (in Italy, digital animator) questionnaire evaluating the ICT infrastructure and the teaching assistance available in school.

CIL achievement has been analysed in relation to students and school SES as well as by territorial area, classified in 3 Macro-areas: North, Centre and South of Italy.

The data have been processed by the INVALSI statistical service with the integration of the school ESCS (Economic, Social and Cultural Status) index, derived from the INVALSI national standardized assessment. The school ESCS index is the average of all the students' ESCS attending the same school and it is therefore an indicator of the background of the school's catchment area. The ESCS index was constructed through the principal component analysis based on three indicators: the employment status of the parents – HISEI –, their level of education – PARED – and their possess of cultural resources and favourable conditions to learning at home – HOMEPOS – (Campodifiori *et al.*, 2010).

The S_NISB, indicating student SES in ICILS, similarly has been derived through principal component analysis considering highest occupational status of parents, highest educational level of parents, and home literacy resources (Fraillon *et al.*, 2019). The reason for including the student S_NISB and school ESCS is twofold. Firstly, the SES at the school level was not provided in the ICILS dataset. Secondly, ESCS is calculated on data collected from a representative sample of students and schools in Italy, and it considers all students and schools in the country. This makes it a reliable and useful proxy for school-level SES in our analyses. Both indices are calculated based on similar variables and following the same methodological indications widely used in the international context. The school ESCS and the student S_NISB were both categorized using quartiles into: Low, Medium-Low, Medium-High, and High. The ESCS categorization into quartiles is commonly used in Italian National Reports (INVALSI, 2018; INVALSI, 2019; INVALSI, 2022). Quartiles are frequently used in research on socioeconomic status because they simplify the interpretation of complex SES data and allow for comparability between groups of students (Ward, 2008). The quartile division of the scores of a continuous index measuring SES is also used in OECD studies[5] to describe results in terms of broader categories. The ESCS index is even divided in only three groups[6]: High – Medium – Low and, a such categorization is used in the ICILS 2018 International Report for the Parental Occupation scale (SEI)[7].

[5] *Italia: Education at a Glance 2021*: OECD Indicators, OECD iLibrary (oecd-ilibrary.org).

[6] https://www.invalsiopen.it/indicatore-escs-valutazione-equa/

[7] The SEI scale is continuous and ranges from 16 to 90 score points but it is divided into three groups "low occupational status", "medium occupational status" and "high occupational

In this study, the characteristics and practice of schools in the digital area have also been examined. In ICILS 2018 a scale has been derived from the items of the student questionnaire about learning different tasks[8] of ICT at school, higher scores on this scale indicate greater attribution of CIL learning in school environment (Fraillon *et al.*, 2020). We categorized the ICT learning index into 4 categories using quartiles (Low, Medium Low, Medium High, High) to observe the differences in CIL performance among the four groups of students with different ICT learning.

4.1. Analyses and results

This descriptive study aims to provide an overview of risk conditions and factors that could impact the digital divide at the national level. All study analyses have been conducted in R, using R Analyzer for Large-Scale Assessments (Mirazchiyski, 2021). RALSA is an R package specifically designed for analysing data from studies using complex sampling and assessment designs, such as the ICILS study. RALSA allows the conversion of data from SPSS file format into an R dataset, and the output is written into an MS Excel workbook.

Confirming that the variability in CIL scores may be greater within countries than between countries, Italy has high differences by Macro-area in CIL scores and the percentage of underperforming students (Figs. 3-4).

Comparing the South (435) with the North (478) and the Centre (471) of Italy, the difference in the average CIL score is significant while the difference between the North and the Centre is negligible (Fig. 3).

Similarly, the percentage of underperforming students (Fig. 4) is higher in the South (74%) than in the North (55%) and in the Centre (60%) of Italy.

status”. Depending on the category to which the students are referred, the results are observed and commented.

[8] Examples of ICT tasks at school are:
- Provide references to internet sources;
- Search for information using ICT;
- Present information for a given audience or purpose using ICT;
- Work out whether to trust information from the internet;
- Decide what information obtained from the internet is relevant to include in schoolwork;
- Organize information obtained from internet sources;
- Decide where to look for information on the internet about an unfamiliar topic;
- Use ICT to collaborate with others.

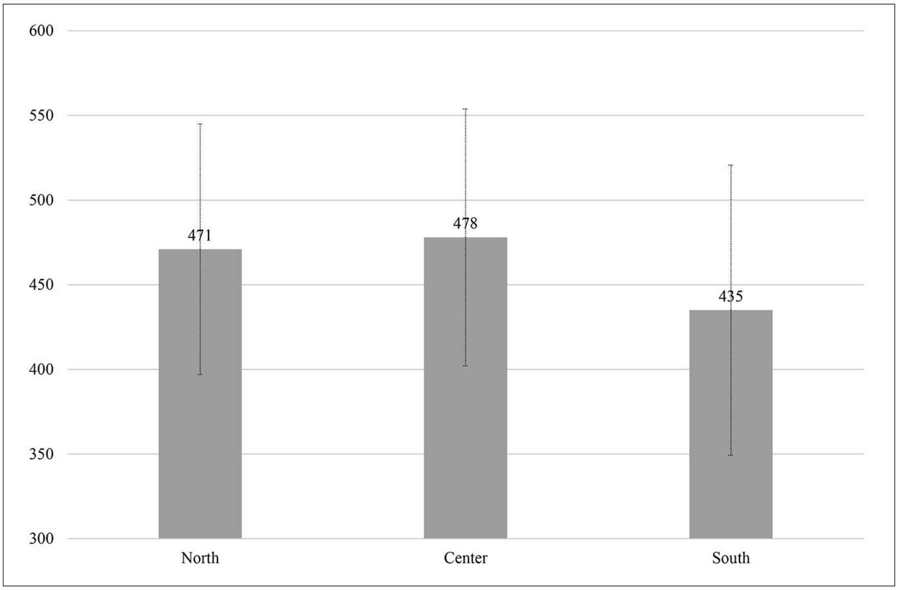

Fig. 3 – CIL average scores by Macro-area

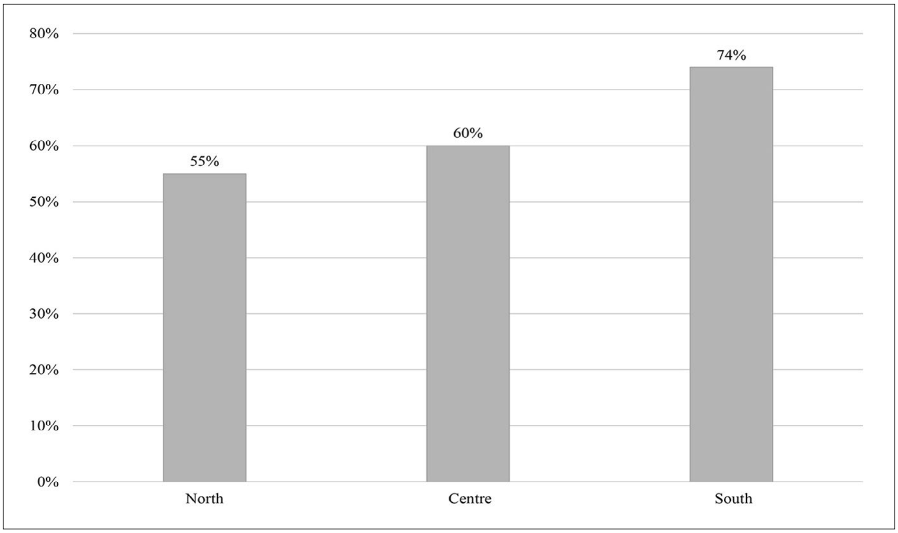

Fig. 4 – Percentage of underperforming students in CIL by Macro-area

As shown by international data, Socioeconomic Status (SES) has a strong influence on digital literacy, with students from disadvantaged backgrounds having significantly lower scores than those with high SES. The analysis of the Italian sample confirms a positive association (r =.32, $p < .05$) between students' CIL scores and their SES, as measured by S_NISB. The graph below (Fig. 5) shows the average CIL scores by student SES.

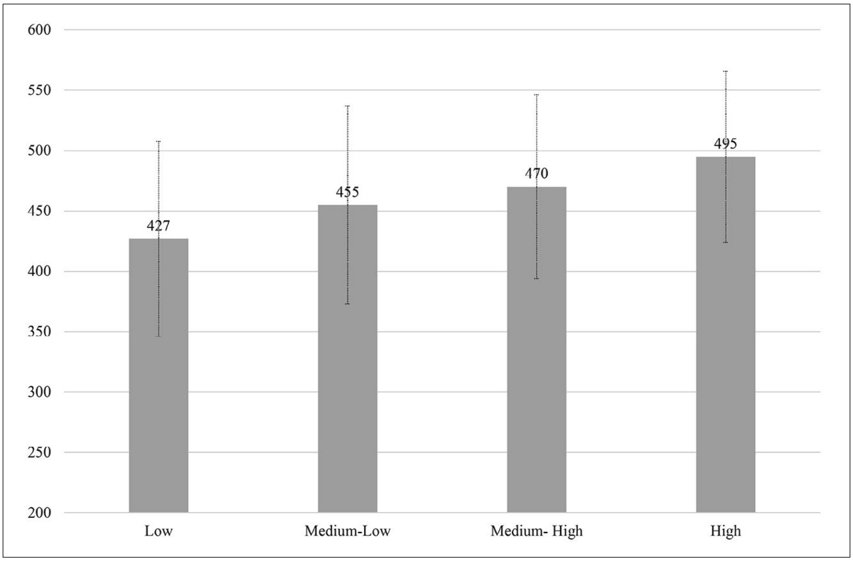

Fig. 5 – CIL average score by student socioeconomic background

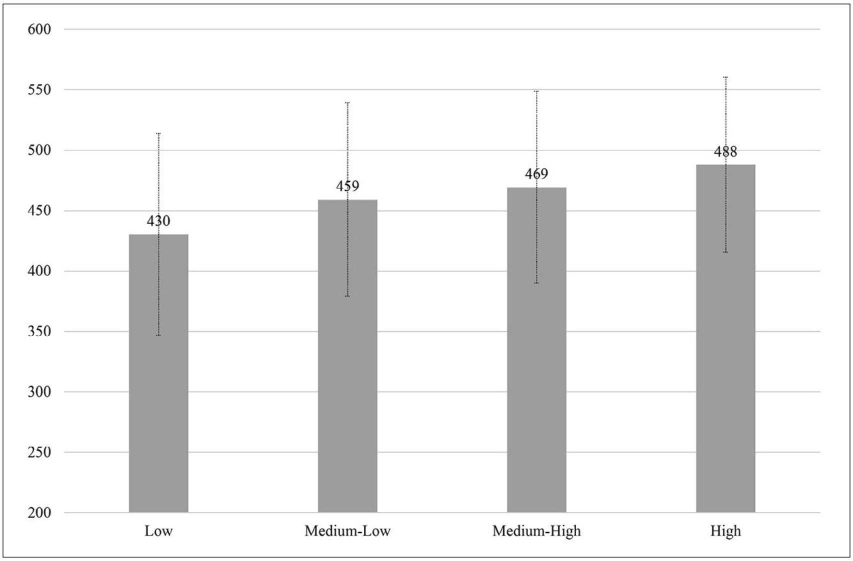

Fig. 6 – CIL average score by school ESCS

As we can see, there is a difference of 68 points between the average CIL score of low SES students and that of high SES students. The effect of SES is certainly detectable in schools with students from homogeneous background. The following graph (Fig. 6) compares the average CIL score by school ESCS.

