THE SCHOOL AND ITS PROTAGONISTS: THE STUDENTS

V 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.

 ISBN 9788835153153

THE SCHOOL AND ITS PROTAGONISTS: THE STUDENTS

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

Edited by
Patrizia Falzetti

INVALSI

FrancoAngeli

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

For several years, in Italian school there are fewer and fewer students; this trend is offset neither by the increase of schooling nor by the rise of students with an immigrant background.

In addition to this issue, there's also the heavy problem of the school dropout. The latest data about ELET (Early Leavers from Education and Training) for Italy show that in 2020 13.1% of young people aged between 18 and 24 (or about 543 thousand) hasn't earned an upper secondary school diploma; this figure is lower compared the previous years but still higher than European Union average (ISTAT, 2020). The situation gets worse if, in addition to the figure about ELET, the implicit school drop-out is considered. Thanks to INVALSI data policy makers, researchers and all stakeholders have the chance to become aware of how many students, despite they go to school and have obtained their qualifications, do not acquire the required skills (Ricci, 2019).

Studies and researches on students achievement are therefore highly relevant because they may focus on their difficulties and strengths and could provide ideas and suggestions to help them.

This book collects six papers submitted during the V edition of the Seminar "INVALSI data: a tool for teaching and scientific research" held in Rome from 25th to 28th February 2021, with the aim of detecting which variables affect are determinant for a proper everybody's educational achievement stressing national INVALSI survey and other international ones.

Social-economical-cultural status (ESCS index) combined with emotional aspects are correlated with lower performance (chapter 2); in some cases, such as the one for the pandemic emergency from Covid-19, this social-economical-cultural background may enforce educational inequalities (chapter 3) with different sides in Italian regions (chapter 4).

In chapters 1 and 5 international survey data will be stressed, instead. In the first chapter, authors will present a study on possible learning improvement with ad additional year of school, analysing OECD-PISA data. In the fifth chapter, authors will focus on IEA ICILS 2018 data: here, what emerges is that students, even if they are digital natives, are not mastering digital skills.

The book ends with chapter 6: some first results of a qualitative research will be told. Here, the goal is to highlight some factors which are determinant for building a student perception of difficulty about Maths tasks.

Papers collected in this book are a limited and non-exhaustive example of the utility of INVALSI data. Year by year, the catalogue of the open-access series "INVALSI per la ricerca" has been greatly enriched as further confirmation of a worthwhile mutual dialogue between academics and school actors. As Statistical Service we hope this partnership could last and may generate many other research works.

1. How much do the skills of Italian 15-year-olds improve over one year of schooling?

by Francesco Avvisati, Nathan Viltard

We provide quasi-experimental measures of the learning gain that results from an additional year of schooling around the age of 15 in Italy, based on international metrics for Reading, Mathematics and Science learning established by the Programme for International Student Assessment (PISA). We combine PISA data (about 15-year-olds) for 2015 and 2018 with data about 14- and 16-year-olds in grade 10, collected as part of the Italian national extension of the PISA study to a grade-based sample. We focus, in particular, on grade repeaters (15-year-olds in grade 9) and on students with an early-school starting age (15-year-olds in grade 11), and compare them to 16- and 14-year-olds in grade 10. Results show that the grade gain from grade 10 to grade 11 for students ahead of schedule is between 11 and 13 score points (depending on the subject); while the grade gain from grade 9 to grade 10, for students one year behind schedule, is about twice as large, between 20 and 26 score points.

Questo contributo propone misure quasi-sperimentali del guadagno di apprendimento che risulta da un anno aggiuntivo di scuola intorno ai 15 anni in Italia; tale guadagno viene espresso in termini di punteggi OCSE-PISA che consentono di comparare l'apprendimento della Lettura, della Matematica e delle Scienze a livello internazionale. I dati PISA (sui 15enni) per il 2015 e il 2018 vengono combinati con i dati sui 14 e 16enni che frequentano il grado 10, raccolti nell'ambito dell'estensione nazionale italiana dello studio PISA a un campione basato sul grado. In particolare, i risultati degli studenti ritardatari (15enni che frequentano il grado 9) e anticipatari (15enni in grado 11) vengono confrontati con i risultati dei 16 e 14enni in grado 10. I risultati mostrano che il guadagno di apprendimento tra il grado 10 e il grado 11 per gli anticipatari è compreso tra 11 e 13 punti (a seconda della materia); men-

tre il guadagno di apprendimento tra il grado 9 e il grado 10, per gli studenti in ritardo di un anno, è circa il doppio, tra 20 e 26 punti.

1. Introduction

Between March 2020 and February 2021, 15-year-olds in Italy have not been able to attend regular classes in school; due to the enduring Coronavirus public health crisis, schools, instead, have organised remote support for home-based learning during much of this 12-month period. Several observers have questioned the effectiveness of these schooling surrogates, either in general or for particular types of students. The focus on "learning losses" due to school closures, however, has paid little attention to the fact that the typical learning progression observed in normal times also differs across students and schools, and remains largely unknown.

Comparing the pace of learning – i.e. learning gains associated with one grade of schooling, or grade gain – across students, regions, or countries, however, is not simple. Assessments administered to students of different grades need to be longitudinally linked, and testing conditions must remain consistent. Previous studies have quantified the average grade gain for a cohort of students based on longitudinally-linked assessments, within a single education system (Prenzel *et al.*, 2006; Nagy *et al.*, 2017; Andrabi *et al.*, 2011; Chetty, Friedman and Rockoff, 2014; Kane and Staiger, 2008). Only a few studies have been able to compare the grade gain across countries, based on international assessments (Singh, 2019; Jones *et al.*, 2014). This paper tries to address this gap by providing quasi-experimental measures of the effect of an additional year-of-schooling and year-of-age on performance based on common metrics for Reading, Mathematics and Science learning established by the Programme for International Student Assessment (PISA). We quantify the learning gain that results from an additional year of schooling around the age of 15 in Italy, and compare it to published estimates of the grade gain on PISA tests for Germany.

We combine PISA data for Italy (about 15-year-olds) with data about 14- and 16-year-olds in grade 10, collected as part of the Italian national extension of the PISA study to a grade-based sample. We focus, in particular, on grade repeaters (15-year-olds in grade 9) and on students with an early-school starting age (15-year-olds in grade 11), and compare them to 14- and 16-year-olds in grade 10. In order to overcome the limitations of a cross-sectional design and to interpret our estimates as reflecting causality, we use matching methods (inverse-probability weighting). Results show that the grade gain from grade

10 to grade 11, for students ahead of schedule, is between 11 and 13 score points (depending on the subject); while the grade gain from grade 9 to grade 10, for students one year behind schedule, is about twice as large, between 20 and 26 score points. We discuss the extent and direction of possible biases that may affect these estimates due to selection based on unobservables.

This paper is organised as follows. Section 2 introduces the data and samples used in this study. Section 3 details the identification strategy – i.e. the regression function that is estimated and the assumptions that allow to interpret one of the estimated parameters as the causal effect of an additional year of schooling on learning. Section 4 presents the results, and Section 5 discusses the results and compares them to other published results.

2. Data and context

The data used in this paper are based on datasets collected in 2015 and 2018 as part of the Programme for International Student Assessment (PISA), a large-scale, cross-national assessment of the Reading, Mathematics and Science performance of 15-year-olds students. PISA has been administered to samples of 15-year-olds students across almost 100 countries in total, every three years since 2000 (participation of countries has generally increased over time, but not all countries participated in every assessment cycle since they began taking part in PISA).

The regular PISA samples are representative of students who are enrolled in grade 7 or above and who are between 15 years and 3 months and 16 years and 2 months at the time of the assessment administration (generally referred to as 15-year-olds in this article). Together with regular PISA samples, we exploit additional samples collected by INVALSI at the same time as the regular PISA data collection. These "grade-based" samples are designed to be representative of students in the modal grade for 15-year-olds students (grade 10 in Italy). Such grade-based samples were collected in order to link the PISA dataset to a national pupil database.

PISA participants are selected from the population of 15-year-olds students in each country according to a two-stage random sampling procedure. In the first stage, a stratified sample of schools is drawn; in the second stage, students are selected at random in each sampled school. A similar two-stage sampling design was also used to draw the additional samples of grade-10 students. All statistical inference accounts for this complex sample design[1].

[1] All estimates are computed using the Stata package repest (Avvisati and Keslair, 2014).

Because of the large overlap between the PISA target population and the grade-10 target population, the first stage of sampling is common for both populations, and only within-school samples differ (although to minimise costs, the overlap between the two within-school samples is maximised, meaning that PISA-eligible students in grade 10 are automatically selected for the grade-10 sample). Different sets of weights are computed for the two samples.

We combine observations from the grade-based sample with observations from the "regular" PISA sample for the corresponding years, each weighted according to its own weights. Sample selection rules are further detailed below (*Sample selection rules and identification strategy*).

In order to match the school identifiers from both datasets, we use the national version of PISA data. Both datasets (national PISA database and grade-based database) are available upon request from the national institute for the evaluation of the education system (INVALSI)[2].

Each dataset also includes a set of balanced-repeated-replication (BRR) weights, for use in the computation of standard errors to account for the two-stage stratified sampling design. When combining observations from the two datasets, we recompute the BRR weights in one of the two datasets to match the Hadamard matrices used for the other database; thereby accounting for the interdependence between the two samples. In other words, when a school's weights are inflated in one sample, we similarly inflate all weights from the same school in the other sample; and when a school's weights are deflated, we similarly deflate all weights from the same school in the other sample.

Students' results in Reading, Mathematics and Science tests are reported on the PISA metric, a norm-referenced scale derived using item-response theory from student responses to the test. For each subject, the test norm was set to a mean of 500 and a standard deviation of 100 across students from OECD countries in a baseline year (which varies by subject), and all later tests have since been reported on the same scale. The data also include a set of background variables collected through PISA questionnaires and sampling forms; these include information about students' age, gender, socio-economic status and family background (e.g. immigrant background).

3. The quasi-experiment

In order to identify the average grade gain from cross-sectional data, we construct a pseudo-longitudinal dataset by building matched groups of stu-

[2] https://invalsi-serviziostatistico.cineca.it/.

dents who differ only in their age and length of schooling. This is made possible by the existence of significant variation in the grades of PISA students within the regular sample: in particular, we focus on students who are not in their expected grade level at age 15, and on the corresponding types of students within grade 10.

3.1. The Italian context

In Italy, the variation in grade levels for students with the same year of birth originates mostly from the practice of early school enrolment, on the one hand, and of grade repetition, on the other hand.

In particular, students can be one year ahead of their expected grade if they start primary school in the year they turn 5 ("anticipo scolastico"), a relatively common practice among students born at the beginning of the year (who would otherwise be the oldest in their school-starting cohort) and in some regions of the country. In contrast to other countries, the practice of skipping grades is virtually non-existent in Italy. Among 15-year-olds students surveyed in PISA in 2015 and 2018, 7.2% were found in grade 11 (one year ahead of schedule). Of these students, 92% were born between January and April, 79% reported (retrospectively) that they had started first grade at age 5, and 66.9% were enrolled in schools in Southern regions (compared to 38.4% among all 15-year-olds students).

Students can also fall one or more year behind their expected grade if they repeat grades. Grade repetition in Italy is very rare in primary school (grades 1 through 5) and relatively uncommon in lower-secondary school (grades 6 through 8), but is common in the first grades of upper-secondary education (where it may also coincide with a change in programme orientation). Among students who reached grade 10, the grade-based samples for 2015 and 2018 show that 1.0% of students had repeated one or more grades in primary school, 3.6% of students had repeated one or more grades in lower-secondary school, and 10.9% in upper-secondary school (based on retrospective self-reports).

3.2. Sample selection rules and identification strategy

The identification of grade effects relies on the construction of matched groups of students who differ only for being one year of age and one school grade apart. We construct each set of matching groups using strict selection rules, and further balance the observable characteristics in the two groups using inverse-probability weighting. We then use simple mean comparisons to estimate the difference in performance between the two groups.

The idea behind inverse-probability weights was originally proposed by Horvitz and Thompson (*A Generalization of Sampling Without Replacement from a Finite Universe*, 1952) in the context of sampling with varying probabilities of selection; it has since been extended to the analysis of causal effects in Rubin's potential outcome framework (Rubin, 1974). In this framework, average treatment effects correspond to the average of individual differences between a "treatment" and a "no treatment" condition, only one of which is observed. To identify and estimate causal effects, one must infer the counterfactual distribution of (potential) outcomes for the "treated" cases – i.e. the outcomes that would have been observed, had they not been treated – from a "control group". Such a control group can be built *ex ante*, through random assignment; or, in quasi-experimental methods, by invoking assumptions of conditional independence, whereby conditional on certain (observable) characteristics, the fact that a case is observed in the "treated" or "not treated" condition provides no information on any other (unobserved) characteristic which might be related to the outcome of interest.

In the present study, we estimate the difference in PISA scores between cases observed at an older age and in a higher grade (treated) and cases observed at a younger age and in a lower grade (not treated). The interpretation of these differences as reflecting point estimates of the grade gain relies on the assumption that, among the students who meet the selection rules to be included in our samples, and conditional on the observable characteristics used in the computation of inverse-probability weights, the observed performance differences between the two groups can be attributed entirely to the extra year-of-age and year-of-schooling in one of the two groups. This requires, for example, to rule out selection effects whereby particular groups of low- or high-achieving students (not identified solely by variables included in the computation of inverse-probability weights) are present only in one of the two matching groups. It also requires to rule out cohort-specific trends: if the performance of the group in the lower grade does not reflect the (unobserved) performance of the group in the upper grade, one year earlier, the estimates reflect a combination of this trend and of the grade effect. In particular, if performance at all grade-levels is improving over successive cohorts, then the estimated difference will under-estimate the real grade gain; if performance is declining, the estimated difference will over-estimate the grade gain.

In detail, we construct two sets of matching groups, by selecting students who are exactly one year ahead of schedule (Set 1) and students who are exactly one year behind schedule (Set 2).

Set 1 (students ahead of schedule) includes all students born in 2000 or in 2003 and enrolled in grade 10 (from the grade-based samples for 2015 and 2018, respectively), as well as all students born in 1999 or in 2002 and enrolled in grade 11 (from the corresponding regular PISA datasets). Set 2 (students behind schedule) includes all students born in 1999 or in 2002 and enrolled in grade 9 (from the regular PISA datasets), as well as all students born in 1998 or in 2001 and enrolled in grade 10 (from the grade-based datasets).

Table 1 illustrates how the PISA samples and the grade-based samples contribute to each of the two sets.

Tab. 1 – Sample selection rules

	Grade-based samples			*Regular age-based samples*			
	Grade	*Year of birth*	*N*	*Grade*	*Year of birth*	*N*	
Behind schedule (2 years)	10	1997/2000	681	8	1999/2002	172	Not used in the analysis
Behind schedule (1 year)	10	1998/2001	1,957	9	1999/2002	3,468	Set 2
On track	10	1999/2002	21,435	10	1999/2002	21,435	Not used in the analysis*
Ahead of schedule (1 year)	10	2000/2003	1,006	11	1999/2002	1,224	Set 1

* The same "on-track" students are included in both the grade-based sample and the PISA sample.

Source: Italian national PISA 2015 and 2018 datasets

Within each dataset (Set 1 and Set 2), we use inverse-probability weights to balance the two groups defined by their grade (and origin dataset) with respect to gender, location (5 regional dummies), month of birth, the PISA index of economic, social and cultural status (ESCS), immigrant background and curriculum orientation (academic, technical or vocational; 4 dummies). This results in two distinct estimates of grade effects, both of which may not be readily generalised to students who are "on track" or to students in different grades[3].

[3] The Stata "teffects ipw" command is used to estimate average treatment effects.

3.3. Descriptive statistics and balancing tests

In this section we present mean comparisons for background characteristics across students who are born one year apart and are enrolled in contiguous grades. We focus on our two groups of interest: students who are exactly one year ahead of schedule, and students who are exactly one year behind schedule.

As explained in the previous section, students who are ahead of schedule are expected to be a select group of students, who are born mostly in the first months of the year and in particular regions (mostly in the South)[4]. However, there are no particular reasons to expect stark differences between students who are ahead of schedule by grade 11 (and are included in the regular PISA sample) and students who are ahead of schedule by grade 10. Students who start primary education earlier than expected, and remain ahead of schedule up until 10th grade, are at low risk of repeating grades later on.

Indeed, Table 2 shows that both groups of students have similar characteristics, and few differences are statistically significant (given the number of characteristics considered, such differences may well reflect the expected variability across samples). We nevertheless use an inverse-probability weighting estimator to balance, by appropriately re-weighting the observations, the characteristics of students in the two groups.

In contrast, we expect somewhat greater differences between students who are one year behind schedule in grade 9 and students who are one year behind schedule in grade 10. Indeed, some students in the latter group may have been "on track" until they reached grade 10, and only repeated the current grade; other students who were one year behind schedule in grade 9 may have fallen further behind track by the time they reach grade 10. These two types of students are included only on one side of the comparison, and as a result, background characteristics may not be well balanced across the two sides.

[4] There are multiple explanations for such regional differences. The demographic decline in recent decades has been most pronounced in the Southern regions, and has left more room for admitting 5-year-olds in primary school. At the same time, early childhood and care services are less developed in the South, increasing demand for early entry into the (free) preschool (scuola dell'infanzia).

Tab. 2 – Characteristics of students ahead of schedule in Italy in grades 10 and 11

	Grade 10 (grade-based sample)		*Grade 11 (regular PISA sample)*		*Difference*	
	Mean	*SE*	*Mean*	*SE*	*Dif.*	*SE*
Curriculum						
Academic high school (liceo)	0.60	(0.02)	0.69	(0.02)	**0.09**	(0.02)
Lower secondary school	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Technical high school	0.27	(0.02)	0.20	(0.01)	**-0.07**	(0.02)
Vocational institute	0.11	(0.02)	0.09	(0.01)	-0.01	(0.02)
Vocational training	0.01	(0.00)	0.01	(0.00)	0.00	(0.00)
ESCS score	0.14	(0.04)	0.17	(0.04)	0.03	(0.05)
Macro-area						
Centre	0.19	(0.02)	0.17	(0.02)	-0.02	(0.02)
North-East	0.07	(0.01)	0.07	(0.01)	0.00	(0.01)
North-West	0.08	(0.01)	0.09	(0.01)	0.01	(0.01)
South	0.39	(0.02)	0.39	(0.02)	0.01	(0.02)
South and Islands	0.28	(0.02)	0.28	(0.02)	0.00	(0.02)
Sex (girl)	0.52	(0.02)	0.55	(0.02)	0.03	(0.02)
Immigrant background	0.04	(0.01)	0.04	(0.01)	-0.01	(0.01)
Month of birth (Jan = 1, Dec = 12)	2.21	(0.06)	2.23	(0.07)	0.02	(0.09)
Grade repetition						
at ISCED 1	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
at ISCED 2	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
at ISCED 3	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Missing information						
Macro-area	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
ESCS score	0.05	(0.01)	0.01	(0.00)	**-0.04**	(0.01)
Girl dummy	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Immigrant background	0.04	(0.01)	0.02	(0.01)	-0.02	(0.01)
Month of birth	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Curriculum	0.01	(0.00)	0.00	(0.00)	0.00	(0.01)

Note: Statistically significant differences are marked in bold. No adjustment is included for multiple-hypothesis testing.