A Low background school is a school whose average socioeconomic level of students is in the lower quartile of the index of economic, social, and cultural status of the schools in the sample. There is a difference of 58 points in CIL average between Italian schools with Low ESCS and those with High ESCS. Therefore, it is worth to check the percentage of "disadvantaged" students by Macro-area (Fig. 7).

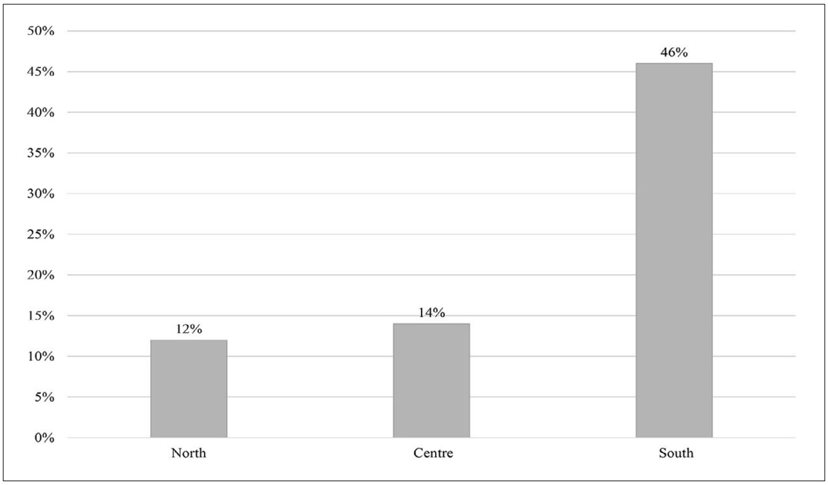

Fig. 7 – Percentage of students in schools with low ESCS by Macro-area

In the South of Italy 46% of schools are with Low ESCS, while in the North are 12%, and in the Centre are 14%; despite the inhomogeneous distribution of schools with low ESCS within the national territory, it could be interesting to see the average CIL score by different level of ESCS and by Macro-area (Tab. 2).

Tab. 2 – CIL average score by school ESCS and Macro-area

	Low	*Medium-Low*	*Medium-High*	*High*
North	460,3 (7,1)	470,3 (6,0)	480,6 (7,0)	494,8 (8,5)
Centre	469,8 (6,7)	441,6 (12,4)	471,0 (8,0)	486,1 (8,2)
South	413,9 (8,5)	442,9 (7,5)	447,1 (11,2)	475,6 (13,2)

The differences in CIL by SES appear less severe in the North and in the Centre Macro-areas; the CIL average score by minimum and maximum levels of school ESCS were highlighted in Figure 8.

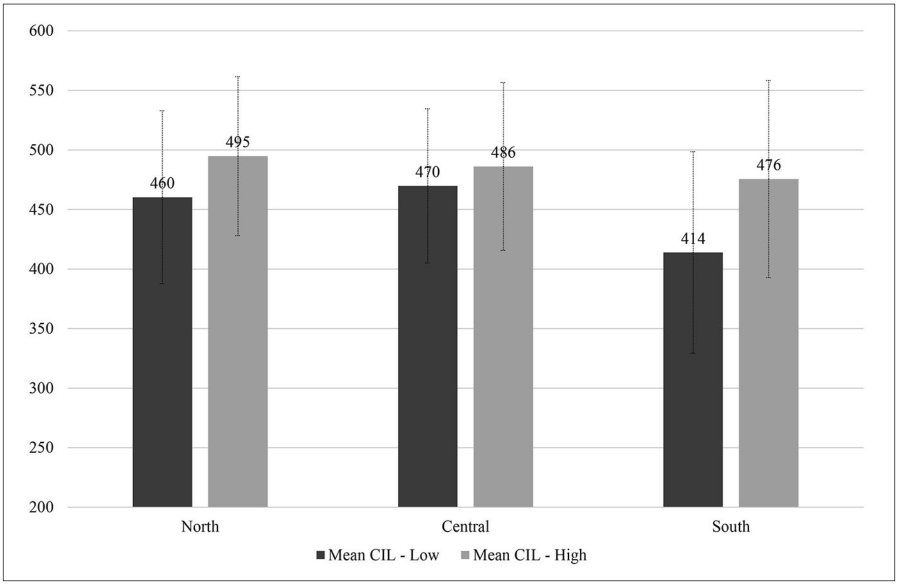

Fig. 8 – CIL average score by school ESCS and Macro-area

The differences in CIL average score between the South and the other Macro-areas are alarming for low SES students, on the contrary there are only 10 points of difference between the South and the Centre among high SES students, a similar difference to that between the North and the Centre of Italy. In addition to the differences in SES that give a first reading key for the lower CIL average score in the South, we also checked for other differences among the three Macro-areas. Regarding the devices for student use, the major differences between the Macro-areas concern the laptops and notebooks available for student use by school (Fig. 9).

Schools in the North and, even more so, in the Centre of Italy, report a higher average number of laptops for student use. The possibility of using laptops in the classroom and taking them home makes them very useful for student learning. Schools in the South of Italy mainly report desktop computers, which may date back to the realization of laboratories under the (NOP) National Operational Program ERDF (European Regional Development Fund) "Learning Environments" 2007-2013. According to the ICILS 2018 international results, the use of ICT is hindered by the limited availability of computers with internet connection at school. Below (Fig. 10), we show the percentage of devices that are not connected to the internet in Italian schools by Macro-area.

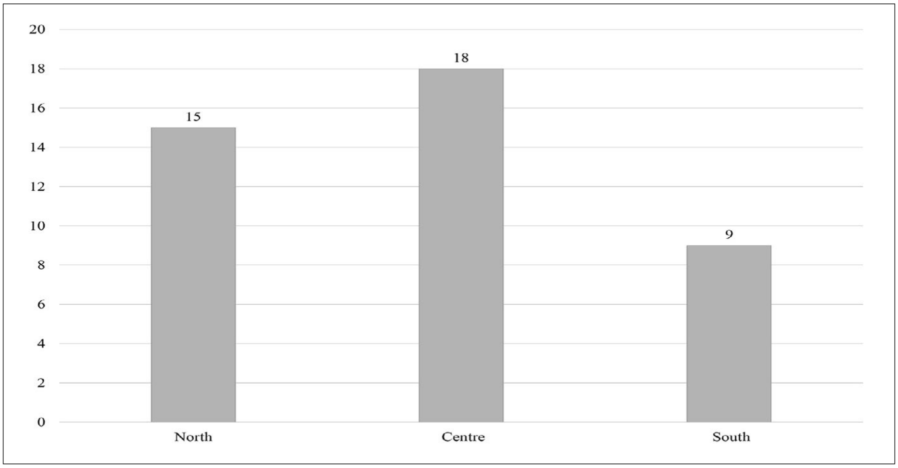

Fig. 9 – Average number of laptops/notebooks available for student use by school

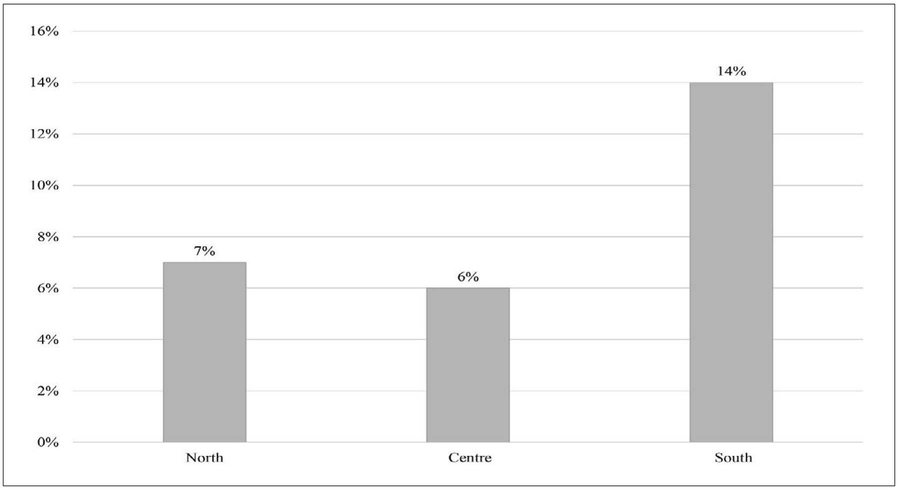

Fig. 10 – Percentage of not connected devices at school by Macro-area

The percentage of devices not connected to the Internet is, on average, twice as high in schools in the South (14%) as in schools in the North (7%) and more than twice as high in schools in the Centre (6%). In 2021 the MIUR has launched a call for the implementation of local networks within the educational institutions of all Italian regions, within the PON 2014-2020 actions.

To further investigate differences by Macro-area, home computer availability (Fig. 11) and internet connection at home (Fig. 12) were also analysed. In both cases, we found lower availability for students in the South Macro-area.

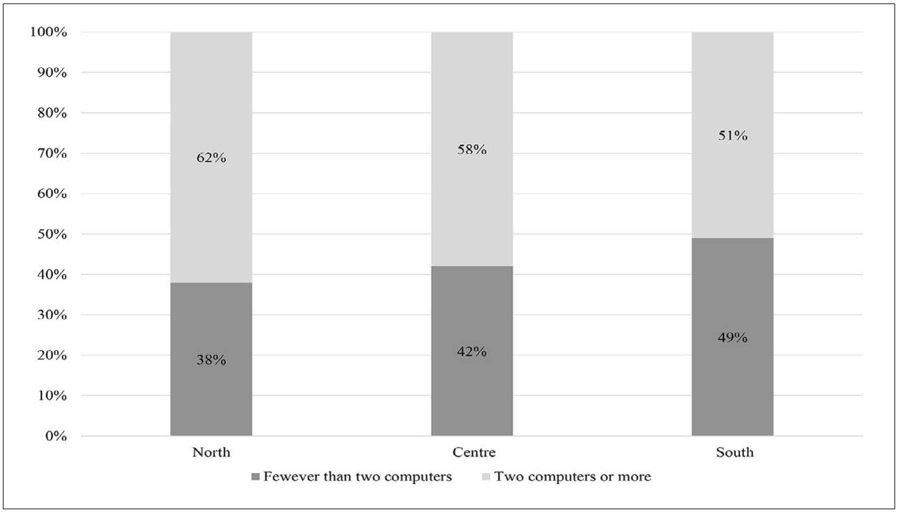

Fig. 11 – Percentage of students by computer availability at home and Macro-area

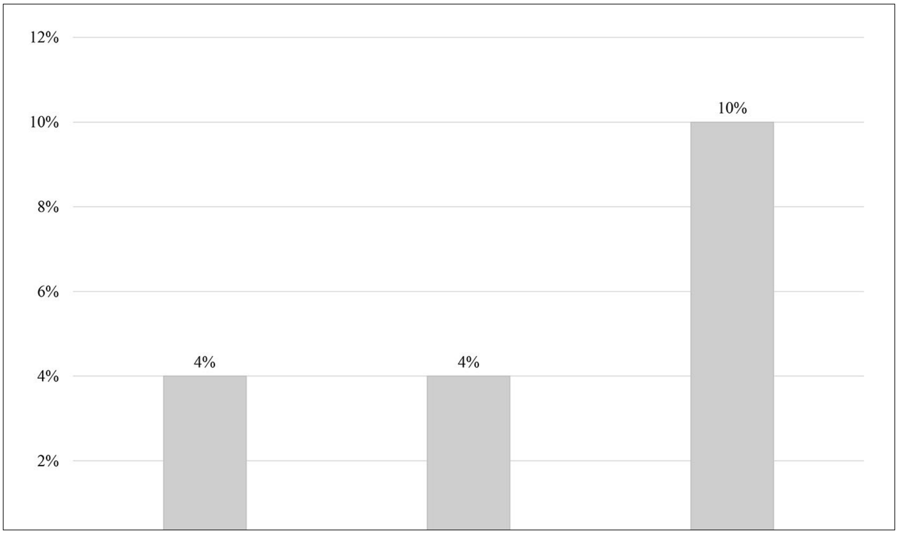

Fig. 12 – Percentage of students without internet access at home by Macro-area

The percentage of students with fewer than two computers at home is higher in the South (49%) than in the North (38%) and in the Centre (42%) of Italy. Moreover, the greater lack of Internet access in the Southern schools is amplified by the greater lack of Internet access at home (10%), more than twice as much as in the other Macro-areas (4%).

Italian ICILS 2018 data confirm that students' CIL results are positively correlated with home availability of 2 or more computers; the Spearman correlations are all significant at $p < .001$, although they differ considerably by Macro-area (North: rho = .11; Centre: rho = .17; South: rho = .22). The figure below (Fig. 13) show data compared by computer availability at home in the three Macro-areas.

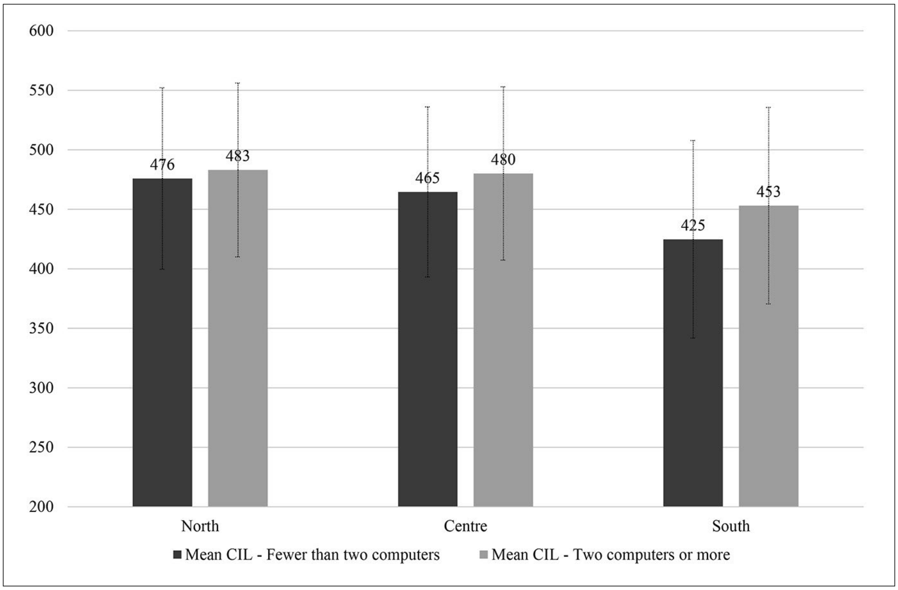

Fig. 13 – CIL average by computer availability at home and Macro-area

In the South of Italy, the difference in average CIL score between students who have at least 2 computers at home and those who have fewer (or none) is deeper than in the other Macro-areas; these data will be commented on along with those on students' average CIL score by internet access at home (Fig. 14).

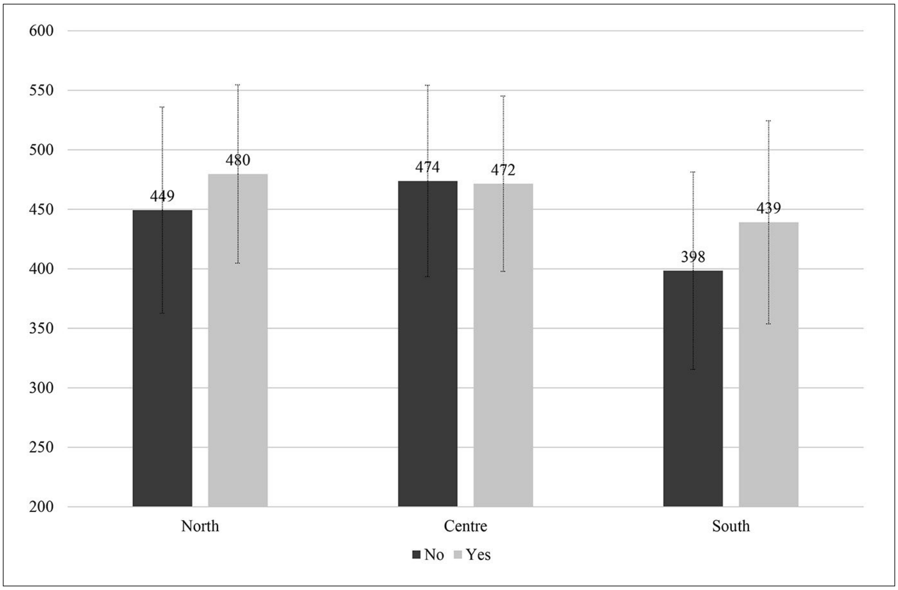

Fig. 14 – Students' CIL average score by internet access at home and Macro-area

The difference between the results of students who have access to the Internet at home and those who do not is greatest (41 points) in the South (rho = .14; p < .001) and absent in the Centre (rho = – 0.01; p = .59). These data are likely related to the fewer specific resources for ICT learning at school detected in the South. Another relevant variable for learning ICT tasks is the pedagogical support in the classroom, which is also not homogeneous among Macro-areas (Tab. 3).

Tab. 3 – Percentage of schools with different level of support for use of ICT by Macro-area

	Nota at all	*Very little*	*To some extent*	*A lot*
North	2,6	37,9	48,2	11,3
Centre	6,9	37	19,3	36,9
South	20,6	42,1	30,8	6,5

The minimum and maximum levels of support for the use of ICT in school by Macro-area were highlighted below (Fig. 15).

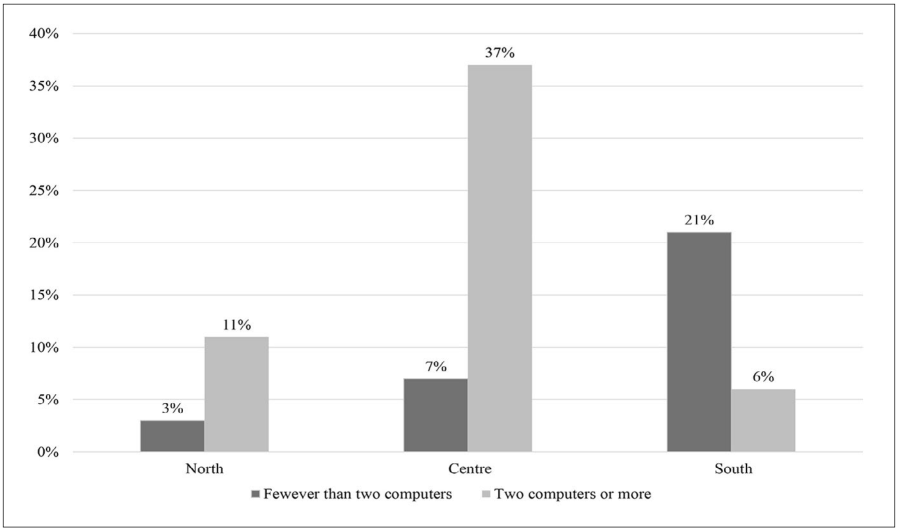

Fig. 15 – Percentage of schools with extreme level of support for use of ICT by Macro-area

The complete absence of support is found in 21% of Southern schools, 3% of Northern schools, and 7% of Central schools. Conversely, a lot of pedagogical support has been found only in 6% of Southern schools, 11% of Northern schools, and 37% of Central schools.