Source: Italian national PISA 2015 and PISA 2018 datasets

Table 3 shows indeed a few more significant differences; students who are one year behind schedule in grade 9 are more likely to have an immigrant background, to be born later in the year, and are less likely to have repeated a grade in high school, compared to students who are one year behind schedule in grade 10. To account for these differences when estimating the grade gain for students who are one year behind schedule by grade 9, we use an inverse-probability weighting estimator.

Tab. 3 – Characteristics of students one year behind schedule in Italy in grades 9 and 10

	Grade 10 (grade-based sample)		*Grade 11 (regular PISA sample)*		*Difference*	
	Mean	*SE*	*Mean*	*SE*	*Mean*	*SE*
Curriculum						
Academic high school (liceo)	0.22	(0.02)	0.26	(0.02)	0.04	(0.02)
Lower secondary school	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Technical high school	0.37	(0.02)	0.36	(0.02)	-0.01	(0.02)
Vocational institute	0.32	(0.02)	0.29	(0.02)	-0.04	(0.02)
Vocational training	0.08	(0.01)	0.09	(0.02)	0.01	(0.02)
ESCS score	-0.53	(0.02)	-0.48	(0.03)	0.04	(0.04)
Macro-area						
Centre	0.19	(0.02)	0.17	(0.02)	-0.02	(0.02)
North-East	0.20	(0.02)	0.25	(0.02)	**0.05**	(0.02)
North-West	0.29	(0.02)	0.25	(0.02)	-0.04	(0.02)
South	0.17	(0.01)	0.16	(0.01)	-0.01	(0.01)
South and Islands	0.15	(0.02)	0.17	(0.02)	0.02	(0.02)
Sex (girl)	0.39	(0.02)	0.38	(0.02)	-0.01	(0.02)
Immigrant background	0.24	(0.01)	0.20	(0.02)	**-0.04**	(0.02)
Month of birth (Jan = 1, Dec = 12)	7.26	(0.10)	6.88	(0.09)	**-0.38**	(0.14)
Grade repetition						
at ISCED 1	0.08	(0.01)	0.06	(0.01)	-0.02	(0.01)
at ISCED 2	0.34	(0.02)	0.27	(0.02)	**-0.07**	(0.02)
at ISCED 3	0.62	(0.02)	0.75	(0.02)	**0.13**	(0.02)
Missing information						
Macro-area	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
ESCS score	0.02	(0.00)	0.08	(0.01)	**0.06**	(0.01)
Girl dummy	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Immigrant background	0.04	(0.01)	0.06	(0.01)	**0.02**	(0.01)
Month of birth	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Curriculum	0.01	(0.00)	0.00	(0.00)	0.00	(0.01)

Note: Statistically significant differences are marked in bold. No adjustment is included for multiple-hypothesis testing.

Source: Italian national PISA 2015 and PISA 2018 datasets

4. Results

Results (Table 4) show that the grade gain from grade 10 to grade 11, for students ahead of schedule, is between 11 and 13 score points (depending on the subject); while the grade gain from grade 9 to grade 10, for students one year behind schedule, is about twice as large, between 20 and 26 score points.

Tab. 4 – Grade and age effects in Italy among students ahead and behind schedule

Group (grade comparison)	*Grade and age effect*						*Number of observations*
	Mathematics		*Reading*		*Science*		
Students one year ahead of schedule (grade 10 to 11)	12.6	(4.5)	13.4	(4.6)	11.2	(4.2)	2,230
Students one year behind schedule (grade 9 to 10)	25.9	(4.2)	25.5	(4.5)	20.1	(4.0)	5,425

Note: grade and age effects for each group of students are estimated using separate regressions. Standard errors in parentheses.

Source: Italian national PISA 2015 and PISA 2018 datasets

The statistical uncertainty around these estimates means, however, that it is not possible to exclude equal effects across both subgroups at conventional levels. Nevertheless, the pattern suggesting larger grade gains for students who are lagging behind, compared to students ahead of schedule, invites some further explanation.

A possible reason behind this pattern may be found in the difference in the grades considered. Past research has documented that grade gains are often larger in lower grades, compared to higher grades (Bloom *et al.*, 2008), perhaps because of some form of decreasing returns to schooling.

A second possible reason may have to do with possible selection effects, affecting, in particular, the estimate for students behind schedule. Indeed, the group of 9th grade students in this case comprises also students who will end up repeating grade 9 or exiting the school system after grade 9; while the group of 10th grade students comprises also students who had not repeated any grade before reaching grade 10. It is plausible that the group of students who are one year behind schedule in grade 10 differs from the group of students who are one year behind schedule in grade 9 in terms of academic potential, and that our strategy of controlling for observables (through inverse-probability weighting) does not sufficiently account for such selection effects.

To explore this possibility, we provide, in Table 5, an additional set of estimates which refer only to students who reported not having repeated any grade in upper-secondary education. While this may ensure more comparable comparison groups, it also greatly reduces the sample size, and shifts the attention to an even more a-typical group of students (in Italy it is very rare to repeat grades before reaching grade 9). The estimated grade gain for this group of students is somewhat smaller (by about 25%), lending some credibility to the possibility that estimates in Table 4 are biased upward by selection effects that originate from grade repetition in grade 9 and 10; however, the large standard errors do not allow strong conclusions.

A further possibility is that the lower grade gain for students ahead of schedule reflects differences in school productivity between Northern and Southern Italy. Indeed, in the early 2000s, the practice of early enrolment was much more common in the Southern regions: as shown in Table 2, about two thirds of all students ahead of schedule are found in schools in the South, which represent only about one third of the total student population. In contrast, the regional distribution of students behind schedule is approximately proportional to that of the total student population (Table 3). Southern regions – Italy's "Mezzogiorno" – lag behind Northern regions in Italy not only economically, but also on standardised national (INVALSI, 2019) and international student assessments (Agasisti and Cordero-Ferrera, 2013; Hippe, Jakubowski and Araújo, 2018). Unfortunately, splitting the samples by geographic location results in small samples and low statistical power, meaning that it is not possible to identify statistically significant variation in grade effects by geographic area.

Tab. 5 – Grade and age effects in Italy among students who fell behind schedule before they started upper secondary education

Group (grade comparison)	*Grade and age effect*						*Number of observations*
	Mathematics		*Reading*		*Science*		
Students one year behind schedule who did not repeat grades in upper secondary education (grade 9 to 10)	18.3	(7.0)	19.1	(6.4)	14.4	(5.8)	2,432

Note: Grade and age effects for each group of students are estimated using separate regressions. Standard errors in parentheses.

Source: Italian national PISA 2015 and PISA 2018 datasets

5. Discussion and interpretation

In this final section, we jointly discuss the estimates of grade gains presented in previous sections, their interpretation, as well as their implications for policy and research.

5.1. Interpretation of grade-and-age effects

Throughout this paper, we have referred to our quantity of interest as the combined effect of being older, by one year, and having completed an additional grade level. The distinction between age-at-testing effects and the "pure" effect resulting from additional schooling, however, may be relevant when discussing the potential impact of school closures, for example. Even during periods of school closure, in the absence of any form of school instruction, maturity effects (which result from physiological ageing and from the accumulation of out-of-school experiences) would continue to contribute to skill acquisition.

Previous literature does not provide much guidance regarding the relative importance of school instruction and out-of-school experiences on skill acquisition. To answer this question, several studies in the United States have examined seasonal patterns in test scores, and some of these have highlighted a "summer learning loss": if test scores fall during the summer, then all test score gains over the school year must result from the additional schooling, rather than from maturity effects. However, the finding of an average fall in test scores during the summer break in elementary school, reported in an influential meta-analysis of early studies (Cooper *et al.*, 1996), may not reflect a real loss of skills, but rather the use of different test forms, which are only imperfectly linked, to measure Reading and Mathematics ability at the end of a school year and at the beginning of the next school years (von Hippel and Hamrock, 2019). Indeed, when more comparable tests and better scaling techniques are used to examine seasonal patterns of learning, the finding of a "learning loss" during the early school years does not always replicate (von Hippel and Hamrock, 2019). A more recent study, using a large dataset spanning eight grades of schooling (grades 1 to 8), has found that test scores decline during the summer months, but that this average loss is decreasing as students move from elementary to lower secondary grades (Atteberry and McEachin, 2020).

This literature, moreover, is largely limited to the United States. But evidence about the elementary-school years in the United States may not apply to 15-year-olds students in Italy; and the contribution of school instruction and out-of-school experiences to the skills measured in a non-curricular test, such as PISA, may differ from their contribution to learning as measured in tests more closely aligned to the school curriculum. Finally, even if some test scores do decline over the summer months, it remains unclear whether this is driven by a loss in student skills or by differences in student motivation to take the test.

The interpretation of our estimates must therefore remain tentative. If school instruction is the most important factor influencing the acquisition of reading, mathematics and science skills measured by PISA, then our estimates reflect mostly "pure" schooling effects, and periods with no instruction will likely see a stagnation or even decline in PISA test scores. If, on the other hand, school instruction complements and adds to other experiences through which students learn and practice these skills, then our estimates reflect both schooling and age-at-testing effects, and periods with no instruction will (at most) see a slowdown in test gains.

5.2. Using grade-gain estimates as a benchmark

The estimates reported in the present paper, based on quasi-experimental methods, indicate that for 15-year-olds students, the grade gain in Italy is, at best, around 25 score points, and possibly as low as 11 score points. The average grade effect estimated in the present paper for particular categories of students (grade repeaters and students ahead of schedule) can also be used, with due caution, as a benchmark for assessing the practical significance of other performance differences observed in PISA. For example, in 2018, the average proficiency in reading in the North-East of Italy was 62 score points above the average in the South-Islands region (INVALSI, 2019), a gap that is significantly wider than this grade gain. This means that it would take students in the South several years of schooling to reach the current level of their Northern peers. While tempting, a simple conversion of any difference to years-of-schooling equivalents should however be avoided; this would indeed require an extrapolation from the effect of a single grade, around the age of 15, and for somewhat particular types of students, to the cumulative effect of multiple years of schooling, for a different group of students.

5.3. Implications for future research

The present paper quantifies the learning gain that results from an additional year of schooling. To our best knowledge, it is the first paper to provide estimates, for Italy, of schooling and age effects that can credibly be interpreted as causal, and which can be compared with similar estimates for other countries. There are, however, important limitations.

A first limitation is that we cannot provide an estimate for a nationally representative group of students, and instead focus on two groups of students which are somewhat a-typical, albeit in opposite ways: grade repeaters and students ahead of schedule.

A second limitation is related to the fact that PISA data were not designed with the intention of measuring the grade gain; the statistical uncertainty associated with our estimates is large. This makes it almost impossible to determine whether learning gains differ in meaningful ways for different types of students, given that confidence interval for differences in effects will almost always include 0.

But what is the optimal design of a study whose aim is to quantify learning gains using the PISA test?

A possible answer to that question would be a longitudinal design which follows students over a complete year of schooling, and administers students a second test at one year of interval. Such a design presents an intuitive appeal, due to its reliance on within-student variation, but also considerable logistic challenges, and is not exempt from significant threats to validity. These are related to the possible selection bias that results from attrition in follow-up waves; to the possible confounding role of test design and test-administration conditions, and of student effort and motivation, on measurement invariance and on test results. In fact, such a design has already been implemented twice in Germany (based on the 2003 and 2012 PISA cohorts) and once in Poland (based on the 2009 PISA cohort). The first German study showed that over a one-year period (which corresponds both to a different age and a different grade) students gained about 25 score points in the PISA Mathematics test, on average, and progressed by a similar amount (21 points) in a test of Science (Prenzel *et al.*, 2006) – a result that is broadly in line with the estimates presented in this paper. The second study however was affected by selective attrition (Heine *et al.*, 2017) and significant test-motivation effects, and, after controlling for these (based on strong assumptions), found much smaller average performance "gains" than in the previous study; in fact, the learning gains were not significantly different from 0 in reading and in science (Nagy *et al.*, 2017).

A better design, in our view, would be one in which the PISA cohort is extended to encompass two full cohorts (24 months of birth), possibly in conjunction with an increase in sample size. Or, alternatively, one in which the PISA sample is extended to cover two contiguous grades. The simultaneous administration of PISA to two cohorts would enable greater uniformity of test administration conditions and avoid the possible negative effect on motivation of a repeated administration of a low-stakes test (Demars, 2007); while the larger size of the comparison groups would reduce the statistical uncertainty and allow more granular investigations of learning gains (e.g. by school type or region).

References

Agasisti T., Cordero-Ferrera J. (2013), "Educational disparities across regions: A multilevel analysis for Italy and Spain", *Journal of Policy Modeling*, 35 (6), pp. 1079-1102.

Andrabi T., Das J., Ijaz Khwaja A., Zajonc T. (2011), "Do Value-Added Estimates Add Value? Accounting for Learning Dynamics", *American Economic Journal: Applied Economics*, 3 (3), pp. 29-54.

Atteberry A., McEachin A. (2020), "School's Out: The Role of Summers in Understanding Achievement Disparities", *American Educational Research Journal*, 000283122093728.

Avvisati F., Keslair F. (2014), *REPEST: Stata module to run estimations with weighted replicate samples and plausible values*, Statistical Software Components S457918, Boston College Department of Economics, retrieved on June 15, 2021, from https://ideas.repec.org/c/boc/bocode/s457918.html.

Bloom H., Hill C., Black A., Lipsey M. (2008), "Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions", *Journal of Research on Educational Effectiveness*, 1 (4), pp. 289-328.

Chetty R., Friedman J., Rockoff J. (2014), "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates", *American Economic Review*, 104 (9), pp. 2593-2632.

Cooper H., Nye B., Charlton K., Lindsay J., Greathouse S. (1996), "The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review", *Review of Educational Research*, 66 (3), pp. 227-268.

Demars C. (2007), "Changes in Rapid-Guessing Behavior over a Series of Assessments", *Educational Assessment*, 12 (1), pp. 23-45.

Heine J. H., Nagy G., Meinck S., Zühlke O., Mang J. (2017), "Empirische Grundlage, Stichprobenausfall und Adjustierung im PISA-Längsschnitt 2012-2013", *Zeitschrift für Erziehungswissenschaft*, 20 (S2), pp. 287-306.

Hippe R., Jakubowski M., Araújo L. (2018), *Regional inequalities in PISA: the case of Italy and Spain*, Publications Office of the European Union, Luxembourg.

Horvitz D., Thompson D. (1952), "A Generalization of Sampling Without Replacement from a Finite Universe", *Journal of the American Statistical Association*, 47 (260), pp. 663-685.
INVALSI (2019), *OCSE PISA 2018: i risultati degli studenti italiani in lettura, matematica e scienze*, retrieved on June 15, 2021, from https://www.invalsi.it/invalsi/ri/pisa2018/docris/2019/Rapporto_Nazionale.pdf.
INVALSI (2019), *Rapporto nazionale: rapporto prove INVALSI 2019*, retrieved on June 15, 2021, from https://invalsi-areaprove.cineca.it/docs/2019/Rapporto_prove_INVALSI_2019.pdf.
Jones S., Schipper Y., Ruto S., Rajani R. (2014), "Can Your Child Read and Count? Measuring Learning Outcomes in East Africa", *Journal of African Economies*, 23 (5), pp. 643-672.
Kane T., Staiger D. (2008), *Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation*, National Bureau of Economic Research, Cambridge.
Nagy G., Lüdtke O., Köller O., Heine J.-H. (2017), "IRT-Skalierung der Tests im PISA-Längsschnitt 2012/2013: Auswirkungen von Testkontexteffekten auf die Zuwachsschätzung", *Zeitschrift für Erziehungswissenschaft*, 20 (S2), pp. 229-258.
Prenzel M., Baumert J., Blum W., Lehmann R., Leutner D., Neubrand M., Pekrun R., Rolff H.G., Rost J., Schiefele U. (*eds.*) (2006), *PISA 2003: Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*, Waxmann Verlag GmbH, Münster.
Rubin D. (1974), "Estimating causal effects of treatments in randomized and non-randomized studies", *Journal of Educational Psychology*, 66 (5), pp. 688-701.
Singh A. (2019), "Learning More with Every Year: School Year Productivity and International Learning Divergence", *Journal of the European Economic Association*, 18 (4), 1770-1813.
von Hippel P., Hamrock C. (2019), "Do test score gaps grow before, during, or between the school years? Measurement artifacts and what we can know in spite of them", *Sociological Science*, 6, pp. 43-80.

2. Some insights on the relationship between student performance and test related emotional aspects

by Mariagiulia Matteucci, Stefania Mignani, Giada Spaccapanico Proietti

Large-scale educational measurement studies usually collect a wide range of data about students' achievements and collateral information. In this study, we investigate the relation between the mathematics performance of students and background variables, by taking into account the students' emotional component defined by the test anxiety. INVALSI 2018 data for grades 5, 8, and 10 are considered. Two different approaches are implemented: a parametric one, the latent regression model, and a non-parametric one, namely regression trees. The overall results show that the emotional component plays a key role in explaining the student Mathematics score. In particular, regression trees show that the groups of students associated to the lowest mean performance are also characterized by considerable relevant anxiety levels.

Le indagini su larga scala per la valutazione degli apprendimenti tipicamente raccolgono un ricco insieme di dati sia sulla performance degli studenti sia su informazioni di contesto. L'obiettivo di questo lavoro è quello di studiare la relazione tra la performance in Matematica e le variabili di contesto, tenendo in considerazione lo stato emotivo vissuto dal rispondente, in particolare per quel che riguarda l'ansia nei confronti del test. L'analisi viene realizzata sui dati della prova INVALSI di Matematica del 2018 data per i livelli 5, 8 e 10. Vengono implementati due diversi approcci, uno parametrico, il latent regression model, e uno non-parametrico, noto come alberi di regressione. I risultati mostrano che l'impatto dell'emozione gioca un ruolo importante nello spiegare le performance in matematica. In particolare, gli alberi di regressione mostrano che i gruppi di studenti associati alle performance medie più basse sono anche caratterizzati da livelli di ansia piuttosto considerevoli.

1. Introduction

The availability of a large amount of data regarding both student performance and background information allows studying several aspects that go beyond the simple analysis of the test results. However, an in-depth analysis of these relationships may lead to interesting research developments, which could be beneficial for the continuous improvement of the evaluation instruments. In a learning analytics perspective, INVALSI makes data available including not only the test responses, but also a set of variables dealing with socio-demographic and economic characteristics, the educational path and, for some grades, student self-reported information about individual and emotional aspects. In the literature, an interesting issue deals with the emotional aspects induced by the test administration.