According to international ICILS results, students had better CIL outcomes when they have learned about ICT at school to a greater extent; the positive association appeared stronger in the countries with highest CIL average (Denmark, Korea, Finland). Italian students CIL scores were significantly but smaller correlated with ICT learning index ($r = .15$; $p < .05$). Below (Fig. 16) the average CIL scores of the students by level of the ICT learning index are shown.

We can observe a higher average score in CIL as ICT learning Index increases, with an average difference of 25 points in CIL between Low and High level of ICT learning at the national level; this trend is found in all Macro-areas (Tab. 4).

Therefore, in all Macro-areas, students achieve better if they learn and frequently practice ICT tasks at school. The analysis on the Italian sample confirms both the association between students' CIL scores with ICT learning at school ($r = .15$; $p < .05$) and socioeconomic background ($r = .32$, $p < .05$), although the latter is higher. In the South Macro-area, where the performance in CIL is lower, there are 40% of students with low SES; in addition, the Southern students is more affected by the lack of ICT resources and infrastructures both at school and at home. In summary, higher performance and shorter difference in CIL by SES emerge in the North and Centre of Italy, where the percentage of disadvantaged students is smaller, while adequate resources and ICT pedagogical support at school are greater; these results seem to support the importance of school in fighting the digital divide.

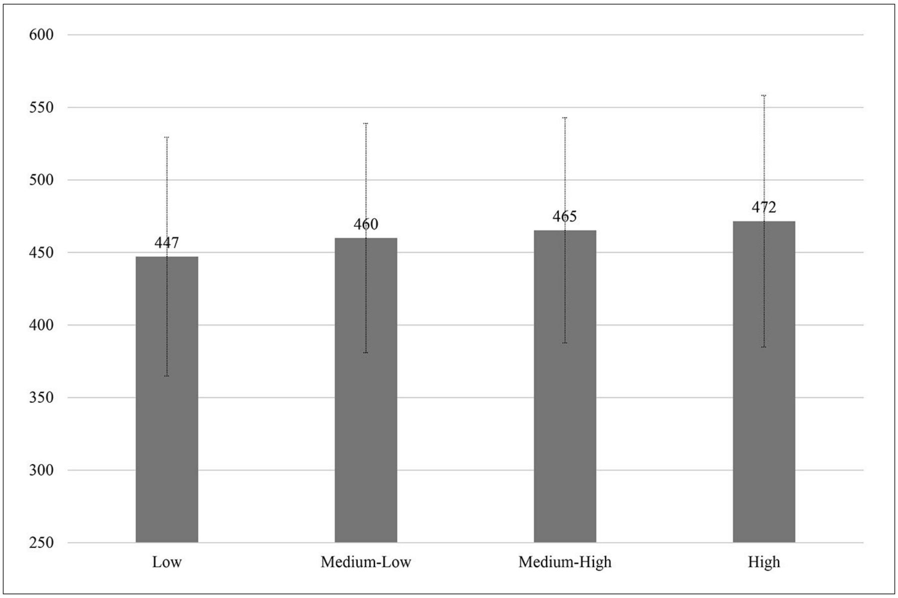

Fig 16 – Students' CIL average score by ICT learning at school

Tab. 4 – CIL score (standard deviation) by ICT Learning level and Macro-area

ICT Learning level	*North*	*Centre*	*South*
Low	459.2 (5.5)	464.1 (9.2)	422.4 (7.2)
Medium-Low	479.4 (5.2)	468.4 (6.4)	432.1 (5.9)
Medium- High	480.9 (4.8)	472.2 (9.3)	441.8 (7.9)
High	492.3 (6.4)	479.6 (9.7)	442.2 (8.8)

5. Conclusions

As highlighted at the international level, CIL scores are associated with the socioeconomic background of students. The critical results of Italian students in ICILS 2018 demonstrate a deep digital divide within the country, particularly in the South Macro-area, whose schools enroll the most students with low SES. Characteristics in the digital domain are also relevant, such as the number of devices and internet access at home, as well as at school, the number and type of computers available for student use, and pedagogical support for ICT tasks. The lower socioeconomic background of students in South Italy is paired with a greater lack of adequate resources at home and at school. The reduced access to the Internet at home is further amplified by the same condition at school; the lower availability of computers at home is coupled with fewer laptops and notebooks at school for student use, which are types of computers that could also be taken home. Differences between Macro-areas are also identified in the availability of pedagogical support for ICT use at school. In the South of Italy, the situation is aggravated by the poverty of the surrounding area, resulting in fewer extracurricular educational offerings (SVIMEZ, 2018). The 2018 SVIMEZ[9] Report highlights how the quality of learning cannot disregard the socioeconomic context. Unemployment, widespread poverty, social exclusion, lack of public services and cultural institutions synergistically affect education. The Adequate resources appear to be more lacking in schools with "disadvantaged" catchment areas, just where it would be necessary to fight the digital divide with respect to educational equity. However, students in each Macro-area perform better in CIL when digital and informative skills are learned at school; in addition, students with low backgrounds perform better in Macro-areas where schools are better equipped with adequate resources. Therefore, the hypothesis that schools can play a key role in developing CIL and countering the digital divide is supported in the presence of favourable conditions for learning and practicing ICT tasks. The integration of CIL in teaching should be implemented through specific actions. It would be urgent to strengthen the educational offer in "disadvantaged" schools to counteract unfavourable conditions for learning.

Make grammar and syntax corrections: These initial results offer interesting insights into the nature and extent of the digital divide. However, the results and hypotheses generated by this pilot study should be subject to analytical studies for generalization. It would be relevant for future studies to consider a finer geographic grid. Additionally, a multilevel model could be considered to investigate the combined effect of schools and students on CIL scores. Additional data could be used to investigate the effect of other school characteristics, such as school size and availability of adequate space.

[9] Association for the Development of Industry in Southern Italy is a private, non-profit association whose corporate purpose is the study of the economic conditions of Southern Italy.

It would also be possible to consider teacher characteristics. For example, in the South of Italy, teachers are more likely to be precarious for longer and have a higher average age than in other Macro-areas.

Finally, it will be important to observe the data from the next cycle, ICILS 2023, which will investigate CIL outcomes in the post-pandemic era. What has changed and in which direction? Certainly, "the pandemic experience has catapulted education systems, which traditionally lag behind in innovation, years ahead in what would have been a slow move towards smart schooling" (OECD, 2021b). Likely the pandemic has given the Italian school the necessary pressure to modernise and equip itself with better technological infrastructure. In a study on familiarity with technology at school, > 90% of school staff reported improvement after the pandemic period (Promethean, 2022). We hope that students' digital and informational competences will also be significantly improved. Only by reducing the digital divide can we promote active citizenship and social inclusion.

References

Blau A. (2002), "Access isn't enough: Merely connecting people and computers won't close the digital divide", *American Libraries*, 33 (6), pp. 50-52.

Campodifiori E., Figura E., Papini M., Ricci R. (2010), *Un indicatore di status socioeconomico-culturale degli allievi della quinta primaria in Italia*, Working Paper INVALSI, n. 2, retrieved on May 9, 2023, from https://www.invalsi.it/download/wp/wp02_Ricci.pdf.

Directorate-General for Education, Youth, Sport and Culture – European Commission (2019), *The 2018 International Computer and Information Literacy Study (ICILS): Main findings and implications for education policies in Europe*, Publications Office of the European Union, Luxembourg, retrieved on January 30, 2023, from https://education.ec.europa.eu/sites/default/files/document-library-docs/icils-2018-policy-note.pdf.

Directorate General for Education, Youth, Sport and Culture – European Commission (2020), *Education and Training monitor 2020 – Italy*, Publications Office of the European Union, Luxembourg, retrieved on January 30, 2023, from https://op.europa.eu/webpub/eac/education-and-training-monitor-2020/countries/italy.html.

Ercikan K., Asil M., Grover R. (2018), "Digital divide: A critical context for digitally based assessments", *Education Policy Analysis Archives*, 26 (51), pp. 1-20.

Fraillon J., Schulz W., Ainley J. (2013), *International Computer and Information Literacy Study: Assessment Framework*, IEA, Amsterdam, retrieved on January 30, 2023, from https://www.iea.nl/publications/assessment-framework/international-computer-and-information-literacy-study-2013.

Fraillon J., Ainley J., Schulz W., Duckworth D., Friedman T. (2019a), *IEA International Computer and Information Literacy Study 2018 Assessment Framework*, IEA, Amsterdam, retrieved on January 30, 2023, from https://doi.org/10.1007/978-3-030-19389-8.

Fraillon J., Ainley J., Schulz W., Friedman T., Duckworth D. (2019b), *Preparing for Life in a Digital World. IEA International Computer and Information Literacy Study 2018 International Report*, IEA, Amsterdam, retrieved on January 30, 2023, from https://www.iea.nl/publications/study-reports/international-reports-iea-studies/preparing-life-digital-world.

Fraillon J., Ainley J., Schulz W., Friedman T., Duckworth D. (2020), *International Computer and Information Literacy Study 2018 Technical Report*, IEA, Amsterdam, retrieved etrieved on January 30, 2023, from https://www.iea.nl/sites/default/files/2020-05/ICILS%202018%20Technical%20Report-FINAL_0.pdf.

Guitert M., Romeu T., Baztán P. (2021), "The digital competence framework for primary and secondary schools", *European Journal of Education*, 56, pp. 133-149.

Hohlfeld T. N., Ritzhaupt A. D., Barron A. E., Kemker K. (2008), "Examining the digital divide in K-12 public schools: Four-year trends for supporting ICT literacy in Florida", *Computers & Education*, 51 (4), pp.1648-1663.

INVALSI (2018), *Rapporto INVALSI 2018*, retrieved on May 9, 2023, from https://INVALSI-areaprove.cineca.it/docs/2018/Rapporto_prove_INVALSI_2018.pdf.

INVALSI (2019), *Rapporto INVALSI 2019*, retrieved on May 9, 2023, from https://INVALSI-areaprove.cineca.it/docs/2019/Rapporto_prove_INVALSI_2019.pdf.