Examination stress and test anxiety have become extensive problems in modern society and the level of test anxiety in both students and teachers has increased as well. Some studies show indeed how the emotional component may influence the test performance (Putwain and Best, 2012).

Thus, to keep track of the impact of this emotional trait on test performance, national and international institutes for student assessment are including in their main tests some items to measure test anxiety (see, e.g., OECD, 2017).

In educational studies, it is well known that Mathematics is considered one of the most difficult and hard school disciplines due to the complex skills required and also for negative attitudes often associated with its learning. Such negative attitudes, the so called math anxiety (MA), involve feelings of worry and apprehension regarding situations involving mathematics competences (Carey *et al.*, 2016; Mammarella *et al.*, 2019).

In Italy, the impact of test anxiety on school performance is still scarcely addressed (D'Agostino *et al.*, 2022). Recently, D'Agostino *et al.* (2022) studied the relationship between anxiety and Reading, Science and Mathematics outcomes of 15-year-old Italian students by using OECD PISA data. However, in-depth studies on INVALSI data are still missing (Falzetti, 2018; Poliandri *et al.*, 2011).

During the INVALSI test administration, students are asked several questions about how they experienced the test. Students are required to make a self-evaluation on an ordinal 4-point scale. From the responses, it is possible to compute a score expressing the emotional status of the respondents (Pintrich and De Groot, 1990; Pintrich *et al.*, 1991). In this work, we investigate the relationship between the emotional component and the student performance, taking into account the interactions with other background variables, such as socio-demographic (gender, geographical area, immigrant status), economic, and those more related to the educational path (school type, enrolment, etc.). These factors are known from literature to influence the test performance or, in general, the child development of competencies (Bradley and Corwyn, 2002; Sirin, 2005; Thomson, 2018).

Also, we make an attempt to stratify the students into groups with different mean performances on the basis of their characteristics. The analysis is conducted on the mathematics test data of the scholastic year 2017-2018. The focus is on grades 5, 8 and 10, as data about the student self-evaluation are available only for those classes.

Regarding the statistical methods, two different approaches, parametric and data-driven, are implemented. In the parametric context, first a graded item response theory (IRT) model (Samejima, 1969) is implemented to measure anxiety level towards the mathematics test. Then, a model known as latent regression model (De Boeck and Wilson, 2013) is applied to simultaneously estimate student ability from the response data and the relationship between the ability and several covariates. Among the independent variables, it is possible to introduce the emotional component previously estimated. The investigation of the relation between the mathematics performance and the explanatory variables is also deepened by using quantile regression (Koenker and Bassett, 1978).

The data-driven approach involves a supervised method, specifically a regression tree (Breiman *et al.*, 1984), where the dependent variable is the estimated Rasch score from the latent regression and the covariates are the same as in the parametric model. This approach allows to identify groups of students with similar performance, and to highlight how the covariates behave in these groups. In particular, by using a recursive procedure, starting from a group containing all students, at each step, the observations are split into subgroups (nodes), gradually more homogeneous with respect to the dependent variable (the mathematics ability, in our example). The resulting tree may confirm or partially disprove what found with the parametric approach.

The main aim of the work is to investigate the effect of the emotional component on the performance from different points of view, by considering both an explanatory interpretation with the latent regression model, and an interpretation on covariates importance and students profiling with the regression tree method.

The work starts, in Section 2, with a description of the data under analysis. It follows, in Section 3, a theoretical and practical explanation of the employed statistical methods. In Section 4, the results of the application are discussed. At the end, Section 5 contains final comments and further developments.

2. Data

The data used in this work have been collected by INVALSI during the Italian national assessments in the 2017-2018 scholastic year (INVALSI, 2018). Two distinct datasets based on the nationally representative sample are used for the statistical analyses. The first dataset contains the responses to test items on mathematics and the second one collects information on student characteristics.

2.1. Cognitive test data

The study considers test responses for students in grades 5, 8, and 10. A different test was administered for each grade. A paper and pencil (P&P) test composed of 41 items was given to grade 5 students. In contrast, older students took a computer-based (CB) test, which is built on an item bank of 206 and 143 items, for grades 8 and 10, respectively. Items in the bank were assigned to students by using an incomplete design.

2.2. Student questionnaire data

The aspects covered in the student questionnaire are functional to the definition of the contextual factors, as well as to the identification of personality characteristics, other than ability, (e.g., self-regulating learning; Pintrich and De Groot, 1990) which may influence the pupils testing behaviour and results. The analysis of the data collected by this survey allows to define and activate ad-hoc intervention programs by providing individually appropriate learning environments for all children, with the aim of reducing substantial disparities. The contextual factors are defined as the intrinsic (i.e., test-related) or extrinsic determinants of the individual test scores. An example of extrinsic factor is the social environment (e.g., school, family) in which the student develops his-her ability and skills.

This work will focus on the following contextual variables at student level: economic, social and cultural status (ESCS), gender, geographical area (with categories Center, North-East, North-West, South and Islands), enrollment (early or regular and late), immigrant status (native, first-generation immigrant, second-generation immigrant). Furthermore, for students in grade 10, the school type is available with categories lyceum, technical and vocational schools.

On the other hand, personality characteristics involve student motivation and self-regulating learning. The first includes the student beliefs about his-her performance and it takes into account a variety of concepts such as perceived competence, self-efficacy, and control beliefs. The latter is related to the coping emotion-focused methods used by the student to appraise the environmental pressure that enables appropriate processing of information and decision making in learning situations (Zimmerman and Risenberg, 1997). More specifically, self-regulated learning depends on cognitive and meta-cognitive strategies, such as self-reported anxiety, rehearsal, organization, self-monitoring, time management and physical and social regulations, adopted by the student to deal with the testing challenge he-she has to tackle or, in general, when he-she has to face a learning task (Pintrich *et al.*, 1991; Schoenfeld, 1992; Pintrich, 1999; Rheinberg *et al.*, 2000; Tanner and Jones, 2003).

Since testing becomes more and more widespread, the consequences associated with test scores, such as the access to higher levels of education, are increasing. Consequently, examination stress and test anxiety have become extensive problems in modern society and the level of test anxiety in both students and teachers has increased as well. Thus, to keep track of the impact of this emotional trait on test performance, national and international institutes for student assessment are including in their main tests some items to measure test anxiety (see, e.g., OECD, 2017). These items constitute a subjective-self-report instrument that is popular in high-stake tests because they are thought to provide the most direct access to a person's subjective experiences in ego-threatening situations. Furthermore, they possess good psychometric properties, are relatively inexpensive to produce, and are simple to administer and score (Zeidner, 1998).

Specifically, in the INVALSI student questionnaire, a part of the items which are intended to measure self-regulation learning refers to the specific test which is attached to. As an example, students are required to answer four graded items regarding their own level of anxiety experienced during the mathematics test. Since the objective of the analysis is to investigate the effect of test anxiety on performance and achievement, this work focuses on this particular set of items. These items are based on the motivated strategies for learning questionnaire (MSLQ) by Pintrich and De Groot (1990). The four anxiety questions start with "Thinking about the mathematics INVALSI test…" and ask students the following: "I was worried beforehand the test", "I was so nervous during the test that I could not find the answers", "While I was taking the test, I had the impression of doing poorly", "While I was taking the test, I felt calm". The answers are expressed on an ordinal 4-points intensity scale with categories: “not at all”, “a little”, “somewhat”, and “very much so”. The last item is reverse-scored with respect to the original.

2.3. Data cleaning and pre-processing

The initial datasets were prepared for the statistical analysis. First of all, for each grade, the dataset containing the responses to the mathematics test was merged with the dataset which holds the student questionnaire data. Secondly, invalid and missing data were treated. In particular, invalid data was considered as a missing value and, afterwards, examinees who reported full missing response patterns in mathematics items and partially missing responses pattern in the covariates were removed. The latter step was required for running the latent regression and the graded IRT model. The last anxiety item was reverse-scored to be interpreted in the same direction of the other items. The resulting datasets had 23,817, 25,377 and 36,342 students, for grades 5, 8 and 10, respectively.

3. Methods

The statistical methods used in this paper are involved into three different steps. In the first step, the graded response model is used to compute an anxiety score from ordinal polytomous items. In the second step, a latent regression model is implemented to estimate the item parameters, the mathematics Rasch ability scores, and the regression coefficients for the relation between the Rasch score and the available covariates, including the anxiety score. The relation between the ability and the covariates is then deepened at various performance levels by using quantile regression. In the third step, regression trees are used to profile students. The models and methods are reviewed in the following.

3.1. Graded response model

The graded response model (GRM) proposed by Samejima (1969) is latent variable model for categorical observed variables (i.e., items) which is able to deal with graded response categories. The GRM is in fact a generalization of the popular two-parameter logistic (2PL) model for binary data (Birnbaum, 1968) to polytomous ordinal data.

Let assume that a set of K_i ordinal responses are possible for item i. The probability of endorsing the response category k, for item , with $k = 1, \dots, K_i$ conditional to the individual latent ability θ, is defined as

$$P(Y_i = k|\theta) = P^*_{ik}(\theta) - P^*_{i(k+1)}(\theta) \qquad (1)$$

Where Y_i Y_i is the ordinal response variable for item i and $P^*_{ik}(\theta)$ describes the operating characteristic curve as follows

$$P^*_{ik}(\theta) = P(Y_i \geq k|\theta) = \frac{\exp[a_i(\theta - b_{ik})]}{1 + \exp[a_i(\theta - b_{ik})]} \qquad (2)$$

In Equation (2), a_i is the discrimination parameter, expressing the capability of the item to distinguish among subjects with different abilities, analogously to the meaning in item response theory (IRT) models for binary data, and the b_{ik}, for $k = 1, \dots, K_{i-1}$, are the difficulty or threshold parameters. The number of difficulty parameters is K_{i-1} as they represent the cut-off points for the different item steps. In particular, each b_{ik} expresses the ability level necessary to have a 0.5 probability to endorse category k or higher. An important feature of the GRM is the possibility to allow for different item discrimination parameters, despite is kept constant within the same item . For this reason, the GRM is more general than Rasch-type models for polytomous items.

The GRM is considered an "undirect" IRT model, as the estimation of the conditional response probabilities requires two steps: the estimation of the operating characteristic curves in (2) and then the computation by difference of the category response curves (1). The GRM works under the assumption of unidimensionality, i.e. the existence of a single latent ability underlying the response process. However, a multidimensional generalization of the GRM is also possible.

The parameters of the GRM have been estimated by using marginal maximum likelihood (MML) with the EM algorithm in the R (R Core Team, 2019) package mirt (Chalmers, 2012).

3.2. Latent regression

Latent regression IRT models (De Boeck and Wilson, 2013) are based on the idea of simultaneously estimating student ability from the response data and the relationship between the ability and several covariates. They can be viewed as explanatory IRT models but also as nonlinear mixed models. Besides the item responses, a set of background variables are usually collected and then incorporated into the IRT model. In this way, not only the usual IRT model parameters are estimated, such as the item parameters and the individual scores, but also the coefficients of the regression of the ability on the explanatory variables. These models are used to explain person-level ability differences and are implemented in large-scale educational assessments such as the *National Assessment of Education Progress* (NAEP).

The measurement model to be used may be any IRT model, unidimensional or multidimensional, for binary or polytomous data. In our specific case, we will work with the unidimensional Rasch model (Rasch, 1960), which can be expressed as

$$P_i(\theta) = P(Y_i = 1|\theta) = \frac{\exp(\theta - b_i)}{1 + \exp(\theta - b_i)} \tag{3}$$

where the response variable Y_i for item i, may take values 0 or 1 for an incorrect or a correct response, respectively, θ is the latent ability and b_i is the difficulty parameter for item i. Besides the IRT model, the following regression model

$$\theta|\boldsymbol{x}_j = \boldsymbol{\beta}\boldsymbol{x}_j + \varepsilon \tag{4}$$

is assumed to describe the relation between the ability and a set of covariates x_j for individual j, where β is the vector of the regression coefficients and $\varepsilon{\sim}N(0; \sigma^2)$. Within this model, it is also possible to allow for random effects, besides the fixed ones.

In operational large-scale testing, the model parameters are usually estimated in two steps. First, the IRT item parameters are estimated by ignoring the subjects' covariates. Then, the remaining parameters are estimated by assuming the item parameters estimated in the previous step as known with the EM algorithm (see, e.g., von Davier and Sinharay, 2010). However, it is also possible to estimate all the model parameters jointly for specific IRT models and when it is allowed by the size of the problem by using, among others, stochastic approximations methods such as the Metropolis-Hastings Robbins-Monro (MH-RM; see Chalmers, 2015). For a review of estimation procedures and recent proposals see Andersson and Xin (2020). The parameters of the latent regression model have been estimated by the package mirt (Chalmers, 2012).

In order to investigate further the relation between the ability and the individual covariates, it is possible to use quantile regression (Koenker and Bassett, 1978). In fact, quantile regression is used to study the linear relation between a dependent variable (in this case, the ability given by the Rasch score) and a set of covariates among the quantiles of the dependent variable. In this way, the classical estimation of conditional mean models is extended to the estimation of a set of conditional quantile functions and it is possible to study the effects of the background variables on the ability at various performance levels. The quantile regression was implemented by using the quantreg package (Koenker, 2019).

3.3. Classification and regression trees

Classification and regression trees (CART; Breiman *et al.*, 1984) belong to recursive partitioning techniques and are used in supervised classification and regression problems. Given a set of variables (Y; X), where Y is the dependent variable (response) and X = (X_1, X_2, …, X_n) are the covariates observed in a sample of dimension n, the CART works by recursively splitting the sample into two groups (binary partitioning), so that the subsequent groups are more homogeneous than the former ones, with respect to the response variable Y. The split is done by choosing the covariate and the specific split point which are able to provide the best fit. This process is repeated until a stopping criterion, such as a maximum number of final groups (terminal nodes or leaves), a target fit, etc., is reached. The graphical representation of this process is made through a tree-based diagram, which makes the interpretation of the results straightforward.

The distinction between classification and regression problems is thus fundamental. In the former case, the response variable Y is continuous, while in the latter one the Y is categorical. In both cases, the covariates may be both quantitative and qualitative. In order to choose the best split, the regression tree works by evaluating the sum of squares of the node *T*, with n_T elements, $SS_T = \sum_{i=1}^{n_T} (y_i - \bar{y})^2$, and the sum of squares for the right and left sons, respectively SS_R and SS_L, for each variable and split point. The splitting criterion to be maximized is then $S_T - (SS_R + SS_L)$, and this is equivalent to maximize the between-groups sum of squares in the analysis of variance. The prevision of the response variable for each node is simply the mean, while the node error $R(T)$ is the variance.

In order to choose the proper tree size, a pruning approach may be adopted. In this case, one lets the tree grow to be as large and complex as possible, and then decide where to make a cut. This choice represents a sort of compromise between the performance of the tree and its complexity. To this end, the following cost-complexity function $R_{\alpha}(T) = R(T) + \alpha|T|$ is computed, where α is a real-valued complexity parameter and $|T|$ is the number of terminal nodes for node T. The optimal solution will be the subtree of the full tree with minimal cost, according to the cost-complexity function. It has been demonstrated that, for each tree size, there is an optimal tree. The choice of α is made by a cross-validation procedure. Another possibility for choosing the best tree is to use the 1-SE rule. According to this method, the standard error (SE) of the risk (error node) by cross-validation is computed. Then, any risk within one SE of the reached minimum is considered to be equivalent to the minimum risk. For this reason, the simplest tree among those with minimum risk is chosen.

In this work the focus is not on the predictive power of trees but on their capability to describe the relation between the response variable and the covariates. In fact, a key advantage of CART is its interpretability, as it is able to stratify the population into strata of high and low outcome (e.g. ability level or score), on the basis of the respondents' (students') characteristics. At the end of the algorithm, it is possible to obtain a certain number of groups of students with similar performance characteristics. Each group is characterized by specific values or categories of the covariates. The tree allows to improve the knowledge about the variables ruling the phenomenon under study and it allows to establish a hierarchy based on the importance of the variables and to delete those which are not relevant for making the groups. Moreover, this method does not require any assumption about the statistical distribution of the data, variables can be both quantitative and categorical, and it is possible to deal with large data, both in terms of observations and number of variables. For more details on trees, see Hastie *et al.* (2009).

4. Results

4.1. The GRM results

The aim of the first step of the analysis is to estimate a score based on the student responses to the four anxiety items (A-D), namely the "anxiety score". To this end, tab. 1 shows the estimates of the item parameters according to the GRM model for the three grades under consideration. In the interpretation of the parameters, one should consider that the anxiety is assumed to be a latent variable with a standardized normal distribution, and the threshold parameters are on the same scale. As can be seen in tab. 1, all items show a high discrimination power as the a-estimates are all above 1.5. In particular, the highest value is observed for item B in grade 10 (α = 3.228). For example, concerning the interpretation of difficulty parameters, the estimate of b_1 = -1.064 for item A in grade 5, means that an anxiety score of -1.064 is necessary to have a 0.5 probability of responding above category 1, i.e. of giving the answers "a little", "somewhat", and "very much so".

Tab. 1 – Item parameter estimates according to the GRM for the anxiety items (standard errors in brackets)

	Grade 5				*Grade 8*				*Grade 10*			
Item A	1.745	-1.064	0.253	1.310	1.924	-1.673	-0.293	0.945	2.580	-0.533	0.661	1.757
[worried]	(0.025)	(0.015)	(0.011)	(0.017)	(0.026)	(0.019)	(0.011)	(0.013)	(0.031)	(0.009)	(0.009)	(0.015)
Item B	1.863	-0.062	1.196	2.232	2.499	-0.180	1.107	2.203	3.228	0.227	1.293	2.176
[nervous]	(0.029)	(0.011)	(0.016)	(0.027)	(0.038)	(0.010)	(0.013)	(0.023)	(0.049)	(0.007)	(0.011)	(0.018)
Item C	1.929	-1.127	0.302	1.330	2.224	-1.394	0.210	1.330	1.745	-1.229	0.305	1.497
[doing poorly]	(0.028)	(0.015)	(0.011)	(0.017)	(0.030)	(0.015)	(0.010)	(0.015)	(0.020)	(0.013)	(0.009)	(0.015)
Item D	2.367	-1.011	0.164	1.178	2.493	-1.534	0.029	1.196	1.931	-1.197	0.409	1.577
[not calm]	(0.037)	(0.013)	(0.010)	(0.014)	(0.035)	(0.016)	(0.009)	(0.013)	(0.022)	(0.013)	(0.009)	(0.015)

In tab. 1, the threshold parameters for item B [nervous] are highlighted in bold as they show a peculiar behavior in comparison to the other items. In fact, the item B is a rather "extreme" item asking whether the student was feeling so nervous to not be able of finding the answers. For this item, b_1 is close to zero or higher, while for the other items it is usually around -1. Also, b_2 and b_3 are pretty much higher on item B than for the remaining items. This means that, for item B, it is less likely to respond in the higher categories than it is for items A, C, and D. This is clearly shown also in fig. 1, where the item category characteristic curves are reported for each item based on grade 10 data.

The plots (fig. 1) show the probability of endorsing each item response category as a function of the ability, i.e. the anxiety level. For item B, it is shown than students with an anxiety score lower than 0 are much more likely to respond "not at all" (black curve) than the other categories.

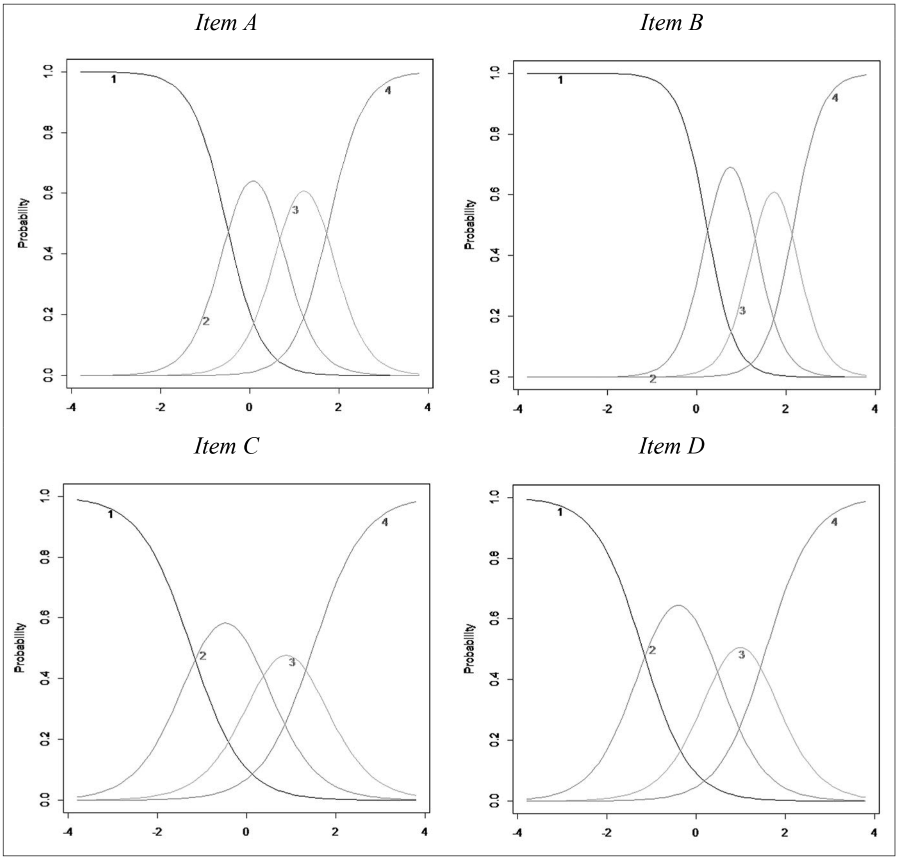

Fig. 1 – Item category characteristic curves for the anxiety items, grade 10 (in the horizontal axis, ability means the latent variable, i.e. anxiety)

4.2. The latent regression results

In the second step of the analysis, we estimated the parameters of a latent regression model, by assuming the Rasch model as the measurement model. Tab. 2 shows some statistics about the estimated difficulty item parameters and the mathematics ability scores by school grades.

Tab. 2 – Descriptive statistics on the estimated item difficulty parameters and mathematics ability scores according to the Rasch model

			Rasch model item difficulty parameters				
Grade	*Min*	*25th Qu.*	*Median*	*Mean*	*75th Qu.*	*Max*	*Sd*
5	-2.627	-0.679	-0.091	-0.148	0.578	1.846	0.996
8	-2.916	-0.726	0.082	0.047	0.776	3.372	1.140
10	-3.564	-1.543	-1.013	-1.040	-0.477	1.765	0.942
			Math ability scores				
Grade	*Min*	*25th Qu.*	*Median*	*Mean*	*75th Qu.*	*Max*	*Sd*
5	-3.306	-0.885	-0.131	-0.101	0.652	3.339	1.094
8	-3.416	-1.004	-0.251	-0.194	0.572	3.416	1.112
10	-4.303	-1.873	-1.093	-0.994	-0.187	2.418	1.153

As can be seen, the mean item difficulty and the mean ability are different from zero that is the usual assumption for IRT models. This difference is due to the presence of the "latent regression" part in the model, as a linear regression is used to model the dependency of the Math ability on the covariates (i.e., the background variables and the anxiety score). For this reason, the mean ability is shifted from zero and depends on the covariates.

In tab. 3, the estimated coefficients for the latent regression model are reported. With respect to the anxiety, a negative significant effect is found for all grades, meaning that the mathematics ability is expected to decrease as there is an increase in the anxiety level. As far as the other covariates are concerned, consistently with the annual INVALSI findings, the results show a negative and significant effect for gender in favour of male, by controlling for the other variables, for grades 5 and 10 on the mathematics performance but not for grade 8. A negative effect is also reported for having a late enrolment, with respect to regular or early, in the scholastic career and for being an immigrant student in comparison to natives (here the only not significant coefficient is for first generation immigrants of grade 10). An advantage of students living in the Northern part of the country, especially in the North-East, is observed with respect to students in the Center and, to a greater extent, to students in the South and Islands. A positive and significant effect of the ESCS is also reported. By considering grade 10 only, lyceum students seem to perform better on average with respect to students in technical and vocational schools.

To deepen the relation between the mathematics ability and the anxiety, a quantile regression model was used conditioning to all the covariates reported in tab. 3. Fig. 2 shows the plots of the estimates of the regression coefficient for anxiety with respect to the mathematics performance quantiles. The plots clearly show that, for all grades, the negative effect of the anxiety score on the mathematics ability increases through quantiles, controlling for the other variables.

Tab. 3 – Estimated regression coefficients for the latent regression model (standard errors in brackets)

	Grade		
	5	*8*	*10*
Intercept	-0.001	0.000	-0.001
	(0.019)	(0.018)	(0.015)
Gender (female)	-0.188***	0.004	-0.285***
[ref: male]	(0.015)	(0.014)	(0.011)
Enrollment (late)	-0.376***	-0.640***	-0.432***
[ref: regular or early]	(0.054)	(0.030)	(0.016)
Immigrant status (first gen)	-0.290***	-0.244***	-0.030
[ref: native]	(0.054)	(0.041)	(0.027)
Immigrant status (second gen)	-0.139***	-0.138***	-0.168***
[ref: native]	(0.028)	(0.029)	(0.021)
Geographical area (North-East)	0.052**	0.134***	0.060***
[ref: North-West]	(0.023)	(0.022)	(0.016)
Geographical area (Center)	0.052*	-0.027	-0.297***
[ref: North-West]	(0.024)	(0.022)	(0.017)
Geographical area (South and Islands)	-0.066***	-0.475***	-0.715***
[ref: North-West]	(0.021)	(0.019)	(0.015)
ESCS	0.270***	0.338***	0.091***
	(0.008)	(0.008)	(0.006)
Anxiety	-0.325***	-0.400***	-0.298***
	(0.008)	(0.008)	(0.006)
School type (technical)	-	-	-0.539***
[ref: lyceum]			(0.013)
School type (vocational)	-	-	-1.191***
[ref: lyceum]			(0.016)

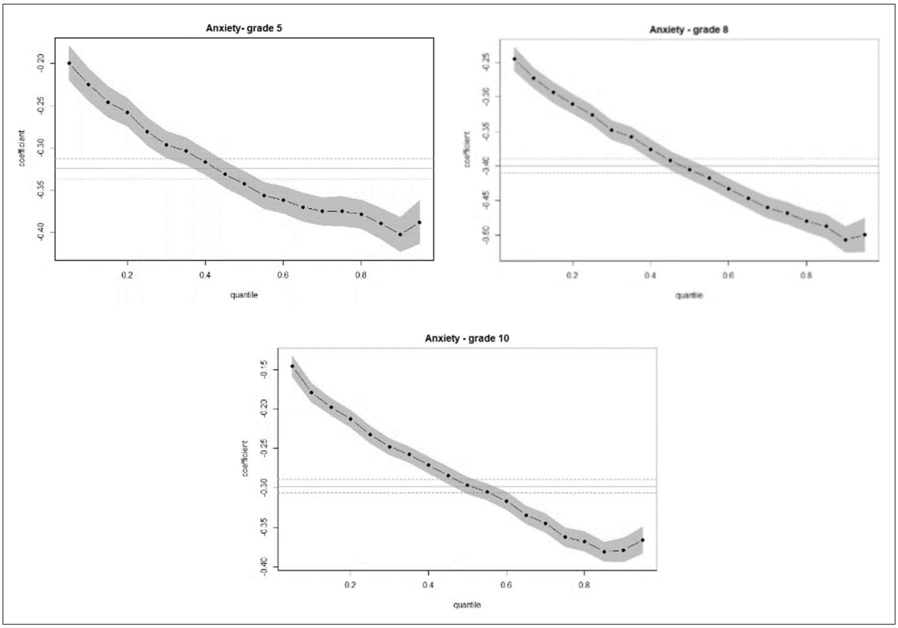

Fig. 2 – Plot of the anxiety regression coefficient for different quantiles of the mathematics ability (confidence bands in grey)

4.3. The regression tree results

As a last step in the data analysis, we employed regression trees for investigating the relation between the Mathematics performance and the covariates. As dependent variable, the INVALSI score with mean 200 has been taken into account. Fig. 3 shows the plot of the regression tree for grade 5. Six terminal nodes have been chosen by using the CART pruning procedure based on a cost-complexity function. Among the sequence of optimal trees identified, the one associated with the lowest cross-validation error (i.e., the variance) has been chosen. Besides, we also checked that the node dimension was not less than 10%. Six terminal nodes have been selected for grades 8 and 10 also. In this last case, only one node consists of about the 9% of the total sample size.

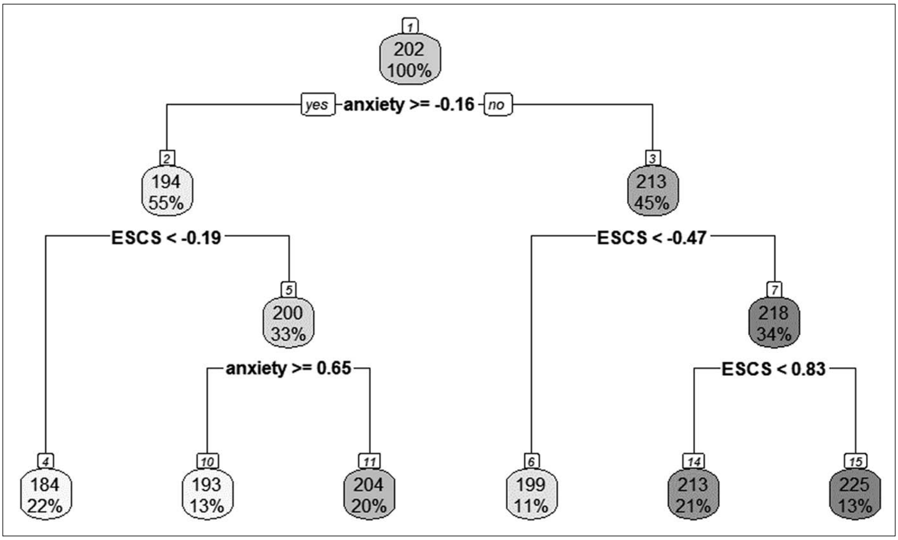

Fig. 3 – Regression tree plot, grade 5

The most interesting aspect of the tree plot in fig. 3 is that only two variables, namely anxiety and ESCS, have been chosen by the tree algorithm to split the nodes. This means that these two are the most important variables in explaining the variability of the INVALSI mathematics score. In order to describe the characteristics of the terminal nodes, for an interpretation from a classification point of view, we recoded both the anxiety score and the ESCS with categories. Both variables have mean around zero, and we set as "medium" level values from -0.2 to 0.2. Lower and higher levels have been defined consequently and by looking at the split values in the tree, as it is reported in tab. 4.

Tab. 4 – Categorization of the covariates, grade 5

Intervals defined by splits				
Anxiety score	<- 0.16 very low/low	[-0.16; 0.65 medium/high		>= 0.65 high/very high
ESCS	< – 0.47 very low	[-0.47; -0.19 low	[-0.19; 0.83 medium/high	>= 0.83 very high

The results of the tree are reported in tab. 5, ordering the nodes by increasing mean performance. The size of the nodes indicates that the number of students classified in the group is similar. Clearly, node nr. 4 (on the left in fig. 3) is associated to the lowest mean mathematics performance (184).

Students in this node are characterized by medium, high or very high anxiety levels and very low or low ESCS. On the contrary, node nr. 15 (on the right in fig. 3) is associated to the highest mean Math score (225) and contains students with a very low or low anxiety score and very high ESCS.

Tab. 5 – Summary of the characteristics of the tree terminal nodes, grade 5

Node nr.	*Node size*	*Anxiety score*	*ESCS*	*Mean performance*
4	22%	Medium; high; very high	Very low; low	184
10	13%	High; very high	Medium; high; very high	193
6	11%	Very low; low	Very low	199
11	20%	Medium; high	Medium; high; very high	204
14	21%	Very low; low	Low; medium; high	213
15	13%	Very low; low	Very high	225

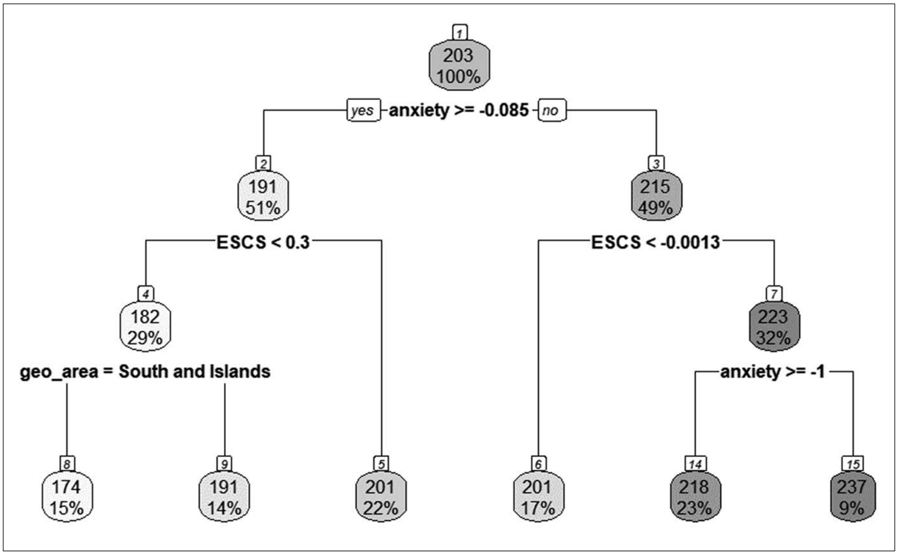

Fig. 4 – Regression tree plot, grade 8

The same approach has been used for grade 8 results. Again, as can be seen in fig. 4, the anxiety variable and ESCS have been chosen in the tree. In addition, the geographical area is used for one split.

The split values have been used to define intervals and interpret the results (see tab. 6). By looking at tab. 7, it can be noticed that students in node nr. 8 have the lowest mean performance (174). They are students again with a medium, high or very high anxiety score and very low, low or medium ESCS. Moreover, they are students living in the South and islands only. On the contrary, students in node nr. 15 have the highest mean Math ability (237) and are characterized by a very low anxiety and medium, high or very high ESCS, whatever the geographical area is.

Tab. 6 – Categorization of the covariates, grade 8

	Intervals defined by splits		
Anxiety score	<-1 very low	[-1; -0.085 [low/medium	>= -0.085 medium/high/ very high
ESCS	< -0.0013 very low/low/ medium	[-0.0013; 0.3 [medium	>= 0.3 medium/high/very high

Tab. 7 – Summary of the characteristics of the tree terminal nodes, grade 8

Node nr.	*Node size*	*Anxiety score*	*ESCS*	*Geo area*	*Mean performance*
8	15%	Medium; high; very high	Very low; low; medium	South and Islands	174
9	14%	Medium; high; very high	Very low; low; medium	North-East; North-West; Center	191
5	22%	Medium; high; very high	Medium; high; very high	Any	201
6	17%	Very low; low; medium	Very low; low; medium	Any	201
14	23%	Low; medium	Medium; high; very high	Any	218
15	9%	Very low	Medium; high; very high	Any	237

With respect to grade 10, the school type seems to be the most important variable as it is used first in the tree (see fig. 5). Besides the school type, we find again the anxiety score and the geographical area. The effect of ESCS seems to be taken up by the school type.

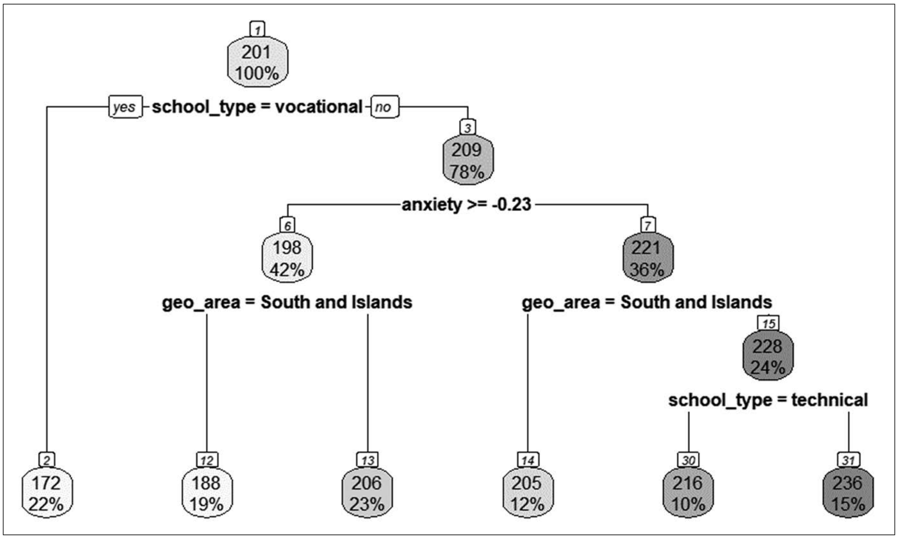

Fig. 5 – Regression tree plot, grade10

The only continuous variable, the anxiety score, has been categorized according to tab. 8.

Tab. 8 – Categorization of the covariates, grade 10

	Intervals defined by splits	
Anxiety score	< -0.23 very low/low	> = 0.23 medium/high/very high

As can be seen in the summary of tab. 9, the node nr. 2 with the lowest mean performance (172) includes students of vocational schools, no matter the anxiety score or the geographical area. On the other hand, the highest performing students in node nr. 31 (mean equal to 236) have very low or low anxiety scores, belong to lyceums, and live in the North or the Center of the country.