INVALSI (2020), *L'indicatore ESCS per una valutazione più equa*, retrieved on May 9, 2023, from https://www.invalsiopen.it/indicatore-escs-valutazione-equa/.

INVALSI (2022), *Rapporto INVALSI 2022*, retrieved on May 9, 2023, from https://invalsi-areaprove.cineca.it/docs/2022/Rilevazioni_Nazionali/Rapporto/Rapporto_Prove_INVALSI_2022.pdf.

ISTAT (2021), *Rapporto annuale: la situazione del paese nel 2021*, retrieved on January 30, 2023, from https://www.istat.it/storage/rapporto-annuale/2021/Rapporto_Annuale_2021.pdf.

Lucisano P. (2021), *La DaD in emergenza: vissuti e valutazioni degli insegnanti italiani. Scelte metodologiche e primi risultati nazionali*, retrieved on May 9, 2023, from 04-sird-studi-e-ricerche-dad.pdf (pensamultimedia.it).

Mirazchiyski P.V. (2021), "RALSA: the R analyzer for large-scale assessments", *Large-scale Assessments in Education*, 9 (1), pp. 1-24.

OECD (2000), *Learning to Bridge the Digital Divide, Schooling for Tomorrow*, OECD Publishing, Paris.

OECD (2019), *TALIS 2018 Results*, Vol. I: *Teachers and School Leaders as Lifelong Learners*, TALIS, OECD Publishing, Paris, retrieved on January 30, 2023, from https://doi.org/10.1787/1d0bc92a-en.

OECD (2020a), *PISA 2018 Results*, Vol. V: *Effective Policies, Successful Schools*, PISA, OECD Publishing, Paris, retrieved on May 9, 2023, from https://doi.org/10.1787/ca768d40-en.

OECD (2020b), *Learning remotely when schools close*, retrieved on January 30, 2023, from https://www.oecd.org/coronavirus/policy-responses/learning-remotely-when-schools-close-how-well-are-students-and-schools-prepared-insights-from-pisa-3bfda1f7/.
OECD (2021a), *Garantire pari opportunità per gli studenti provenienti da diversi contesti socioeconomici,* Italia | Education at a Glance 2021: OECD Indicators | OECD iLibrary (oecd-ilibrary.org).
OECD (2021b), *OECD Digital Education Outlook 2021: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots*, OECD Publishing, Paris, retrieved on May 9, 2023 from https://doi.org/10.1787/589b283f-en.
Promethean (2022), *Il report sul rapporto tra tecnologia e scuola in Italia*, retrieved on May 9, 2023, from https://www.prometheanworld.com/it/microsites/report-sul-rapporto-tra-tecnologia-e-scuola-in-italia/.
SVIMEZ (2018), *Rapporto SVIMEZ 2018: L'economia e la società del Mezzogiorno*, il Mulino, Bologna.
Vuorikari R., Punie Y., Carretero S., Brande L.V., European Commission (2016), *DigComp 2.0: The digital competence framework for citizens*, Publications Office, Luxembourg.
Ward M. M. (2008), "Socioeconomic status and the incidence of ESRD", *American Journal of Kidney Diseases*, 51 (4), pp. 563-572.

7. Do national and international surveys speak the same language?

by Paola Giangiacomo, Patrizia Falzetti, Cecilia Bagnarol

The National Institute for the evaluation of the education and training (INVALSI) has been dealing with national and international surveys for years. These surveys have different focus and characteristics but also have points in common such as the assessment of students' basic skills.

The purposes of the two surveys are different, at an international level an attempt is made to collect information to provide a measure of the differences existing between the participating countries, while at the national level the state of the national school system is investigated to provide self-assessment tools to schools to improve the performance of its students and consequently of the entire Institute.

The research question of this paper is based on the comparability of the two surveys, while considering the different purposes of each of them.

The aim of this work is therefore to compare the results in Mathematics obtained by Italian students in the second year of secondary school who participated in PISA 2018 with those obtained by the same students in the INVALSI national tests in the same year (OECD, 2018; INVALSI, 2018). In the first part of the research, the reference frameworks of PISA survey and INVALSI surveys are analyzed to verify the differences in the setting and implementation of the two surveys. Subsequently, attention is focused on the analysis of differences in student performance through correlational methods using the average scores by gender and geographical macro-area. Both investigations took place in the spring of 2018. The analyses were conducted both at the individual student level, considering those who carried out both tests and at the territorial level, in particular detail to the geographical macro-area. A further study involved the evaluation of the score in Mathematics obtained by students in the INVALSI 2016 test at grade 8 to verify whether, even in terms of previous INVALSI score, there is a relationship in terms of prediction of the outcome achieved by the student in PISA 2018. Our study shows a close link between the performance of students at the end of lower secondary school and the results they obtain in the second year of upper secondary school. Another element to underline is the fact that the two surveys give us a coherent and correlated picture, PISA and INVALSI present, in fact, very similar results, when we consider only the students in the upper secondary grade.

L'Istituto nazionale per la valutazione del sistema educativo di istruzione e formazione (INVALSI) si occupa da anni delle Rilevazioni a carattere Nazionale e Internazionale. Tali Rilevazioni hanno focus e caratteristiche diverse ma hanno anche dei punti in comun come la valutazione delle competenze di base degli studenti. Le finalità delle due Rilevazioni sono differenti, a livello internazionale si cerca infatti di raccogliere informazioni per fornire una misura delle differenze esistenti tra i Paesi partecipanti, a livello nazionale invece si indaga lo stato del sistema scolastico nazionale al fine di fornire strumenti di autovalutazione alle scuole per migliorare i rendimenti dei propri studenti e di conseguenza dell'intero Istituto.

La domanda di ricerca del presente contributo si basa sulla confrontabilità delle diverse indagini, pur considerando le peculiarità e le diverse finalità di ognuna di esse. Obiettivo del presente lavoro è quindi quello di comparare i risultati in Matematica ottenuti dagli studenti italiani al secondo anno della scuola secondaria di II grado che hanno partecipato all'indagine PISA 2018 con quelli ottenuti dagli stessi studenti alle prove nazionali INVALSI nello stesso anno (OECD, 2018; INVALSI, 2018). In una prima parte della ricerca si analizzano i quadri di riferimento dell'indagine PISA e delle Rilevazioni INVALSI al fine di verificare le differenze nell'impostazione e nell'implementazione delle due indagini. Successivamente si focalizza l'attenzione sull'analisi delle differenze nei rendimenti degli studenti attraverso metodi correlazionali utilizzando i punteggi medi per genere e macroarea geografica. Entrambe le indagini si sono svolte nella primavera del 2018.

Le analisi sono state condotte sia a livello di singolo studente, considerando quelli che hanno svolto entrambe le prove, sia a livello territoriale con dettaglio di macroarea geografica. Un ulteriore studio ha riguardato l'introduzione del punteggio in Matematica ottenuto dagli studenti nella Prova INVALSI 2016 al grado 8 al fine di verificare se, anche in termini di punteggio pregresso INVALSI, esista una relazione in termini di predittività dell'esito raggiunto dallo studente in PISA 2018. In particolare, dal nostro studio emerge uno stretto legame tra l'andamento degli studenti alla fine della scuola secondaria di primo grado e i risultati che gli stessi ottengono al *secondo anno della scuola secondaria di secondo grado. Altro elemento da sottolineare è il fatto che le due Rilevazioni ci restituiscono un quadro coerente e correlato, PISA e le Rilevazioni nazionali presentano, infatti, risultati molto simili, quando consideriamo nelle analisi i soli studenti in seconda superiore.*

1. Introduction

In recent decades, the experiences of international surveys on school learning, such as the OECD-PISA, have allowed the different participating countries to have a more accurate picture of the characteristics and effectiveness of their education systems.

Especially in Italy, there has been an effect of socioeconomic and cultural variables on the performance of students, which tends to increase with the transition to higher school grades, up to determine a segregation effect that sees students with the same background attend the same type of school.

Many studies, in fact, have amply demonstrated how the Italian education system is characterized by a low variance within schools and a high variance between schools (OECD, 2018).

The results of international comparative surveys also helped to clarify the origin of these differences: while the variance within schools is likely to be due to individual differences and individual student engagement, the variance between schools seems to be to a large extent the result of the above-mentioned compositional processes.

A challenge for the Italian education system is, therefore, to try to reduce the differences in performance existing between schools that group socially disadvantaged students and schools whose students come from better socioeconomic and cultural conditions.

Although many countries have identified the reduction of the inequality of achievement between students from different social groups or ethnic groups as a major task of the education system, this goal is far from being achieved.

In fact, the results of the most important international surveys have repeatedly highlighted the persistence of a strong association between academic performance and socioeconomic and cultural background of the students' family of origin (OECD, 2016; Mullis *et al.*, 2016; Mullis *et al.*, 2012; Chiu and Xihua, 2008; Ercikan *et al.*, 2005; Sirin, 2005).

Furthermore, the results of Italian students, associated to the socioeconomic indicators at school and student level, continue to highlight strong differences between geographical areas and between different study paths, especially in the secondary education.

The results of the National Surveys conducted by INVALSI showed similar results regarding territorial differences and differences between the fields of study. As for the variance within schools and between schools, the general performance of the outcomes is quite similar, but the values are different, the percentage of variance attributable to differences between schools is less large (INVALSI, 2018).

Hence the importance of comparing the PISA tests with the National Surveys and verify whether and to what extent any differences may be associated with the different survey designs and/or the characteristics of the instruments used (Caponera, Losito and Palmerio, 2019).

In this paper, this exploratory analysis was carried out on the Mathematics tests used in PISA 2018 and on those used in the National Surveys conducted in 2018.

The objective of this contribution is, in fact, to compare the results in Mathematics obtained by Italian students attending the second year of secondary school participants in PISA 2018 with those obtained by the same students participating in the INVALSI national tests in the same year.

The results of the National Surveys conducted by INVALSI on students learning showed similar results to those of the PISA test about territorial differences and differences between study paths. Previous studies using a multilevel analysis model (OECD, 2013; INVALSI, 2015a) have, however, shown greater variation in the percentage of variance between schools, which is wider for PISA tests.

A first part of this paper analyzes the PISA reference frameworks and the INVALSI surveys to verify the presence of differences in the design and implementation of the two surveys and in the construction of the item. The results of the comparison show that PISA tests are built to investigate even complex cognitive processes.

A second part of this paper is focused on the analysis of the differences in the performance of the same students in the two tests and on correlation analysis: the average scores by gender and geographical area obtained by the students in the INVALSI tests are overlapping with those of the PISA tests when considering the students as a whole. Moreover, the correlation coefficients, although high, show that the two tests, national and PISA, are only partially overlapping, providing an opportunity to draw different and complementary information on learning Mathematics.

2. The PISA and INVALSI Frames of Reference for Mathematics

To compare PISA 2018 and INVALSI 2018 in Mathematics, students' results must be compared to grade 10, which corresponds to the second class of upper secondary school.

Both took place in the spring of 2018, the two surveys are aimed at different populations by definition: PISA is aimed at 15 year old students, regardless of the class attended, while the INVALSI national survey is aimed to students in the second class of upper secondary school, regardless of age.

The PISA and INVALSI frameworks for Mathematics are different because the aims of the two surveys are different: PISA is based on skills and is not linked, by definition, to the curriculum, while the INVALSI framework is closely linked to the National Indications and Guidelines. However, as can be seen from the comparison tables below, the intersection is not empty.

Tab. 1 – Differences and similarities in the of the PISA and INVALSI Mathematics test – Focus of Content

INVALSI	*PISA*
Numbers	Quantity
Space & Figures	Space & Shape
Relantionship & Functions	Change & Relationships
Data & Prevision	Uncertainty & Data

For the INVALSI Reference Framework, as well as in the National Indications, explicit reference is made to the objects covered by the discipline (for example, the numbers are the objects covered by Arithmetic) to emphasize that it is through the study of these objects that the disciplinary plant is built.

The choice to classify INVALSI items also according to macro-processes is determined by the fact that one of the fundamental objectives for students, made explicit in the legal indications of all school levels, is the acquisition of the ability to use Mathematics to read and represent reality, as well as of course the Mathematical tools necessary to carry out this Mathematization and obtain the results to be interpreted in the initial context.

The categories used by the PISA survey, consistent with the definition of Mathematical literacy, that is, the ability of a person to formulate, use and interpret Mathematics in various contexts and includes Mathematical reasoning and the use of concepts, procedures, data and Mathematical tools to describe, explain and predict phenomena.

This competence helps individuals to recognize the role that Mathematics plays in the world, to make assessments and to make well-founded decisions that allow them to be citizens with a constructive, committed, and reflective role, highlight the role these contents play in solving real problems.

The aim of the survey is to understand if 15-year-old children, who in most OECD countries are close to the end of compulsory education, are sufficiently prepared to apply their Mathematical knowledge to understand important issues and solve real problems. The PISA survey aims to describe the ability of individuals to "reason Mathematically" and to use Mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena.

To be able to use Mathematical concepts in an active way it is important that the student has a strong understanding of Mathematical concepts typical of pure Mathematics and that he is engaged in explorations of the abstract world of Mathematics. In addition, it is important that the learning of Mathematics is placed in contexts of reference related to real life and the school should provide the student with the opportunity to make this experience.

For example, the category Quantity indicates the scope of content where there are elements of quantification in contextualized problems: "It includes the quantification of attributes of real objects, relationships, situations and entities, the understanding of various modes of representation of such quantifications, and the ability to judge interpretations and arguments based on quantity. To quantify reality, one must understand measurements, counts, quantities, units, indicators, relative dimensions, trends and numerical models" (OECD, 2018).

The purpose of PISA is to provide a complete set of information to understand the functioning of the school system as a whole, for INVALSI is to return to each school the information necessary for a comparative and longitudinal exercise. The Frameworks of Reference in PISA are based on competences and are, by definition, unrelated to the curricula in INVALSI and are linked to the National Indications and Guidelines. The design in PISA is sample (averages affected by estimation errors, no variability between classes within schools) in INVALSI is census.

3. Differences and similarities between the PISA tests and the National test in Mathematical

Although the PISA tests are intended to measure whether and how students, at the end of compulsory education, have acquired Mathematical skills that allow them an active and conscious participation in everyday life (social and working), the basic contents of the construction of the tests are similar to those usually provided for in the national curricula.

The aim of the national curricula should be, in fact, to provide students with the knowledge, skills and tools necessary to deal with situations that require the use of Mathematics. PISA also intends to address these issues. For this reason, the Mathematical contents measured in PISA are largely like those of the national curricula and are:
– Change and relationships;
– Space and shape;
– Quantity;
– Uncertainty and data.

The two reference frameworks have many points in common, although their specific characteristics remain evident. The following table illustrates the contents, processes and contexts of the PISA and National Surveys tests.

Tab. 2 – Differences and similarities in the of the PISA and INVALSI Mathematics test

	INVALSI	*PISA 2018*
Content	Numbers	Quantity
	Space & Figures	Space & Shape
	Data & Previcision	Uncertainty & Data
	Relationship & Functions	Change & Relationships
Process	Problem Solving	Formulate
	Argueing	Employ
	Knowing	Formulate
Contexts	Family	Personal
	Common experience of students	Employment
	Scientific	Scientific
		Public

4. Target population

PISA is a sample survey that takes place every three-year for 15-year-old students, regardless of the school degree pursued, while the National Surveys conducted by INVALSI are annually aimed at the universe of students, in this case, students of the second grade of secondary school classes.

The two surveys are different in methodology and purpose, but first of all in the reference population.

However, a subset of the respective reference populations is in common.

Many students attended, in Spring 2018, in both investigations PISA and INVALSI: the incidence of these students is similar, about 80 percent of the total, for Italy as a whole.

So, we used data from cohort of OECD PISA 2018 and INVALSI 2018. The surveyed students are representative samples of the population of grade 8, 9, 10 and 11 for OECD PISA, that samples according to age, while for INVALSI we use students of grade 10. The resulting dataset then contains all the students who have done the Math tests for both Surveys.

Taking into account the geographical area the composition of the data-set is: 15% in North West, 22% in the North East, 27% in the Centre, 13% in the South and 24% in the South and Islands for a total number of 7,009 students.

Taking into account the gender in adding to geographical aspect (Fig. 1) we found a homogeneous picture. Overall there are 3,553 females and 3,456 males.

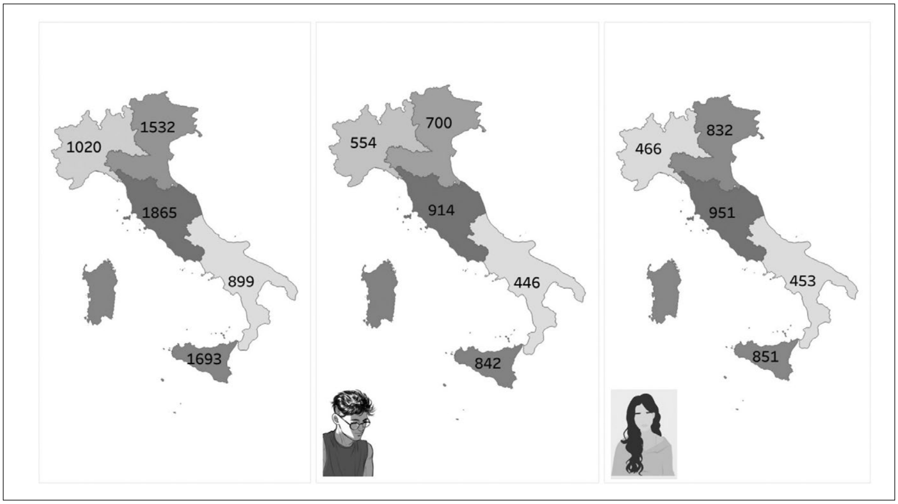

Fig. 1 – Students distribution over Italy considering also gender

Moreover in a further step of the analysis we taking into account students that obtained a score at grade 8 in the INVALSI National Survey of 2016 in the Mathematics test to verify whether, even in terms of previous INVALSI score, there is a relationship in terms of prediction of the outcome achieved by the student in PISA 2018.

So, more precisely, our database include 7,009 students that performed INVALSI and PISA Survey in Spring 2018 and INVALSI in Spring 2016.

5. Objective and Research Hypotheses

Starting from two different surveys, PISA, at International level and INVALSI assessment, at National level, our research was focused to find a common reading key.

Surveys are different especially in strictly methodological aspects, PISA allows comparisons between the participating countries that have different school systems, so it is based on the competences of the students and is independent from the different national curricula. In addition, PISA introduced the innovative concept of Literacy and the last definition of literacy for the field of Mathematics is: the ability of an individual to formulate, employ, and interpret Mathematics in a variety of contexts. It includes Mathematical reasoning and the use of Mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to recognize the role that Mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective citizens (OECD PISA 2018).

The INVALSI National Assessment verifies what the legislation requires students to be able to do. The starting point of the preparation of the Tests are the National indications and the MI Guidelines. Starting from these documents, INVALSI develops Reference Frameworks that detail the skills and knowledge that must be measured through the Tests.

Finally, we aimed to better understand the differences and similarities of the two surveys.

Research questions we intended to answer are:

– Are the two Survey somehow similar?
– Two different frameworks can speak the same language?

6. Data and methodology

In this analysis we decided to use the scores obtained by the students in the two Surveys at the Mathematics test. Although in 2018 the PISA focus was literature, we opted for the Mathematics test because: tests despite their different structures are very similar in terms of content and processes and the last year available for student data at the grade 10 of the INVALSI is 2017/18. In summary we can say that PISA is based on a larger set of open-ended questions while the INVALSI on a smaller set with a smaller number of open-ended questions.

First, we focused the attention on the differences in student performance through correlation analysis. We continued the analysis using a multiple regression model to answer to this question: there is a relationship in terms of prediction of outcome achieved by students in PISA 2018? We conducted the analysis using IDB Analyzer, application developed by the IEA Data Processing and Research Center (IEA-DPC) that can be used to analyse PISA data, in particular it can process the 5 Plausible Value[1] that are the score of PISA.

Despite the great availability of information related to the socioeconomic context of the students in the INVALSI databases in this analysis we decided to use as explanatory variables of the PISA score the socioeconomic condition represented by the ESCS index[2], gender and macro-geographical detail together with INVALSI scores[3] achieved in 2018 and 2016.