Tab. 9 – Summary of the characteristics of the tree terminal nodes, grade 10

Node nr.	*Node size*	*Anxiety score*	*School type*	*Geo area*	*Mean performance*
2	22%	Any	Vocational	Any	172
12	19%	Medium; high; very high	Tecnhical; Lyceum	South and Islands	188
14	12%	Very low; low	Tecnhical; Lyceum	South and Islands	205
13	23%	Medium; high; very high	Tecnhical; Lyceum	North-West; North-East; Center	206
30	10%	Very low; low	Technical	North-West; North-East; Center	216
31	15%	Very low; low	Lyceum	North-West; North-East; Center	236

5. Concluding remarks

The results of this study show that the impact of the emotions on the performance is relevant; in particular, in the latent regression model, the coefficient related to the anxiety score is significant and highlights a negative impact.

For all grades, the output of the quantile regression analysis along the ability distribution reveals that the negative effect of the anxiety score on the mathematics ability increases through the levels of ability, by controlling for the other variables. The results of the trees confirm that anxiety and ESCS are the most important aspects in explaining the variability of the INVALSI mathematics score at grades 5 and 8. With respect to grade 10, the school type seems to be the most important variable but, again, the anxiety score has a significant impact. On the other hand, the effect of ESCS seems to be taken up by the school type.

These analyses point out that it is necessary to account for individual aspects in order to obtain a more accurate interpretation of the results on the performance.

The strategy of the analysis involving two different approaches allowed to explore the multifaceted issues of assessing performances. The parametric approach based on the latent regression model assured the generalization of the results but required the assumption of a linear relation that could not be valid for these complex data. The approach based on regression trees was found to be very flexible, as it is not limited to specific assumptions on the relation between the dependent and the independent variables. Also, the interpretation of the results is straightforward. However, as it is a data-driven method, the results may be sensitive to the specific dataset. In this sense, a systematic study for different grades and years is needed.

For both approaches, the main limitation of this study is that the hierarchical structure of data is not taken into account. In fact, students are nested into classes and schools, and it is likely that the results may be affected by cluster effects. For this reason, future research may use multilevel models.

Further investigations may also consider multidimensional approaches (e.g. Italian language and Mathematics performances and related anxiety) and the use of classification trees with the INVALSI proficiency levels as dependent variable. Besides, a deeper study is needed by taking into account further information coming from the student questionnaire, such as the attitude towards mathematics. Lastly, the interaction of anxiety and gender certainly deserves attention.

Concluding, the emotional factors seem to show a relevant impact on the performances, confirming the importance to include a specific section devoted to their measurement in large-scale standardized test surveys.

References

Andersson B., Xin T. (2020), "Estimation of Latent Regression Item Response Theory Models Using a Second-Order Laplace Approximation", J*ournal of Educational and Behavioral Statistics*, 46 (2), pp. 244-265, retrieved on June 15, 2021, from https://journals.sagepub.com/doi/full/10.3102/1076998620945199.

Birnbaum A. (1968), "Some latent trait models and their use in inferring an examinee's ability, in F.M. Lord, M.R. Novick (*eds.*), *Statistical Theories of Mental Test Scores*, Addison-Wesley, Reading, pp. 397-479.

Bradley R.H., Corwyn R.F. (2002), "Socioeconomic Status and Child Development", *Annual Review of Psychology*, 53, pp. 371-399.

Breiman L., Friedman J.H., Olshen R.A., Stone C.J. (1984), *Classification and Regression Trees*, Chapman & Hall, New York.

Carey E., Hill F., Devine A., Szucs D. (2016), "The chicken or the egg? The direction of the relationship between mathematics anxiety and mathematics performance", *Frontiers in Psychology*, 6, Article 1987.

Chalmers R.P. (2012), "mirt: A Multidimensional Item Response Theory Package for the R Environment", *Journal of Statistical Software*, 48 (6), pp. 1-29.

Chalmers R.P. (2015), "Extended Mixed-Effects Item Response Models with the MH-RM Algorithm", *Journal of Educational Measurement*, 52, pp. 200-222.

D'Agostino A., Schirripa Spagnolo F., Salvati N. (2022), "Studying the relationship between anxiety and school achievement: Evidence from PISA data", *Statistical Methods & Applications*, 31, pp. 1-20.

De Boeck P., Wilson M. (2013), *Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach*, Springer Verlag, New York.

Falzetti P. (2018), *Il questionario studente. 13° Seminario tematico a invito sulla valutazione*, Roma, 13 marzo 2018, retrieved on June 15, 2021, from https://www.istruzione.it/snv/allegati/Invalsi-Falzetti%2013032018.pdf.

Hastie T., Tibshirani R., Friedman J. (2009), *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, Springer Verlag, New York.

INVALSI (2018), *Rapporto nazionale. Rapporto prove INVALSI 2018*, retrieved on June 15, 2021, from https://invalsi-areaprove.cineca.it/docs/2018/Rapporto_prove_INVALSI_2018.pdf.

Koenker R. (2019), *Quantreg: Quantile regression R package version 5.85*, retrieved on June 15, 2021, from https://cran.r-project.org/web/packages/quantreg/quantreg.pdf.

Koenker R., Bassett G. (1978), "Regression Quantiles", *Econometrica*, 46 (1), pp. 33-50.

Mammarella I.C., Caviola S., Dowker A. (2019), *Mathematics Anxiety: What is Known, and What is Still Missing*, Routledge, London.

OECD (2017), *PISA 2015 Results (Volume III): Students' Well-Being*, PISA, OECD Publishing, Paris.

Pintrich P.R. (1999), "The Role of Motivation in Promoting and Sustaining Self-Regulated Learning", *International Journal of Educational Research*, 31, pp. 459-470.

Pintrich P.R., De Groot E. V. (1990), "Motivational and Self-Regulated Learning Components of Classroom Academic Performance", *Journal of Educational Psychology*, 82, pp. 33-40, retrieved on February 21, 2023, from http://dx.doi.org-10.1037-0022-0663.82.1.33.

Pintrich P.R., Smith D.A.F., Garcia T., McKeachie W.J. (1991), *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ*), Ann Arbor: National Center for Research to Improve Postsecondary Teaching and Learning, The University of Michigan.

Poliandri D., Cardone M., Muzzioli P., Romiti S. (2011), *A Rating Scale Model for a Scale of Test Anxiety in Italy", National Institute for the Educational Evaluation of Instruction and Training (INVALSI)*, Working Paper No. 11-2011, retieved on February 21, 2023, from https//ssrn.com-abstract=1856110.

Putwain D., Best N. (2012), "Do Highly Test Anxious Students Respond Differentially to Fear Appeals Made Prior to a Test?", *Research in Education*, 88, 1, pp. 1-10.

R Core Team (2019), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Wien, retrieved on June 15, 2021, from https://www.r-project.org/.

Rasch G. (1960), *Probabilistic Model for Some Intelligence and Achievement Tests*, Copenhagen, Danish Institute for Educational Research.

Rheinberg F., Vollmeyer R., Burns B. D. (2000), "Motivation and self-regulated learning", in J. Heckhausen (*ed.*), *Advances in Psychology*, 131, *Motivational Psychology of Human Development: Developing Motivation and Motivating Development*, Elsevier Science, Oxford, pp. 81-108, retrieved on February 21, 2023, from https://doi.org/10.1016/S0166-4115(00)80007-2.

Samejima F. (1969), "Estimation of Latent Ability Using a Response Pattern of Graded Scores", *Psychometrika*, 34, pp. 1-97.

Schoenfeld A.H. (1992), "Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics", in D.A. Grouws (*ed.*), *Handbook of Research on Mathematics Teaching and Learning: A Project of the National Council of Teachers of Mathematics*, Macmillan Publishing Co., Inc., New York, pp. 334-370.

Sirin S.R. (2005), "Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research", *Review of Educational Research*, 75 (3), pp. 417-453.

Tanner H., Jones S. (2003), "Self-efficacy in mathematics and students' use of self-regulated learning strategies during assessment events", in N.A. Pateman, B.J. Doherty, J. Zilliox (*eds.*), *Proceedings at the 27th Conference of the Int. Group for the Psychology of Mathematics Education*, PME, Honolulu, vol. 4, pp. 275-282.

Thomson S. (2018), "Achievement at School and Socioeconomic Background – An Educational Perspective", *NPJ Science of Learning*, 3 (5), Article 5.

von Davier M., Sinharay S. (2010), "Stochastic Approximation Methods for Latent Regression Item Response Models", *Journal of Educational and Behavioral Statistics*, 35, pp. 174-193.

Zeidner M. (1998), *Test anxiety: The state of the art*, Plenum, New York.

Zimmerman B.J., Risemberg R. (1997), "Self-regulatory dimensions of academic learning and motivation", in G.D. Phye (*ed.*), *The Educational Psychology Series, Handbook of Academic Learning: Construction of Knowledge*, Academic Press, San Diego, pp. 105-125, retrieved on February 21, 2023, from https//doi.org-10.1016-B978-012554255-5-50005-3.

3. Territories and educational poverty: differences and characteristics and new challenges

by Barbara Baldazzi

In Italy, level of education and skills young people are able to achieve still largely depends on social background, socio-economic context and the geographical area in which one lives. These findings give rise to two research questions: What are the main variables that influence inadequate skills of students? In 2020, there was one of the most profound and sudden transformations, moving from a totally in-person teaching to a distance teaching for the last months of the 2019/20 school year and to a mixed teaching in the first months of the 2020/21 school year. Can some variables more related to the student's daily life, and closely linked to the new way of doing school at home, be protective on the achievement of competences?

Presence of books in the home (more than 100) is associated with a probability of achieving sufficient competences twice as high. Protective effect of pre-primary school attendance is weaker but still significant. Speaking with household in Italian facilitates proficiency. Owning a pc and an internet connection helps in the development of competencies. Among children who come from disadvantaged families (ESCS = first quartile) it is interesting to note that being able to use a PC and an internet connection increases the probability of having adequate skills.

In the light of what happened in 2020 and in 2021, the result presented is important. The impact of distance learning and school closures has affected a population of students already affected by profound inequalities of opportunity and, despite national and local policies, the efforts of educational institutions, teachers and families, the effects on skills and school dropout, especially in the most vulnerable segments of the population, could be particularly serious.

In Italia, il livello di istruzione e le competenze che i giovani riescono a raggiungere dipendono ancora in gran parte dall'estrazione sociale, dal

contesto socio-economico e dall'area geografica in cui si vive. Questi risultati danno origine a due domande di ricerca: Quali sono le principali variabili che influenzano le competenze inadeguate degli studenti? Nel 2020 c'è stata una delle trasformazioni più profonde e repentine del Sistema scolastico: si è passati da un insegnamento totalmente in presenza ad un insegnamento a distanza per gli ultimi mesi dell'anno scolastico 2019/20 e ad un insegnamento misto (prevalentemente a distanza per gli studenti della scuola secondaria superiore) nei primi mesi dell'anno scolastico 2020/21. Alcune variabili più legate alla vita quotidiana dello studente, e strettamente connesse al nuovo modo di fare scuola a casa, possono essere protettive sul raggiungimento delle competenze?

Una buona presenza di libri in casa (più di 100) è associata a una probabilità doppia di raggiungere competenze sufficienti. L'effetto protettivo della frequenza della scuola pre-primaria è più debole ma comunque significativo. Parlare in italiano all'interno della famiglia e possedere un pc e una connessione internet aiutano nello sviluppo delle competenze. Tra i bambini che provengono da famiglie svantaggiate (ESCS = primo quartile) è interessante notare che poter utilizzare un PC e una connessione internet aumenta la probabilità di avere competenze adeguate.

Alla luce di quanto accaduto nel 2020 e anche nel 2021, il risultato presentato è importante. L'impatto della formazione a distanza e della chiusura delle scuole ha, quindi, colpito una popolazione di studenti già segnata da profonde disuguaglianze di opportunità e, nonostante le politiche nazionali e locali, gli sforzi delle istituzioni scolastiche, degli insegnanti e delle famiglie, gli effetti sulle competenze e sulla dispersione scolastica, soprattutto nelle fasce più deboli della popolazione, potrebbero essere particolarmente gravi.

1. Introduction

In Italy, despite a steady decrease over the last 10 years, the share of young people leaving the education and training system early remains high (Baldazzi and Cascioli, 2019). Dropping out of school is only the tip of the iceberg. The difficulty for some young people to continue successfully in education and training starts early in school. Skill levels are unequally influenced by gender, citizenship, socio-economic and cultural status of the family (Save the Children, 2020). Inadequate competences are perpetuated over the years and influence school choice, learning and, ultimately, the decision to leave the school system.

In literature, economic inequality can be considered from two different angles. "The ex post view looks at differences in individual economic results or 'outcomes', like economic well-being, living standards, earnings, income, etc. The ex ante view looks at how different the circumstances involuntarily inherited or faced by individuals and affecting their economic achievements are. [...] The ex post view is referred to as inequality of outcome. [...] The ex ante view is referred to as inequality of opportunity" (Stiglitz *et al.*, 2018, p. 102).

It is important, therefore, to monitor ex ante inequality, also defined as inequality of opportunity, as a key determinant of ex post inequality. This is true for example for competencies. It should also be noted that these individual characteristics can be seen as circumstances contributing to inequality in individual outcomes, but also as an outcome of inequality explained by family-related characteristics.

In the field of education and training, inequality of opportunity is referred to as educational poverty. Every child and young person has the right to learn, train, develop his skills, competences and aspirations in the most fruitful way possible and with the best opportunities; when this right is not guaranteed, the child is in a condition of educational poverty, suffers from inequality of opportunity, which strongly and negatively affects their growth (ISTAT, 2019).

These findings give rise to two research questions: What are the main variables that influence inadequate skills of students? In 2020, there was one of the most profound and sudden transformations, moving from a totally in-person teaching to a distance teaching for the last months of the 2019/20 school year and to a mixed teaching in the first months of the 2020/21 school year. Can some variables more related to the student's daily life, and closely linked to the new way of doing school at home, be protective on the achievement of competences?

2. Data and methods

The Italian National Institute for the Evaluation of the Education and Training Educational System (INVALSI), annually, carries out standardized tests to assess the performance of all Italian students at the end of the second year of upper secondary school, or in other words, after 10 years of school attendance.

The data used concern students in the second year of upper secondary school for the school year 2018/2019. The 2018/19 school leverage at grade 10 is equal to 485,895 students taking the Reading test and 486,690 taking the Mathematics test. The first descriptive analyses were done on the two distinct populations; the subsequent multivariate analysis models were developed on a subgroup of 392,206[1] students who participated in the Reading and Mathematics tests. All the variables included in the first step and in the second step are illustrated in the tab. 1.

Tab. 1 – Variables included in the analisys

Variables, variable description and salues

- Inadequate competence in Italian literacy: share of students in grade 10, second year of upper secondary education, performing below the baseline level of proficiency in literacy competence (levels 1 e 2 out of 5 levels)
- Inadequate competence in Numeracy literacy: share of students in grade 10, second year of upper secondary education, performing below the baseline level of proficiency in numeric competence (levels 1 e 2 out of 5 levels)
- Attainment of a sufficient level in both skills: percentage of students in grade 10, second year of upper secondary education, performing above basic level in numerical and literacy skills (levels 3, 4 and 5 out of 5 levels)

Explanatory variables and values

- Gender (female, male)
- Geographical area (North, Centre, South and Islands)
- Nationality (Italian, first generation foreigner, second generation foreigner)
- Typology of Institute (high school, technical institute, professional institute)
- Regularity in the school (yes, no = lost at least one year)
- ESCS: Economic Social Cultural Status index, four quartiles (first, second, third, fourth)
- Number of books at home (0 to 25 books, 26 to 100 books, more than 100 books)
- DAD: Having an internet connection and a PC (yes, no)
- Having attended pre-primary school (yes, no)
- Language spoken at home (Italian, other language)
- Mother occupational status (employed, housewife)

The next step in the analysis was to understand if some variables more related to the daily life of the student, and linked to the new way of doing school at home, could be protective on the achievement of competences. The variables analysed, taken from the INVALSI database, are the following: *Number of books at home*; *Having an internet connection and a PC* (a single variable called *DAD)*; *Having attended pre-primary school*; *Language spoken at home*; *Mother occupational status*.

The analysis was carried out with a logistic regression model which took into account the variables illustrated above (Hosmer and Lemeshow, 1990). The dependent variable is *achievement of a sufficient level in both skills* and is worth 1 if the student achieves a sufficient level in Mathematics and Reading,

[1] The number of observations used for the logistic regression model is smaller than the number of total observations because students who participated in only one of the two tests (or only in the Mathematics test, or only in the Reading test) and students who presented a "non-response" in at least one of the dependent variables were excluded.

vice versa it is worth 0[2]. In a second step, the interactions between the variables were assessed which have a protective effect: the interactions between the DAD variable and other dependent variables were examined (two by two).

The fixed effects estimated in the model are shown in terms of odds ratios (OR). These represent the ratio between the odds of those exposed to a given risk factor and those in the target category. In this study the odds are given by the probability of having at least sufficient competence in Mathematics and Reading in relation to its complementary probability. In other words, the OR measures the association between the response variable and the covariate under examination: it is 1 in the absence of this association, it is greater than 1 when the probability of having at least sufficient competence increases in the presence of the risk factor; it is less than 1 when it decreases. In the next figures significant probabilities (OR) are in dark grey and the black line represents the absence of association.

3. Results

Students of grade 10 in the 2018/2019 school year with inadequate competence in Italian literacy are more numerous among males (39.1%) than females, among those living in the South and Islands (41.9%), among first-generation foreigners (54.2%), among students attending vocational or professional institutes (66.7%), and among those living in lower socioeconomic and cultural level families[3] (46.5%) (fig. 1).

Students of grade 10 in the 2018/2019 school year with inadequate competence in Numeracy, the percentages track the same categories but are much higher. Insufficient competence in Numeracy is present for 53.5% of those living in the South and Islands, 53.6% of first-generation foreigners, 73.4% for those attending vocational and professional institutes, and 51.7% for students living in families with low socioeconomic and cultural level. The only exception is the higher proportion of girls with inadequate skills compared to boys 42.2% vs. 33.5% (fig. 2).

[2] For the definition of the levels see the National Report Invalsi Tests 2019, https://invalsi-areaprove.cineca.it/docs/2019/Rapporto_prove_INVALSI_2019.pdf.

[3] In order to measure the socio-economic and cultural status of students, INVALSI constructs, by integrating several variables an indicator called ESCS (Economic Social Cultural Status index), standardised so that the zero value corresponds to the Italian average and each unit above or below it to the standard deviation of the distribution of values. The first quartile corresponds to the score below which 25% of the ESCS scores, in ascending order, are located. The second quartile (or median) is the score below which 50% of the measures are found, and so on.

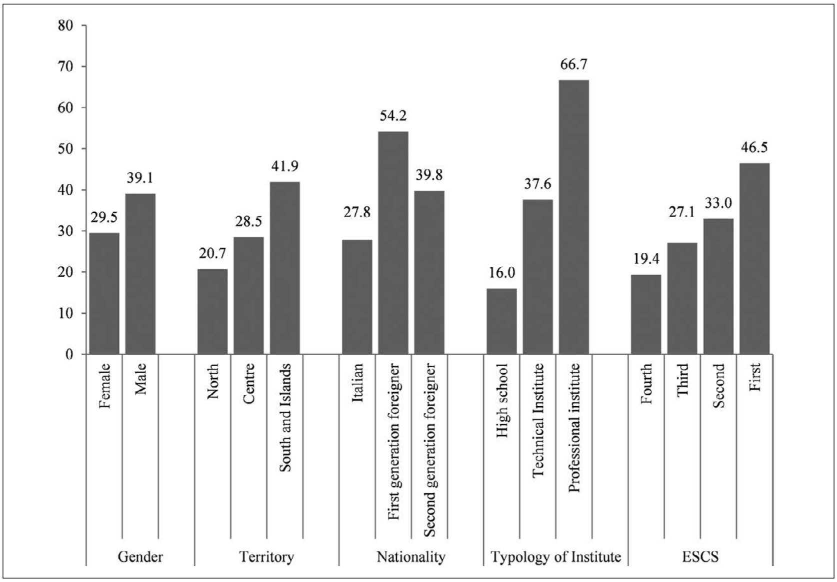

Fig. 1 – Students of grade 10 in 2018/2019, Inadequate competence in Italian literacy by gender, territory, nationality, tipology of institute, ESCS (Economic Social Cultural Status Index)

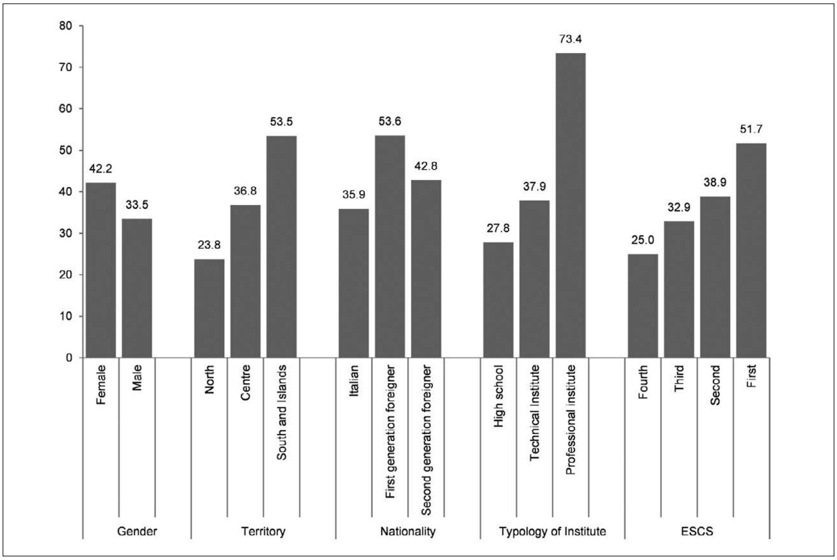

Fig. 2 – Students of grade 10 in 2018/2019, Inadequate competence in Numeracy by gender, territory, nationality, tipology of institute, ESCS (Economic Social Cultural Status Index)

On a territorial level, students living in the south and islands perform worse, and even looking at the maps of Italian provinces, it can be observed that competences in Italian literacy are better than competences in Numeracy (fig. 3).

It clearly emerges that some structural variables such as gender, nationality, socio-economic and cultural status, type of school, territory can be protective on the achievement of competences.

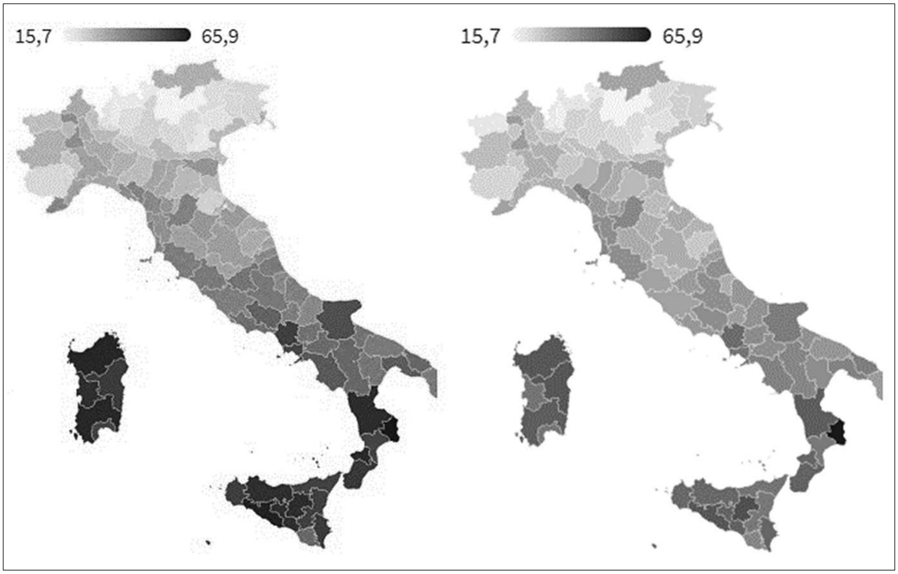

Fig. 3 – Students of grade 10 in 2018/2019, Inadequate competence in Numeracy (on the left) and in Italian literacy (on the right) by provinces

Apart from these factors, it is interesting to analyse which other stimuli can be protective with respect to the risk of not reaching an adequate level of competence in the two subjects, such as: having books, an internet connection and a PC in one's home; having attended pre-school; speaking mainly Italian at home rather than another language.

The probability of having at least sufficient skills is significantly higher for those living in the North than for those living in the South and Islands (3 times more likely); there is a slightly higher probability for boys than for girls; students who are regular in their studies are about three times more likely to have sufficient skills than those who have failed (fig. 4). Moreover, first-generation natives are more disadvantaged (significant probability less than 1).

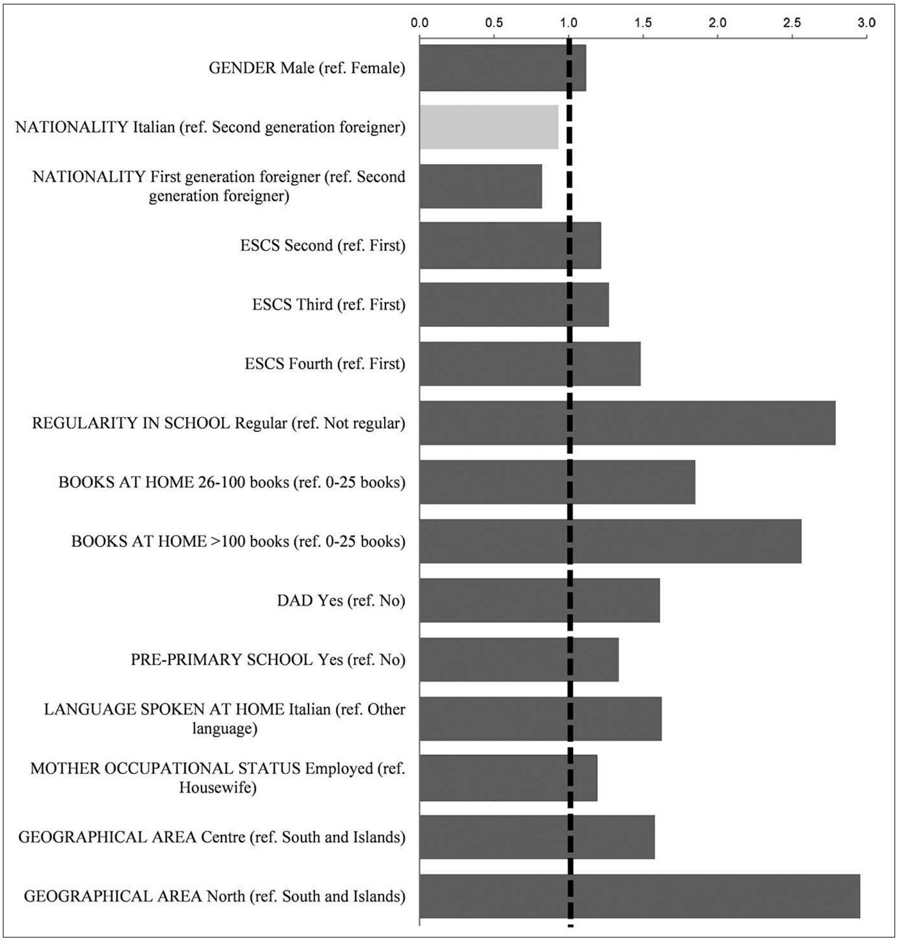

Fig. 4 – Estimates from the logistic regression model on the probability of students in grade 10 performing above basic level in Numeracy and Italian literacy skills (levels 3, 4 and 5 out of 5 levels), year 2018/2019, odds ratio (the light coloured bar identify a value which is not significantly different from 1)

As the socio-economic and cultural level of families increases, the probability of having good skills in Numeracy and Italian literacy increases. As a proxy for household cultural level, being able to count on a large presence of books in the home (more than 100 books), is associated with a probability of attaining sufficient competences 2.5 times higher than not having books or having less than 25 books.

The protective effect of attending pre-primary school is weaker but still significant, with a 34% higher probability among those who went to pre-primary school of having adequate competences compared to those who did not attend. The connection between the mother's employment status and skills is less important.

Speaking with household in Italian, even for everyday exchanges, facilitates skills (63% more than for those who habitually speak a language other than Italian).

Owning a personal computer and an Internet connection helps in developing skills: 59% more likely than those who do not have a connection and a personal computer.

For the second step of the analysis with the logistic regression model, the interaction between some variables and presence of a PC and internet connection (DAD) is examined. Associations between socio-economic and cultural level (ESCS) and presence of a PC and internet connection (DAD) was significant.

Among students with a low socio-economic and cultural level, having a PC and an internet connection increases the probability of having good skills, more than for students belonging to families with higher socio-economic and cultural levels (fig. 5).

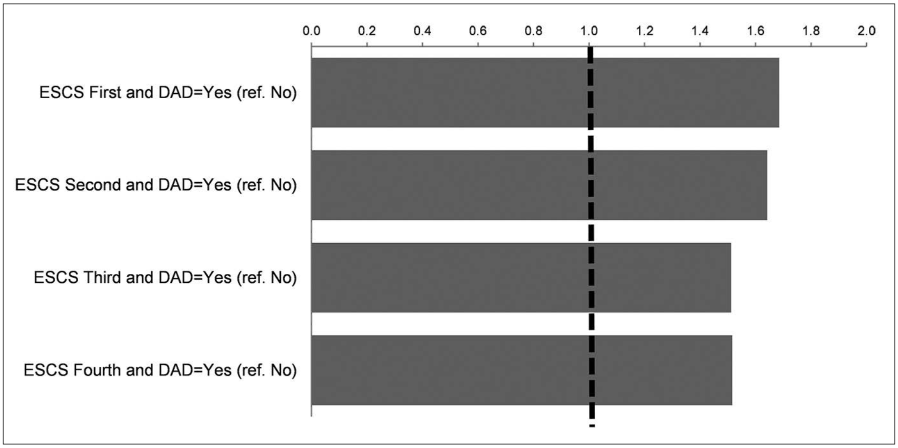

Fig. 5 – Interaction between ESCS and DAD, logistic regression model on the probability of students in grade 10 performing above basic level in Numeracy and Italian literacy skills (levels 3, 4 and 5 out of 5 levels), year 2018/2019, odds ratio

4. Conclusions

In the year 2020 the pandemic has brought a change in the forms of teaching and "doing school". There are still few data that allow to understand what has happened: for students' sociality, for skills (United Nations, 2020). Can some variables more related to the student's daily life, and closely linked to the new way of doing school at home, be protective on the achievement of competences?

A good presence of books in the home (more than 100) is associated with a probability of achieving sufficient competences twice as high. Protective effect of pre-primary school attendance is weaker but still significant (Stringher and Cascella, 2020). Speaking with household in Italian facilitates proficiency. Owning a pc and an internet connection helps in the development of competencies. Among children who come from disadvantaged families (ESCS = first quartile) it is interesting to note that being able to use a PC and an internet connection increases the probability of having adequate skills.

In the light of what happened in 2020 and in 2021, the result presented is important. In 2020, there was one of the most profound and sudden transformations, moving from a totally in-person teaching to a distance teaching for the last months of the 2019/20 school year and to a mixed teaching (mainly at a distance for upper secondary school students) in the first months of the 2020/21 school year (UNICEF, 2021).

Therefore, it becomes even more important to have a good connection and a PC or electronic device to be able to interact with school and teachers. The ISTAT survey on the "Integration of students with disabilities in schools" (ISTAT, 2020b) showed that educational institutions have equipped themselves in various forms of distance learning[4], but despite the efforts of institutions, teachers and families, 8% of children and youth in schools of all levels (23% among students with disabilities) remained excluded from any form of distance learning and did not take part in video lessons with the class group. Central regions have the lowest proportion of excluded students (5%), while in the South and Islands the proportion has almost doubled (9%). In this very particular phase of learning, having a connection and a PC, besides

[4] Distance learning (DAD): Distance learning activities involve the reasoned and guided construction of knowledge through an interaction between teachers and pupils. Whatever the medium through which didactics is exercised, the aim and the principles do not change. It takes place through direct or indirect connection, immediate or deferred, by means of videoconferences, video lessons, group chats; the reasoned transmission of teaching materials, by uploading them onto digital platforms and the use of class registers in all their communication and teaching support functions, with subsequent revision and discussion carried out directly or indirectly with the teacher, interaction on systems and interactive educational applications that are strictly digital. The mere sending of materials or the mere assignment of tasks, which are not preceded by an explanation of the content in question or which do not provide for a subsequent clarification or restitution by the teacher, are devoid of elements that can stimulate learning and are therefore not considered part of distance learning.

being a fundamental predictive factor for an adequate development of skills, becomes a requirement for access to education (ISTAT, 2020a, 2021).

The impact of distance learning and school closures has, therefore, affected a population of students already affected by profound inequalities of opportunity (OECD, 2020; Ricci, 2019) and, despite national and local policies, the efforts of educational institutions, teachers and families, the effects on skills and school dropout, especially in the most vulnerable segments of the population, could be particularly serious.

References

Baldazzi B., Cascioli R. (2019), *The early school leavers as indicator of educational poverty*, presented at Educational Poverty: Definitions, measures and Mapping, Lunch-to-lunch Workshop, Pisa, November 4-5.

Hosmer D.W. Jr, Lemeshow S. (1990), *Applied logistic regression*, John Wiley & Sons, New York.

INVALSI (2019), *Rapporto nazionale, Rapporto prove INVALSI 2019*, retrieved on June 15, 2021, from https://invalsi-areaprove.cineca.it/docs/2019/Rapporto_prove_INVALSI_2019.pdf.

ISTAT (2019), *2019 SDGs Report. Statistical Information for 2030 Agenda in Italy*, retrieved on June 15, 2021, from https://www.ISTAT.it/it/files//2019/08/SDGs-2019_inglese.pdf.

ISTAT (2020a), *2020 SDGs Report. Statistical Information for 2030 Agenda in Italy*, retrieved on February 21, 2023, from https://www.istat.it/en/archivio/247720.

ISTAT (2020b), *L'inclusione scolastica degli alunni con disabilità – A.S. 2019-2020*, retrieved on June 15, 2021, from https://www.ISTAT.it/it/files/2020/12/Report-alunni-con-disabilit%C3%A0.pdf.

ISTAT (2021), *BES 2020 Il benessere equo e sostenibile in Italia*, retrieved on June 15, 2021, from https://www.ISTAT.it/it/files//2021/03/BES_2020.pdf.

OECD (2020), *L'istruzione scolastica al tempo del Covid-19: gli insegnanti e gli studenti erano preparati?*, Country note OCSE, retrieved on June 15, 2021, from https://www.invalsi.it/invalsi/ri/pisa2018/Country_Note.pdf.

Ricci R. (2019), *Dalla licenza media alla maturità. Il percorso visto attraverso i dati INVALSI*, retrieved on June 15, 2021, from https://www.invalsiopen.it/.

Save the Children (2020), *L'impatto del Coronavirus sulla povertà educative*, retrieved on June 15, 2021, from https://www.savethechildren.it/cosa-facciamo/pubblicazioni/impatto-del-coronavirus-sulla-poverta-educativa.

Stiglitz J.E., Fitoussi J.P., Durand M. (2018), *For Good Measure Advancing Research on Well-being Metrics beyond GDP*, OECD Publishing, Paris.

Stringher C., Cascella C. (2020), "What do we know about preschool quality in Italy? Preschool effects on child outcomes: a pseudo-longitudinal exploration",

in P. Falzetti (*ed.*), *INVALSI Data: Methodologies and results. III Seminar "INVALSI data: a research tool"*, FrancoAngeli, Milano.
UNICEF (2021), *Covid-19 and school closures, One year of education disruption*, retrieved on June 15, 2021, from https://data.unicef.org/resources/one-year-of-covid-19-and-school-closures/.
United Nation (2020), *Shared responsibility, Global solidarity: Responding to the socio-economic impacts of COVID-19*, retrieved on June 15, 2021, from https://unsdg.un.org/sites/default/files/2020-03/SG-Report-Socio-Economic-Impact-of-Covid19.pdf.

4. The influence of socio-economic-cultural background on academic results in the INVALSI tests of Italian and Mathematics in two Southern regions: Puglia and Abruzzo[*]

by Sergio Di Sano, Caterina Balenzano

Among the various factors that impact academic performance, the socio-economic-cultural status (ESCS) of the student plays a significant role. This impact is often intertwined with that of territorial disparities (Bagnarol and Donno 2020; Russo *et al.*, 2020). In line with previous studies carried out on specific regions (e.g., Martini, 2020), this work intends to investigate the relationship between ESCS and academic results, comparing Puglia and Abruzzo; two Southern regions where the average learning score is slightly higher than what one might expect from an estimate based on GDP. The objective of this study is to examine the impact of socio-economic-cultural status (ESCS) on academic achievement as measured by INVALSI tests in Italian and Mathematics. Additionally, the study aims to investigate how this influence may vary based on other factors such as the student's gender, origin, delay in schooling, and the average ESCS score of the class. The results of students in the INVALSI tests of Italian and Mathematics for lower secondary school (grade 8) for the school year 2018-19 were used. Descriptive statistical analyzes were performed to investigate learning outcomes and proficiency levels in relation to the ESCS quartile of students. Different regression analysis models were then applied to see how much the influence of socio-economic variables on learning outcomes acts independently of the influence of other variables, such as gender and origin of the student, delay in the path scholastic and average ESCS of the class. The results confirm the influence of the ESCS on school results in the INVALSI tests, but also the contribution of other variables in modifying this influence. The comparison

[*] The work was conceived and discussed together, even if the drafting of paragraphs 2.2, 4.1, 4.3, 5 must be attributed to Sergio Di Sano, the drafting of paragraphs 2.3, 2.4, 2.5, 3, 4.2 must be attributed to Caterina Balenzano while the drafting of 1, 2.1 and 6 must be attributed to both.

between the models and between the relative weights of the predictors in the two regions is interpreted to provide useful information in defining regional school policies.