7. Results

7.1. Descriptive Statistics

A general overview of our data shows that the scores obtained from the sample of students examined in our analysis are in line with those of the national sample of INVALSI National Survey 2018, which stands at an average of 200 and a standard deviation of 40.

Regarding results obtained in the PISA 2018 Mathematics test, we note a higher value than the Italian average of 487, in line with the OECD of 489: in fact, in our sample there is a general average score of 501.64.

Observing the data, considering gender, in Tab. 3 differentials can be found in line with those present in the National/International Surveys: in fact in PISA 2018 boys get 23 points more than girls, which is more than double the difference found in other OECD countries, while in INVALSI 2018 the difference, statistically significant, drops to 8 points.

[1] International surveys such as PISA report student performance through plausible values. The cognitive data in PISA are scaled with the Rasch Model and the performance of students is denoted with plausible values (PVs). For minor domains, only one scale is included in the international databases. For major domains, a combined scale and several subscales are provided. For each scale and subscale, five plausible values per student are included in the international databases. Plausible values are imputed values similar to the individual test scores, meaning they have approximately the same distribution as the measured latent trait. Plausible values have been developed as an approximation to obtain consistent estimates of population characteristics in assessment situations where students are given too many items to allow accurate estimates of their abilities.

[2] The index is built considering educational level and occupational condition of parents and some family's resources available (Campodifiori *et al.*, 2010).

[3] We use in the analysis the WLE score, built on the basis of Raschs' Model taking into account students' ability and difficult of items.

Tab. 3 – Descriptive statistics by gender

		INVALSI 2018		*PISA 2018*	
	N	*Mean*	*Std Dev.*	*Mean*	*Std Dev.*
Female	3,553	197.33	38.01	490.24	86.04
Male	3,456	205.61	41.53	513.31	90.58
Total	7,009	201.42	40.00	501.64	89.07

Even considering the geographical disaggregation, looking by geographical macro-area, we found the same differentials and the well-known picture that contrasts the areas of Northern Italy with those of Southern part (Tab. 4).

Tab. 4 – Descriptive statistics by macro-area

		INVALSI 2018		*PISA 2018*	
	N	*Mean*	*Dev. Std*	*Mean*	*Dev. Std*
North West	1,020	215.35	40.05	537.14	82.06
North East	1,532	213.96	37.39	529.19	78.08
Centre	1,865	198.68	38.24	506.50	83.46
South	899	190.90	37.12	467.41	85.18
South & Islands	1,693	183.11	36.72	456.71	87.35
Total	7,009	201.42	40.00	501.64	89.07

7.2. Correlation Analysis

First we focused the attention on the differences in student performance through correlation analysis reported in Table 5. It's possible to see that, for combo INVALSI 2018 and PISA 2018, there is a strong positive relation between scores of the two survey, regardless gender and/or macro-area, suggesting that higher the PISA score higher the INVALSI score.

The results also show a level of correlation ranging from 0.65 in the South to 0.70 in the North West while there are no particular differences considering the gender (0.71 for males and 0.70 for females).

Moreover we included in the analysis the INVALSI score at grade 8 from National Survey INVALSI of 2016: the correlation between PISA 2018 score and INVALSI 2016 score is still positive and strong but however less strong than the previuos one and keep the same scheme of correlation considering gender and/or geographical detail. It should be underlined the low correlation (0.38) between PISA 2018 score and INVALSI 2016 score in the South and Island. The low correlation could be due to a reduced match between student information in 2016 compared to 2018 particularly in the southern regions, where abstentionism from tests is higher. These results are in line with what we expected: the score of the current year is potentially a better predictor than the previous score even if the high and positive value compared to 2016 suggests a consistency and effectiveness of the National Tests over time.

Tab. 5 – Correlation results by gender and macro-area

		Correlation INVALSI 2018 and PISA 2018	*Correlation INVALSI 2016 and PISA 2018*
Gender	Male	0.71	0.56
	Female	0.70	0.50
Macro-area	North West	0.70	0.64
	North East	0.66	0.67
	Centre	0.68	0.55
	South	0.65	0.51
	South & Islands	0.68	0.38
Total		0.71	0.60

7.3. Model

The results reported in Tab. 6 refer to the first implemented model that includes: ESCS index at student level, gender (female is the baseline), macro-area (North West is the baseline) and the INVALSI WLE score, express on Rash scale, at Grade 10. It's possible to see that ESCS index and INVALSI score have a positive relation on PISA's score and have the greatest weight, especially the INVALSI score; live in the South Italy is a disadvantage while be a male and live in the North West is an advantage. Moreover, we included in the analysis the INVALSI score at grade 8 from National Survey of 2016. So, more precisely, our database includes students that performed INVALSI and PISA Survey in Spring 2018 and INVALSI 2016. Also in this case, there is a strong relationship between PISA score 2018 and INVALSI 2016 score, less strong but still of high intensity.

We repeated the multiple regression model including also the INVALSI score of 2016. The results are very similar to the previous model (Tab. 6): the INVALSI scores are important regressors and, as we expected, the 2018 INVALSI score is more important. So, in conclusion we can say that there is a strong relationship between the Surveys and the analyzes that give us a coherent and correlated picture.

Tab. 6 – Regression outcome first model

	Regression Coefficient	*Regression Coefficient (SE)*	*Stndrdzd. Coefficient*	*Stndrdzd. Coefficient (t.value)*
Constant	229.51	8.64		
ESCS_STUDENTE_18	8.47	1.63	0.09	5.13
WLE_MAT_200_18	1.40	0.04	0.63	39.59
Male	10.25	2.72	0.06	3.84
North East	-5.88	4.06	-0.03	-1.45
Centre	-7.05	4.54	-0.03	-1.54
South	-34.03	4.99	-0.16	-6.45
South & Islands	-33.48	5.18	-0.13	-6.73

Tab. 7 – Regression outcome second model

	Regression Coefficient	*Regression Coefficient (SE)*	*Stndrdzd. Coefficient*	*Stndrdzd. Coefficient (t.value)*
Constant	188.70	8.04		
ESCS_STUDENTE_18	7.69	1.63	0.08	4.68
WLE_MAT_200_16	0.50	0.05	0.22	9.94
WLE_MAT_200_18	1.10	0.05	0.49	22.00
Male	9.83	2.66	0.06	3.75
North East	-4.44	3.67	-0.02	-1.21
Centre	-8.33	4.28	-0.04	-1.93
South	-37.38	4.69	-0.18	-7.43
South & Islands	-40.75	5.08	-0.16	-8.25

8. Conclusions

The comparative examination of the results of PISA 2018 and INVALSI conducted by INVALSI in the school year 2017-2018 showed a remarkable degree of consistency between the two sources.

This consistency is generally enhanced when the focus is on the results of the common subjects between the two sources, namely 15-year-old secondary school students. We are talking about about 75% of each of the two surveys, one (PISA) referring to 15-year-old students in general (including anticipating students, late students and those in vocational training), the other (INVALSI) referring to those attending secondary school (including, therefore, students over 15 years but late and those in advance, who are in II despite having less than 15 years).

A first objective of this study was to compare the theoretical reference frameworks for Mathematics used as a basis for developing the cognitive tests of PISA and the National Surveys.

The analysis carried out in the first part of this study, relating to the comparison between the two theoretical reference models that are the basis for the construction of the tests, showed a substantial similarity as regards the contents presented in the two tests, PISA and INVALSI. Analysis of the data also shows two important results. First, the average scores by gender and geographical area obtained by students at the INVALSI tests are substantially overlapping with those of the PISA tests. Secondly, the correlation coefficients, although high, show that the two tests, national and PISA, are only partially overlapping, providing information complementary to the learning of Mathematics. The fact that a large part of the variance between the two tests is not common could be attributable both to the different purposes of the two tests – the national test is explicitly linked to the national curricula unlike the PISA tests – both to the fact that the characteristics of the tests are different. In PISA, in fact, there are more articulated open-ended questions that allow to collect data that have a greater variability and that better differentiate the high levels from the advanced ones.

Even when we calculate the correlation between PISA 2018 score and INVALSI 2016 score we verify that it is still positive and strong but less than the previuos one and keep the same scheme of correlation. Overall, the results are entirely in line with those expected. But what interests us most here is the fact that the two surveys give us an absolutely coherent and unique picture. PISA and INVALSI present very similar results, a substantial confirmation of what has already emerged from the descriptive analysis.

References

Baker D., Goesling B., Letendre G. (2002), "Socioeconomic Status, School Quality and National Economic Development: A Crossnational Analysis of the 'Heyneman-Loxley Effect' on Mathematics and Science Achievement", *Comparative Education Review*, 46, 3, pp. 291-312.

Baye A., Monseur C. (2016), "Gender differences in variability and extreme scores in an international context", *Large-scale Assessments in Education*, 4 (1), retrieved on January, 30, 2023, from http://dx.doi.org/10.1186/s40536-015-0015-x.

Boyden J., Mann G. (2005), "Children's risk, resilience, and coping extreme situations", in M. Ungar (*ed.*), *Handbook for Working with Children and Youth: Pathways to Resilience Across Cultures and Contexts*, SAGE, Thousand Oaks, pp. 3-26.