Tra i fattori che influenzano i risultati scolastici, un ruolo centrale è giocato dallo status socio-economico-culturale (ESCS) dello studente, il cui impatto a volte si intreccia con l'influenza dei divari territoriali (Bagnarol e Donno, 2020; Russo et al.*, 2020). In linea con precedenti indagini, svolte su specifiche regioni (per es., Martini, 2020), il presente lavoro intende indagare la relazione tra ESCS e risultati scolastici, ponendo a confronto la Puglia e l'Abruzzo; due regioni del Sud in cui il punteggio medio di apprendimento risulta di poco superiore a quello prevedibile facendo una stima basata sul PIL. L'obiettivo dello studio è indagare l'influenza dell'ESCS sui risultati scolastici alle prove INVALSI di Italiano e Matematica ed esaminare come questa influenza si modifichi in base ad altre variabili: genere e origine dello studente, ritardo nel percorso scolastico, ESCS medio della classe. Sono stati impiegati i risultati degli studenti alle prove INVALSI di Italiano e Matematica per la secondaria di primo grado (grado 8) per l'anno scolastico 2018/19. Sono state svolte analisi statistiche descrittive sui risultati di apprendimento e i livelli di competenza in relazione al quartile di ESCS degli studenti. Sono stati, poi, applicati diversi di modelli di analisi della regressione per vedere quanto l'influenza delle variabili socio-economiche sui risultati di apprendimento agisca in modo indipendente rispetto all'influenza di altre variabili, quali genere e origine dello studente, ritardo nel percorso scolastico e ESCS medio della classe. I risultati confermano l'influenza dell'ESCS sui risultati scolastici alle prove INVALSI, ma anche il contributo di altre variabili nel modificare tale influenza. Il confronto tra i modelli e tra i pesi dei predittori nelle due regioni viene interpretato per fornire indicazioni utili per la definizione di politiche scolastiche regionali.*

1. Introduction

If we examine the INVALSI data, there are many factors that influence school results and, among these, a central role is played by the socio-economic-cultural status (ESCS) of the student. The importance of this variable for understanding inequalities in education has always been a cornerstone of sociological analysis, in which various theories have been proposed (Besozzi, 2017). This issue has also been investigated in the context of school psychology, with reference to the "Social Justice" perspective (Shriberg *et al.*, 2020), which has made an important contribution to the definition of school policies.

Some more recent contributions (Bagnarol and Donno, 2020; Russo *et al.*, 2020) examining the INVALSI English language test, have underlined the importance of investigating the relationship between social stratification and territorial gaps, focusing attention on the need to consider not only the differences on the North-South axis, but also those between individual regions or specific territories. As Russo and colleagues (2020) note, the gap between central-northern and Southern students is shorter for socially advantaged students (high ESCS level), while it widens when you go down the steps of social stratification.

As Easterbrook and Hadden (2021) point out, we must take the local context seriously if we are to apply effective policies to reduce the effect of inequalities in education, and in particular the disadvantage of minorities. A policy that works well in a local context, indeed, may not work as well in a different context. Therefore, it is important to compare different regions in terms of the impact of socio-economic-cultural status on academic performance, as well as the variables that may be related to this influence.

In line with previous surveys carried out on specific regions (e.g. Martini, 2020b), the purpose of this study is to investigate the influence of the ESCS on school results in the INVALSI tests of Italian and Mathematics in two Southern regions, Puglia and Abruzzo, and to examine how this influence is modified by a number of other variables: gender and origin of the student, delay in school and average ESCS of the class.

The interest of this comparison is based on the fact that these two regions have different GDP but the same relationship between GDP and learning outcomes; in particular, the average academic achievement in both regions is slightly higher than what would be expected based on an estimate derived from GDP (Martini, 2020a).

2. A framework for the effect of social inequalities on learning outcomes

2.1. A multidisciplinary perspective on educational inequalities

Understanding the influence of socio-economic inequalities to undertake effective change paths requires a multidisciplinary perspective able to examine the topic at the macro and micro level, thus taking into account both the social structure and the agency of individuals.

At the macro level, sociology has developed complex analyses on social stratification and the reproduction of inequalities, capturing the multiform nature of the conflict between social actors, who pursue different interests. However, relying only on these types of analysis does not offer transformative perspectives if the individual is not considered as endowed with agency, and able to undertake effective paths of change, through awareness and empowerment.

At the micro-level, psychology, at least starting from the work of Bronfenbrenner (1979), has developed in-depth studies of contextual influences at the proximal and distal level, trying to understand how the individual builds his own development trajectory, starting from significant social experiences in specific contexts. But there is still a lack of full awareness of the complex intertwining between psychological variables and processes of social change, which is essential for undertaking a fight against inequalities. Only in more recent times, thanks to the contribution of Shriberg (2008), school psychology has focused attention on social justice as a field of research and intervention.

Acknowledging both the strengths and limits of sociological and psychological perspectives, in the following we focus on four interpretative lines, to be considered as the pieces of a mosaic, which together provide on the one hand the picture of powerful social forces that tend to reproduce social stratification, through family and school influences and, on the other, they open the way to intervention paths in which new social policies are flanked by psychological work on the individual, which goes in the direction of promoting empowerment, motivation for change and civic engagement.

2.2. The "Social Justice" perspective in school psychology

In the psychological field, there is a growing interest among researchers in the topic of social inequalities which represents a significant problem in a world characterized by increasing globalization. In the context of school psychology, the interest involves three different areas: advocacy for human rights, multiculturalism and the perspective of social justice.

As for the defense of human rights, it should be remembered that the UN Convention on the Rights of the Child in Article 28 (right to education), holds that education must be ensured that is accessible to all and in Article 29 (purposes of education) that education must promote the development of the child's potential. More generally, Article 2 reports the commitment to protect children from all forms of discrimination or sanctions motivated by social conditions. In this context, an interesting contribution by Nastasi and Naser (2020) proposes a vision of the school psychologist as a "human rights advocate" integrated into an ecological-evolutionary framework inspired by the theory of Bronfenbrenner (1979). The idea is that the school psychologist, in his action to defend the rights of the child, operates at the level of the mesosystem, that is, the relationship between the microsystems of the school, the family, the peer group and the community; also considering indirect influences (which characterize the esosystem). At the same time, the guiding principles and categories of rights represent a "meta-macrosystem", in the sense they influence the macrosystem, made up of society's culture, values and norms.

As for multiculturalism, the emphasis is placed on the importance of recognizing and valuing the differences between students, including their dimensions of intersection of individual identity and cultural background. Multiculturalism emphasizes diversity and the wide variety of individual differences in relation to aspects such as age, disability, gender, gender identity, race, ethnicity, national origin, religion, sexual orientation, language, and socio-economic status (Proctor, 2018). In this context, several scholars have highlighted the importance of speaking not of "minorities" but of "minoritization" to refer to individuals and groups that are treated unfairly through laws, policies and practices that lead to marginalization (Newell *et al.*, 2021). Speaking of minoritization means shifting the emphasis from a passive identity, based on the deficit, to the actions that individuals and groups of individuals take to marginalize minorities. The concept of "intersectionality" allows us to grasp the simultaneous experience of different social categories, such as race, gender, class, and nationality (Proctor, 2018). Social categories interact and are the basis of the multiple identity of students, who may perceive that they are subject to prejudice not based on a single characteristic but on a specific combination of different aspects of their identity.

As for the perspective of social justice, reference is made to the commitment of school psychologists in the advocacy of students, and of marginalized students. This perspective is closely connected to multiculturalism, as social justice cannot be promoted without multiculturalism (Newell *et al.*, 2021). Among the aspects that characterize the perspective of social justice, in the context of school psychology, are the protection of students' rights and the promotion of opportunities (Grapin and Shriberg, 2020). Social justice involves the promotion of non-discriminatory practices and the empowerment of families and communities by contributing to schools and communities characterized by equity. Shriberg *et al.* (2020) examine some previous research that investigated social justice from the point of view of school psychology and found a convergence on three basic ideas: rights, access, and respect. The idea of "rights" refers to the importance of ensuring compliance with the provisions of the Convention for the rights of the child. The idea of "access" concerns the importance of guaranteeing all children the use of society's resources, such as the access to education. The idea of "respect" refers to how people are treated, it is not enough to guarantee the right to go to school and access to school; it is also necessary to ensure that children are not discriminated against, by promoting their participation and active involvement.

It therefore becomes important to train school psychologists and teachers who are competent in terms of multicultural education and education for social justice. As Proctor (2018) notes, it is not enough to provide culturally relevant psychological services to provide adequate support to students with different needs. A broader view of multiculturalism, in the practice of school psychology and teacher work, involves five distinct aspects: a) examining and challenging personal attitudes, perceptions and beliefs of different students and populations; b) acquiring knowledge on different populations; c) understanding school issues that impact different student populations (curriculum, assessment policies and practices); d) actively engaging in policies that promote equity for different student populations; e) acting proactively to promote the values underlying multiculturalism.

2.3. Sociological theories on inequalities in education

The theme of inequalities in education and training processes represents a classic field of sociology studies, both in its theoretical dimension, when exploring the mechanisms through which cultural disadvantages are transmitted on an intergenerational level, and in its dimension of potential social change, when attempting to use the results of empirical surveys to ensure equity in educational institutions.

On the theoretical-interpretative level, various explanatory models have tried to clarify the reasons why students from disadvantaged families have higher risk of educational poverty, not reaching adequate training standards and/or adequate levels of higher education. In line with literature according to which family background (Fitzpatrick and Yoels, 1992) and socio-economic context (Vajello and Dooley, 2013) represent primary determinants for some social groups' academic success or early school leaving, a first line of research has hypothesized the idea of a "cognitive disadvantage" for students with low socio-economic status. These, compared to their peers from advantaged social backgrounds, would appear since the beginning of school less equipped from a linguistic and cultural point of view, due to a not very stimulating family environment.

In Bernstein's hypothesis (Bernstein, 1971), for example, the difference in academic success between students belonging to lower and upper social classes would be related to the two different linguistic codes generated by the different social structures: the "elaborated" code, which is used in educational contexts by wealthy classes, and the "restricted" code, used by the lower classes. These linguistic codes are not defined in terms of vocabulary, but in terms of the kinds of options speakers take up to organize what they have to say. According to Bernstein, middle-class children's success at school would be due to the use of the "elaborated code" with their parents at home. It is explicit, "context-independent" and universalistic, so these students are more able to cope with school tasks. On the contrary, students from lower families usually use a restricted code, which is "implicit', "context-bound", and "particularistic"; so, they find difficulties in coping adequately to the school demands, which are based on generalizations, classifications and conceptual abstractions.

A second line of studies focused on the role of inherited "cultural capital" (Bourdieu, 2015) which translates into certain cultural habits and greater accessibility to resources (Bourdieu and Passeron, 1971) for students from upper classes. These factors would help children from families with more high status in achieving meaningful learning goals and skills more easily. Specifically, according to theorists studying social and cultural reproduction, social class impacts school success independently from students' attitudes, as the influence of family status on student success appears to be mediated above all by family cultural capital. It, together with ethos and lifestyle (Berzano and Genova, 2011), directs cultural practices (Bourdieu, 2003), would explain the intergroup differences in student performance and would act as a reproductive mechanism of social stratification (Bourdieu and Passeron, 1971; 1972). In this logic, inequalities are reproduced both through the economic aspects that determine access to material resources (Schizzerotto and Barone, 2006), and through weak cultural practices (Devine and Savage, 2000), who increase the risk of educational failure (Suh, Suh and Houston, 2007). A close relationship has been proved, in fact, between the socio-economic and cultural status of families and academic performance (Naidoo *et al.*, 2014), as students who can count on a more stimulating environment reach, on average, higher levels of ability of those observed in lower-class children, also in international comparisons.

More specifically, scholars who have studied which mechanisms are involved in the intergenerational transmission of cultural capital highlighted the role of several factors. First, the easy access to higher education opportunities for students from advantaged families – e.g. access to extra-scholastic activities, cultural resources, and tools for learning, but also a differing ways of spending leisure time (Lareau, 2011). On the contrary, the poor social networks and the low probability of participating in extracurricular activities in disadvantaged families (Ream and Rumberger, 2008) exposes these students to considerable obstacles in realizing their full potential, and therefore their chances to achieve better Education and Training. Furthermore, in line with the hypothesis of Boudon (1979), which adds to the background resources the importance of aspirations and attitudes dictated by the social models of reference, also a difference in terms of implicit values and attitudes toward education was evidenced. While the conception of education as investment in higher classes increases students' motivation and engagement, aspirations, desires, and ambitions, and consequently, better chances of achieve good results (Fiorio and Leonardi, 2010), the values and tacit messages linked to education in lower classes are consistent with students' elevated levels of disengagement toward school and education. Specifically, literature has highlighted that families with medium-high income invest more resources in the education of their children, and that the more educated parents are able to better support and monitor the learning processes of their children (Cappellari, 2006), are more involved in school activities than parents of less well-off and educated classes who have a more peripheral role in children's learning (Butler, 2014).

According to this interpretative model, the cultural heritage becomes a school inheritance not only thanks to the quality of the family environment (Oseguera, Conchas, and Mosqueda, 2011), but also through an implicit transmission of the value of education.

2.4. The role of individual and contextual factors in the acquisition of skills

Beyond the mechanisms hypothesized to explain the relationship between status and academic performance, the predictive power of students' socio-economic-cultural background on student performance has been highlighted in numerous studies carried out in the Italian context (see Ballarino and Checchi, 2006; Checchi, 2010; Schizzerotto and Barone, 2006), which confirm how the advantages of the context of belonging are systematically transformed into greater educational success (Raimondi *et al.*, 2013). This characteristic of the student, operationalized by INVALSI with the metric approach of the Economic Social Cultural Status (ESCS) index, represents in fact one of the most significant predictors for the analysis of the results of the learning tests in all levels of education. However, the research (Thrupp, 1995; Gottfried, 2014) has shown that the educational success of a student is also influenced by other factors, the consideration of which seems to reduce the weight of his socio-economic and cultural status index (Economic, Social and Cultural Status, ESCS): a) gender; b) nationality; c) whether or not they are in good standing in the course of their studies; d) the average status level of students in the class and school attended (Martini, 2020b).

With respect to the gender dimension, the research found that girls show greater competence in linguistic tests, unlike their male peers, who obtain better results in mathematical-scientific tests (Martini, 2005). Studies also show that students' nationality on academic achievement has an effect independent of status; in fact, the difficulties experienced by immigrant students depend not only on the scarce family endowment, but above all on linguistic and social integration difficulties that inevitably hinder their performance (Schnepf, 2007). As expected, being late in one's studies is also a risk factor that is associated with poor basic skills (Bertozzi, 2015). Finally, in socio-educational research, the effect of the learning context is considerable, both about the characteristics of the territory in which the school operates, and to the characteristics of the school and of the class group (peer effect), in terms of social-cultural background, skills and motivation (Martini, 2010). The effect of the context is expressed both through the indirect role of teacher quality and teaching processes, and through the direct influence of peers, who represent a model to follow and influence the values and behaviors of students, as a result of competition (Conti *et al.*, 2013), and social interactions (Thrupp, 1999).

Argentin *et al.* (2017) highlighted the existence of a growing level of segregation from elementary schools to high schools, particularly for teachers of Italian. This segregation refers to the systematic association of teacher quality indicators and the student's social level. Along the same lines, a subsequent work by Abbiati and Argentin (2019) highlighted "influences on average of modest entity" of the ascribed characteristics of teachers (gender, age, social origins) on the learning of their students.

As the variance of the results in the INVALSI tests related to the differences between schools and classes (between variance) is greater in Southern Italy (Martini, 2020a), studying the role of context is particularly important in analyzing the performance of students in the Southern regions, where a process of segregation is observed. The distribution of pupils in schools and classes, indeed, appears more unbalanced with respect to individual characteristics that affect learning, as pupils from socio-culturally advantaged or vulnerable backgrounds are concentrated in different schools and classes.

Examining the role of all these factors, in combination with the role of student's status, is therefore very important. From a theoretical-interpretative point of view, this approach allows to identify the best set of explanatory variables that best predict students' performance, while from a planning view, it can suggest the identification of targeted territorial policies capable of reducing the gap deriving from gender and family background, through strategies capable of promoting equity in education systems.

2.5. Social stratification, territorial differences, and educational vulnerability

Focusing on the issue of territorial gaps in students' skills that emerge from the INVALSI national surveys on learning, the literature debate and socio-educational research have focused above all on the relationship between stratification, education, and social vulnerability. Looking at the phenomenon of marginalization through the indicators considered by Save the Children (2020) to map the educational risk in Italy, the most critical levels are highlighted in the South, both in the implicit school dropout rates of 14-16-year-olds who do not reach the minimum skills in Italian and Mathematics, and in the percentages of Early school leavers (explicit school dropout of 18-24 year olds) and NEET (young 15-24 year olds not employed and not in education and training).

In line with the numerous studies that have highlighted territorial gaps in school success (Braga and Checchi, 2010; Falzetti, 2019; Argentin *et al.*, 2017; Murrau and Scicchitano, 2008; Donno *et al.*, 2020), data from the INVALSI National Report (INVALSI, 2019) show that the education system in the South would be not only less effective, but also less equitable. In the South of Italy, in fact, the percentages of pupils with low socio-economic status who do not reach adequate levels in the assessment tests of learning are higher than in the North, suggesting that the national education system appears inadequate in ensuring pupils the same educational opportunities. This lack of equity, since the schools in the South operate in a poorer socio-economic context (Checchi and Peragine, 2005; Braga and Checchi, 2010) where lower socio-economic and cultural status levels in the adult population are observed (Martini, 2020a), tends to amplify in the higher grades of education (Falzetti, 2019). Furthermore, there are significant differences in learning outcomes, not only in the five macro-areas into which our country is divided, but also among the different regions of these macro-areas.

In this sense, in accordance with some recent contributions that have examined English language skills (Bagnarol and Donno, 2020; Russo *et al.*, 2020), we believe it is interesting to investigate the relationship between social stratification and territorial gaps, focusing attention on the differences between specific territories, to explore regional peculiarities that can highlight strengths or weaknesses that can be used for planning evidence-based school policies.

3. The research question: regional differences in the relationship between social inequalities and learning

Little is known about regional differences in the relationship between social inequalities and learning. A first work in this direction was carried out by Martini (2020b) as far as the Veneto region is concerned. In this research, the effect of the socio-economic-cultural status index (ESCS) on the results of the INVALSI tests was investigated, through various regression analyzes, showing a reduction in the effect of the ESCS when other variables were gradually considered.

In the logic of this framework, this contribution wants:

- to investigate the influence of the ESCS, intended as a proxy measure of the student's social, cultural and economic capital, on school results in the INVALSI tests of Italian and Mathematics for lower secondary school (grade 8) in two Southern regions, Puglia and Abruzzo;
- to examine how this influence changes as a function of gender and origin of the student, delay in schooling, and average class ESCS.