Breen R., Jonsson J.O. (2005), “Inequality of Opportunity in Comparative perspective: recent research on Educational Attainment and Social Mobility”, *Annual Review of Sociology*, 31, pp. 223-243.
Buchmann C. (2002), “Measuring family background in international studies of education: Conceptual issues and methodological challenges”, in A.C. Porter, A. Gamoran (*eds.*), *Methodological Advances in Cross-National Surveys of Educational Achievement*, National Academy Press, Washington, DC, pp. 150-197.
Caponera E., Losito B., Palmerio L. (2019), “Le prove nazionali INVALSI e l’indagine internazionale PISA 2015: un confronto tra i risultati in Matematica e Lettura”, in P. Falzetti (a cura di), *Uno sguardo sulla scuola. II Seminario “I dati INVALSI: uno strumento per la ricerca”*, FrancoAngeli, Milano, text available at https://series.francoangeli.it/index.php/oa/catalog/book/372, date of consultation 22/4/2024.
Cerna L. (2016), *Immigration Policies and the Global Competition for Talent*, Palgrave Macmillan, Basingstoke.
Chiu M.M., Xihua Z. (2008), “Family and Motivation Effects on Mathematics Achievement: Analyses of Students in 41 Countries”, *Learning and Instruction*, 18 (4), pp. 321-336.
Crane J. (1996), “Effects of home background, SES, and maternal test scores on Mathematics achievement”, *Journal of Educational Research*, 89 (5), pp. 305-314.
Ercikan K., McCreith T., Lapointe V. (2005), “Factors Associated with Mathematics Achievement and Participation in Advanced Mathematics Courses: An Exami-nation of Gender Differences from an International Perspective”, *School Science and Mathematics*, 105 (1), pp. 5-14.
Feinstein L., Duckworth K., Sabates R. (2008), *Education and the Family: Passing Success across the Generations*, Routledge, London.
Hanushek Eric A., Woessmann L. (2007), *The Role of Education Quality for Economic Growth*, Policy Research Working Paper, No. 4122, World Bank, Washington, DC.
INVALSI (2018), *Rilevazioni nazionali sugli apprendimenti 2017-2018. Rapporto risultati*, retrieved on January, 30, 2023, from https://www.INVALSI.it/INVALSI/doc_evidenza/2018/Rapporto_prove_INVALSI_2018.pdf.
Mullis I.V.S., Martin M.O., Foy P., Drucker K.T. (2012), *PIRLS 2011 International Results in reading*, TIMSS & PIRLS International Study Center, Boston College, Chestnut Hill.
Mullis I.V.S., Martin M.O., Foy P., Hooper M. (2016), *TIMSS 2015 International Results in Mathematics*, TIMSS & PIRLS International Study Center, Boston College, Chestnut Hill.
OECD (2019a), *PISA 2018 Assessment and Analytical Framework*, PISA, OECD Publishing, Paris, retrieved on January, 30, 2023, from https://dx.doi.org/10.1787/b25efab8-en.
OECD (2019b), *PISA 2018 Results*, Vol. II: *Where All Students Can Succeed*, PISA, OECD Publishing, Paris, retrieved on January, 30, 2023, from https://doi.org/10.1787/b5fd1b8f-en.

Osterlind S.J. (1998), *Constructing test items: Multiple-choice, constructed-response, performance, and other formats*, Kluwer Academic Publishers, Boston.
Sandoval-Hernandez A., Cortes D. (2012), *Factors and Conditions that Promote Academic Resilience: A Cross-Country Perspective, Paper presented at the annual meeting of the 56th Annual Conference of the Comparative and International Education Society*, Caribe Hilton, San Juan.
Scharenberg K., Rudin M., Müller B., Meyer T., Hupka-Brunner S. (2014), *Education Pathways from Compulsory School to Young Adulthood: The First Ten Years*, TREE, Basel.
Sirin S.R. (2005), "Socioeconomic Status and Academic Achievement: A Meta-analytic Review of Research", *Review of Educational Research*, 75 (3), pp. 417-453.
White K.R. (1982), "The relation between socioeconomic status and academic achievement", *Psychological Bulletin*, 91 (3), pp. 461-481.
Willms J.D. (2006), *Learning Divides: Ten Policy Questions about The Performance and Equity of Schools and Schooling Systems*, UIS Working Paper No. 5, UNESCO Institute of Statistics, Montreal.

The authors

Francesco Annunziata, graduate in Sociology and Policy-Making from a Territorial Perspective at the University of Salerno, works in the International Surveys department at INVALSI monitoring test administrations and verifying the consistency of data for OECD and IEA surveys.

Gianluca Argentin is an Associate Professor at the Department of Sociology and Social Research of the University of Milan Bicocca. He mainly deals with three lines of investigation: a. the counterfactual impact evaluation of interventions aimed at tackling school inequalities and at improving students' performances; b. the investigation of the teaching force in the Italian education system, both focusing on their occupational conditions and on their role in the reproduction of social inequalities; c. the analysis of education inequalities in the school system.

Cecilia Bagnarol graduated in Statistics, Economics and Business at the Alma Mater Studiorum of Bologna. After severals years in INVALSI, where she worked at Statistical Service supporting activities for statistical analysis on large data bases of national and international surveys on learning, now she works at INAPP's Statistical Service as a member of ESS and PIAAC national team.

Caterina Balenzano is Assistant Professor in General Sociology at the University of Bari – Department of Political Science and studies social protection and support policies, programs and interventions aimed at vulnerable minors and families, including measures aimed at children and adolescents living in educational poverty and/or at risk of school dropout.

Elisa Caponera is a researcher at INVALSI. She was Italian National Research Coordinator (NRC) for IEA TIMSS and ICILS projects. Her current themes of research are parent involvement at school, gender difference in Mathematics achievement, school effectiveness and equity of school system.

Paola D'Elia is Research Fellow at the Department of Neuroscience Imaging and Clinical Sciences (University "G. d'Annunzio" of Chieti-Pescara). Postgraduate in School Psychology, her research interest focuses on digital learning and school climate, and how to promote inclusive education and effective learning.

Sergio Di Sano is Assistant Professor of School Psychology at the University of Chieti-Pescara – Department of Neuroscience, Imaging and Clinical Sciences, and studies learning and adaptation processes in the school environment, with particular reference to reading skills and the school climate, collaborating with schools in improvement projects based on participatory research-action.

Patrizia Falzetti, Technologist Director, she is the Head of the INVALSI Area of the Evaluation Research, of the SISTAN Statistical Office and of the INVALSI Statistical Service which manages data acquisition, analysis and return about both national and international surveys on learning (OECD and IEA). She coordinates and manages the process about returning data and statistical analysis to every school and to the Ministry of Education and Merit.

Paola Giangiacomo is a senior researcher at the National Institute for Educational and Educational Education Assessment (INVALSI), where she holds the position of National Data Manager for the surveys promoted by the OECD. Her main activities concern the revision and adaptation of survey instruments, the definition of sampling plans, the statistical analysis of quantitative and qualitative data, the drafting of technical and scientific reports, training activities for data analysis.

Sabrina Girardi is Research Fellow at the Department of Neuroscience Imaging and Clinical Sciences (University "G. d'Annunzio" of Chieti-Pescara). Postgraduate in School Psychology, her research interest focuses on digital learning and school climate, and how to promote inclusive education and effective learning.

Lorenzo Mancini is Senior Data Scientist at CROS NT. PhD in Statistical Science at the University of Bologna. His research interests involves clustering and classification methods, latent variable models and machine learning.

Rita Marzoli held a master degree in Foreign Languages, Literatures, and Linguistics (German studies). Now she is editorial board coordinator for 'INVALSI per la ricerca', a double-blind peer-reviewed series published by Franco Angeli. Her research interests are: systematic review, bibliometrics, open access, research evaluation, information literacy.

Giorgio Monti completed the PhD in Economics at University of Bologna in 2020 with a thesis titled "Essays in Economics of Education". He is currently a researcher for Consilient Research in Hargeisa, Somalia, where he is focusing in education project in humanitarian and development settings.

Riccardo Orlando graduated with a degree in Mathematics at Università degli Studi di Milano and teaches at ISIS Einaudi near Bergamo.

Laura Palmerio senior researcher at INVALSI, is head of the International Surveys department. Italian Responsible for OECD and IEA projects. She is currently a member of INVALSI Scientific Council. Main research interests: equity in education, relations between literacy in Reading and in Mathematics.

Ornella Papa is a Specialist in Psychological Evaluation and a Researcher at the INVALSI. She has been Teaching Assistant at the Chair of "Introduction to Educational and Evaluative Research" at the University of Rome Tor Vergata. She has dealt with the OECD PISA Projects, Evaluation and Improvement, INVALSI Tests about Special Educational Needs. She collaborates with the INVALSI Library in which she deepened the role of School Libraries for student achievement, the themes of the Digital divide and Innovative learning environments. Her current research interests are in the field of educational equity, school dropout.

Ottavio G. Rizzo is assistant professor in Math Education at the University of Milan. His main research interests are the teaching and learning of Mathematics in service courses at the tertiary level, Pre-service Teacher education, cryptographic applications of arithmetic geometry.

Chiara Tamanini has a Master's Degree in History and Philosophy and has been working for IPRASE (Institute for Educational research and Experimentation in the Province of Trento) for more than twenty years, mainly dealing with evaluation and assessment. She has also been cooperating with INVALSI for many years, for which she contributed to projects such as Vales and V&M, to external assessment as part of NEV, and to construction of tests for Secondary School. She is author and editor of many books and papers.

Loris Vergolini is an Assistant Professor at the University of Bologna since 2021, he carries out his research activity in the field of policy evaluation with a particular focus on measures aimed at boosting university participation. His research interests include the study of inequalities in educational opportunities and social cohesion.

INVALSI, as a part of the National System of Evaluation of the Education and Training System (SNV), conducts periodic and systematic tests on students' knowledge and skills. Albeit with some modifications over time, these standardised tests have been objectively measuring for about 20 years students' achieving and learning in some main skills in Italian, Mathematics and English domains.

In addition to conducting the National Survey, INVALSI coordinates and ensures the participation of Italy in certain main international surveys in education promoted by OECD (Organisation for Economic Cooperation and Development) and IEA (International Association for the Evaluation of Educational Achievement) which, both of them, carry out specific tests on some students' literacies and skills.

At the end of each survey, INVALSI makes useful databases available for studying and analysing the Italian education system – with an international comparison as well – and, on the occasion of the VI Seminar "INVALSI data: a tool for teaching and scientific research" (Rome, from 25th to 28th November 2021), the potential of their use became evident. This volume collects some papers presented there.

The book is therefore full of insights on the possible uses of national and international surveys. We hope that from it reading, researchers, teachers and all stakeholders could find further stimuli to better investigate the Italian education system thanks to INVALSI data and beyond.

Patrizia Falzetti, Technologist Director, is the Head of the INVALSI Area of the Evaluation Research, of the SISTAN Statistical Office and of the INVALSI Statistical Service which manages data acquisition, analysis and return about both national and international surveys on learning (OECD and IEA). She coordinates and manages the process about returning data and statistical analysis to every school and to the Ministry of Education and Merit.

FrancoAngeli
La passione per le conoscenze