The choice to compare the two specific regions of Southern Italy, Abruzzo and Puglia, is based on the consideration of two orders of factors. First, the two regions have different GDP but the same relationship between GDP and learning outcomes; in particular, with an average score in the INVALSI tests which in both regions is slightly higher than what one might expect based on an estimate based on GDP (Martini, 2020a). Second, the scores of the individual regions tend to follow those of the macro-area to which they belong and, among the regions of the South, both Abruzzo and Puglia do not differ from the national average either for the performances in Italian or in Mathematics recorded to grade 8. Furthermore, given the high discrepancy between regions highlighted by various Italian contributions (Bratti *et al.*, 2007; Agasisti and Vittadini, 2012), focusing the attention on two defined territorial areas allows us to limit the variability of the data.

This work focuses the attention in particular on grade 8 (third grade of lower secondary school) for a number of reasons: a) greater reliability of the data compared to those of primary school as, unlike the paper tests used for primary school, Computer Based Tests (CBT) used in secondary schools are less influenced by the phenomenon of cheating (INVALSI, 2019); b) greater interpretability than those of upper secondary school because, compared to the upper secondary school tests, in lower secondary school, there is the advantage of not having a differentiation by type of school (high schools/"licei", technical and vocational institutes), which is not easy to interpret; c) greater territorial variability compared to primary school data: while in primary school the territorial gap between the geographical areas of the country does not reach statistical significance, in line with what emerges from international surveys during the secondary school, both for Italian and Mathematics tests' results, students from Southern macro-areas show performances significantly below the national average (Martini, 2020a); d) greater salience of the issue of equity vs. territorial disparities at this level of education: since in our country the common trunk of the educational path ends with the middle school, it should guarantee everyone the same quality of education by ensuring a uniform base of fundamental skills to guarantee the equality of educational opportunities (Martini, 2020b). It seems relevant, in this sense, to compare the two regions in terms of equity.

4. Method and data analysis plan

4.1. Sample

This study is based on secondary data, kindly made available by INVALSI. Analyses are based on sample data that have been weighted. The students who participated in the INVALSI grade 8 tests for the year 2019 numbered 572,229 for 29,231 classes. The students in the sample examined are about 30,994. The extraction of the sample took place with a two-stage method: in the first the schools were extracted and in the second the classes (generally, two for each school sampled). We have chosen to analyze the sample data for their higher quality, as they are collected in the presence of an external observer who monitored 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 Italian and Mathematics learning relating to the 2018/2019 school year, i.e., the most recent data available, given that in the year 2019/2020 the tests were not administered, due to the lockdown linked to the pandemic.

The Italian tests examine two types of language skills, among those provided for by the National Indications and the Guidelines: 1) the ability to understand an authentic text, thanks to the competence in identifying salient information, reconstructing the meanings expressed in the text and reflecting on the content and textual form; 2) the ability to reflect on the language, that is, the knowledge of grammar necessary to express oneself correctly. The grade 8 test includes, specifically, a section for reading comprehension of three texts of various kinds, a vocabulary section, and a grammar section.

The Mathematics test measures some fundamental skills among those provided by the National Indications and the Guidelines, in different areas. The questions proposed often start from real world problems and verify the disciplinary knowledge, the ability to solve problems and reflect on processes by arguing the reasons for the choices made. The content domains tested in the grade 8 test are Numbers, Space and figures, Data and predictions, Relations and functions.

Specifically, the research examined the scores of the students' skills in the tests of Italian (WLE ITA) and Mathematics (WLE MAT), respectively obtained through the estimate made according to the model of Rasch (1980), which allows to put on the same continuum the difficulty of the items and the ability to be measured. These scores, expressed on a scale with mean 200 and standard deviation 40, were used as dependent variables in the tested models.

The following factors were considered as independent variables:

– the values of the student's socio-economic and cultural status index (**ESCS-s**), a variable constructed by INVALSI based on three indicators: a) the employment status of the parents; b) the level of education of the 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 equal to zero and a standard deviation of 1 (Campodifiori *et al.*, 2010);
– **gender**: 1 if the student is male, 0 if she is female;
– **regularity in the course of study**: 1 if the student has a delay in the course of study, 0 if not;
– the **origin**: 1 if the student is immigrant, 0 if is Italian;
– the average ESCS of the class (**ESCS-c**).

To ensure the generalizability of the results, the data used were preliminarily weighted with 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 (Falorsi, Falzetti, and Ricci, 2019).

4.3. Models of analysis

Three types of analyses were carried out. A first descriptive analysis has computed the average scores in the INVALSI tests for the four quartiles of the ESCS, comparing the data of Abruzzo and Puglia with each other and with those of Italy. A second analysis compared the percentages of students belonging to each of the four quartiles for the five competence levels in the Italian and Math tests.

A third analysis, based on regression, was applied to investigate the effect of the ESCS on the results of the INVALSI tests, for the two regions (Abruzzo and Puglia) considered separately. For the analyses we used the SPSS statistical software and the ENTER method in multiple regression. The first model involves only one predictor: student-level ESCS (ESCS-s). Other multiple regression analysis were then carried out by adding other independent variables to the ESCS to compare different models (tab. 1). In the second model we added the variable gender (male/female). In the third model, we added the immigrant status variable (Italian/immigrant) to the previous variables. In the fourth model, we added the variable late schooling (regular/late schooling) to the previous variables. In the fifth model, we added the class ESCS variable (ESCS-c) to the above variables.

Tab. 1 – Summary of tested models

Models	*Predictor variables*
Model 1	Only Student ESCS
Model 2	Student ESCS + Gender (female)
Model 3	Student ESCS + Gender + Origin (immigrant)
Model 4	Student ESCS + Gender + Origin + Regularity in studies (late)
Model 5	Student ESCS + Gender + Origin + Regularity in studies + ESCS at the class level

5. Results

5.1. ESCS and test results

The average scores on the Italian and Mathematics tests, comparing the students grouped into quartiles of ESCS are reported in tables 2 and 3, respectively.

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

	ESCS			
	I quartile	*II quartile*	*III quartile*	*IV quartile*
Abruzzo	187	200	204	218
Puglia	182	198	203	216
Italy	181	198	205	218

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

	ESCS			
	I quartile	*II quartile*	*III quartile*	*IV quartile*
Abruzzo	186	200	201	220
Puglia	182	196	203	213
Italy	183	198	206	218

The results show wide performance variations based on the quartile of ESCS for both tests and for both regions. The average performance of students in the first and fourth quartile of ESCS differs by approximately one standard deviation, both for the Italian and Math, in Abruzzo and in Puglia.

5.2. Competency level and ESCS quartiles

The comparison between the percentages of students belonging to each of the four quartiles for the five competence levels in the Italian and Math for the two regions are shown in the tables 4 and 5. To interpret the results, one must consider that the competence levels 1 and 2 represent unsatisfactory learning outcomes, while the levels 4 and 5 represent good or very good results.

As regards the Italian tests (tab. 4), in both the regions, the levels 1 and 2 show an inverse trend compared to levels 3 and 4. The majority of students with a level 1 in Italian (52% for Abruzzo and 59% for Puglia) fall into quartile 1 of ESCS, while only about 10% (11% for Abruzzo and 7-8% for Puglia) fall into quartile 4. This means that an unsatisfactory level of learning in Italian mainly characterizes students with a low ESCS.

Focusing the attention to the students with a high level of competence, for example level 5, the analysis evidence that only 8% of the students belong to first quartile in the ESCS, while 40-50% of them (41% for Abruzzo and 46% for Puglia) belongs to the fourth quartile of the ESCS. This means that the achievement of the highest competence levels in the Italian tests is a prerogative of students with high ESCS.

Tab. 4 – Average percentages of students belonging to each quartile of ESCS, distributed by competence level in Italian and by region

		ESCS (quartiles)				
Italian		*1*	*2*	*3*	*4*	*Total*
Level 1	Abruzzo	52.2	17.4	19.0	11.3	100.0
	Puglia	59.2	19.5	13.7	7.6	100.0
Level 2	Abruzzo	38.2	23.3	26.0	12.4	100.0
	Puglia	42.2	20.6	22.4	14.8	100.0
Level 3	Abruzzo	24.4	22.6	31.0	22.0	100.0
	Puglia	25.9	19.8	30.8	23.4	100.0
Level 4	Abruzzo	19.6	12.7	26.0	41.7	100.0
	Puglia	19.8	18.4	26.5	35.3	100.0
Level 5	Abruzzo	8.3	21.7	29.1	40.9	100.0
	Puglia	8.3	19.3	26.2	46.2	100.0

Tab. 5 – Average percentages of students belonging to each quartile of ESCS, distributed by competence level in Math and by region

		ESCS (quartiles)				
Math		*1*	*2*	*3*	*4*	*Total*
Level 1	Abruzzo	49.4	15.8	26.5	8.3	100.0
	Puglia	54.3	17.3	16.3	12.1	100.0
Level 2	Abruzzo	33.0	26.3	26.6	14.2	100.0
	Puglia	36.4	21.4	25.4	16.7	100.0
Level 3	Abruzzo	29.3	18.2	27.9	24.6	100.0
	Puglia	30.5	20.4	25.3	23.8	100.0
Level 4	Abruzzo	18.3	16.1	27.8	37.8	100.0
	Puglia	13.5	21.3	31.2	34.0	100.0
Level 5	Abruzzo	9.9	20.4	25.6	44.2	100.0
	Puglia	15.7	15.5	28.4	40.4	100.0

As for the Math tests (tab. 5), the results are largely like those seen for the Math tests. In this sense, the diversity in terms of learning opportunities for students with high ESCS compared to those with low ESCS appears as strong both for Mathematics and for Italian in both regions.

5.3. Results of regression models

Tables 6 and 7 illustrate the results of the regression models for Italian test respectively for Abruzzo and Puglia, while Tables 8 and 9 show the results for the math test.

For the Italian tests (tables 6 and 7) the linear regression model shows a relevant effect of ESCS-s, both for Abruzzo (b = 12.6) and for Puglia (b = 13.7). The inclusion of gender and immigrant status variations does not reduce the influence of the ESCS in both regions. The inclusion of the late education variable reduces the influence of ESCS in Abruzzo but not in Puglia. The further inclusion of the ESCS-c variable reduces the influence of ESCS-s in both regions, but more in Puglia.

Tab. 6 – Results of the regression models for the Italian test in Abruzzo

	Model 1 ESCS-s	*Model 2 + Gender*	*Model 3 + Immigrant*	*Model 4 + Late school*	*Model 5 + ESCS-c*
Constant	201.6***	190.2***	196.2***	207.1***	206.7***
ESCS-s	12.6***	12.5***	12.1***	10.6***	9.6***
Female		7.8***	7.8***	7.5***	7.5***
Immigrant			-5.3***	-3.9***	-4.0***
Late schooling				-10.6***	-10.5***
ESCS-c					7.5***
R-squared	0.124	0.137	0.138	0.155	0.160

Tab. 7 – Results of the regression models for the Italian test in Puglia

	Model 1 ESCS-s	*Model 2 + Gender*	*Model 3 + Immigrant*	*Model 4 + Late school*	*Model 5 + ESCS-c*
Constant	199.7***	186.6***	197.9***	209.0***	208.7***
ESCS-s	13.7***	13.6***	13.4***	13.0***	11.2***
Female		8.8***	8.5***	8.2***	8.1***
Immigrant			-10.6***	-8.3***	-8.9***
Late schooling				-12.6***	-11.3***
ESCS-c					8.5***
R-squared	0.154	0.169	0.171	0.174	0.184

For the Mathematical test, the linear regression model shows a relevant effect of the ESCS-s, both for Abruzzo (b = 13.4) and for Puglia (b = 12.6). The inclusion of gender and immigrant status variations does not reduce the influence of the ESCS in both regions. The inclusion of the late education variable reduces the influence of ESCS in Abruzzo but not in Puglia. The addition of the ESCS-c variable further reduces the influence of ESCS-s in both regions.

Tab. 8 – Results of the regression models for the Math test in Abruzzo

	Model 1 ESCS-s	*Model 2 + Gender*	*Model 3 + Immigrant*	*Model 4 + Late school*	*Model 5 + ESCS-c*
Constant	201.0***	208.6***	215.9***	232.3***	231.6***
ESCS-s	13.4***	13.4***	13.2***	12.0***	10.5***
Female		-5.2***	-5.7***	-6.0***	-5.9***
Immigrant			-6.2***	-3.8***	-3.9***
Late schooling				-17.6***	-17.3***
ESCS-c					10.3***
R-squared	0.132	0.138	0.140	0.151	0.161

Tab. 9 – Results of the regression models for the Math test in Puglia

	Model 1 ESCS-s	*Model 2 + Gender*	*Model 3 + Immigrant*	*Model 4 + Late school*	*Model 5 + ESCS-c*
Constant	198.5***	206.2***	214.1***	228.3***	228.0***
ESCS-s	12.6***	12.7***	12.5***	12.0***	10.3***
Female		-5.2***	-5.4***	-5.8***	-5.9***
Immigrant			-7.3***	-4.3***	-4.7***
Late schooling				-16.2***	-15.1***
ESCS-c					7.8***
R-squared	0.118	0.123	0.124	0.128	0.136

6. Conclusions

This study stems from the recognition that socio-economic cultural status is the most significant factor differentiating performance on INVALSI tests. However, its complexity makes it challenging to fully understand and identify appropriate interventions to mitigate its impact. At the same time, the variety of regional situations should be taken into consideration when choosing policy interventions that can be truly effective.

The INVALSI score in the ESCS quartiles shows large differences (one standard deviation) between the first and the fourth quartile in all groups; both in Abruzzo and in Puglia. The analysis of the distribution of ESCS quartiles according to skill levels reveals that it is completely inverted when comparing levels 1 and 2 on one side with levels 3 and 4 on the other. This applies to both Italian and Mathematical CBT test scores.

Our analysis confirm the strong impact of cultural socio-economic status, for both tests and both regions; even if the problem seems to have a greater incidence in Puglia than in Abruzzo, in the sense of a stronger performance gap based on socio-economic-cultural status. This difference could be due to GDP, since Puglia has a lower GDP than Abruzzo.

Regarding the complex nature of the influence that socio-economic-cultural status has on the results of the INVALSI tests, regression analysis was carried out to investigate if the weight of the ESCS variable on the results of the INVALSI tests is reduced when other variables are entered in the multiple regression model. For the Italian tests, the comparison between the two regions shows the importance in Abruzzo of intervention paths focused on the problem of late schooling and in Puglia on the segregation and composition of classes. For the regression analyzes on mathematical tests, the results show the importance in Abruzzo of intervention paths focused on the problem of late schooling and in Puglia on the segregation and composition of classes.

Overall, these results can be a first step to promote the use of INVALSI data for the definition of effective regional policies aimed at reducing the impact of socio-economic-cultural variables on students' learning outcomes. INVALSI data, indeed, provide in-depth information on the factors related to students' achievement and could be useful to policymakers to inform their decisions both system-wide and at an individual school level.

However, despite Italy's extensive participation in most international and national evaluations, the impact of these data on educational national and regional policies' planning is weak and often inconsistent (Damiani, 2016). Even though Italy ranks among the most unequal countries in Europe (Barone and Ruggera, 2018), reducing inequalities in education has remained a peripheral goal in educational policies (Pensiero, Giancola and Barone, 2019). To decrease the impact of ESCS on learning processes and outcomes, especially in Southern regions where there is a higher incidence of low-SES families and a higher risk of student dropout, several system-level structural actions could be implemented. First, educational policies must be constructed and implemented together with social and economic policies, including the experimentation of innovative infrastructure and resources for informal learning. Second, teachers' pre-service and in-service training programs could be ensured, to provide lower secondary school teachers some innovative teaching methods to stimulate students with lower ESCS motivation and engagement. Finally, both the number of hours students spend in school and extra-school educational activities could be increased, as well as the development of extracurricular projects realized by experts and the involvement of parents in school activities.

We conclude with the hope that in the future, INVALSI will be able to gather a significant amount of data on non-cognitive skills and students' motivational profile, which could contribute to a deeper understanding of the processes through which socio-economic factors have an impact, and thus, enable the formulation of appropriate intervention policies.

References

Abbiati G., Argentin G. (2019), "Caratteristiche sociodemografiche degli insegnanti italiani e loro influenze sugli apprendimenti degli studenti", in P. Falzetti (a cura di), *Uno sguardo alla scuola. II Seminario "I dati INVALSI: uno strumento per la ricerca"*, FrancoAngeli, Milano.

Agasisti T., Vittadini G. (2012), "Regional economic disparities as determinants of student's achievement in Italy", *Research in Applied Economics*, 4 (1), pp. 33-54.

Argentin G., Abbiati G., Gerosa T. (2017), "Non proprio la stessa scuola. Segregazione degli insegnanti tra scuole e abbinamento insegnanti-studenti come meccanismi nascosti di disuguaglianza nel sistema scolastico italiano", in P. Falzetti (a cura di), *I dati INVALSI: uno strumento per la ricerca*, FrancoAngeli, Milano.

Bagnarol C., Donno S. (2020), *Analisi spaziale degli apprendimenti scolastici (INVALSI) in inglese: un confronto tra le regioni del Nord d'Italia*, Paper presented at the XIII Conferenza Espanet, 17-19 settembre, Venezia.

Ballarino G., Checchi D. (2006), *Sistema scolastico e disuguaglianza sociale: scelte individuali e vincoli strutturali*, il Mulino, Bologna.

Barone C., Ruggera L. (2018), "Educational equalization stalled? Trends in inequality of educational opportunity between 1930 and 1980 across 26 European nations", *European Societies*, 20 (1), pp. 1-25.

Bernstein B. (1971), *Class, codes and control*, Routtedge and Kegan Paul, London, vol. 1.

Bertozzi R. (2015), "Le traiettorie scolastiche degli alunni con cittadinanza non italiana tra svantaggio e successo", in M. Santagati, V. Ongini (a cura di), *Alunni con cittadinanza non italiana. Tra difficoltà e successi*, Rapporto nazionale a.s. 2013/14, pp. 57-80, Fondazione ISMU, Milano.

Berzano L., Genova C. (2011), *Sociologia dei lifestyles*, Carocci, Roma.

Besozzi E. (2017), *Società, cultura, educazione*, Carocci, Roma.

Bottero W. (2004), "Class identities and the identity of class", *Sociology*, 38 (5), pp. 985-1003.

Boudon R. (1979), *Istruzione e mobilità sociale*, Zanichelli, Bologna.

Bourdieu P. (2003), "Culture reproduction and social reproduction", in C. Jenks (*ed.*), *Culture. Critical concepts in sociology*, Routledge Press, New York, vol. 3, pp. 63-99.

Bourdieu P. (2015), *Forme di capitale*, Armando, Roma.

Bourdieu P., Passeron J.C. (1971), *I delfini. Gli studenti e la cultura*, Guaraldi, Rimini-Firenze.

Braga M., Checchi D. (2010), "Sistemi scolastici regionali e capacità di sviluppo delle competenze. I divari delle indagini PIRLS e PISA", *Italian Journal of Social Policy*, 3, pp. 1-25.

Bratti M., Checchi D., Filippin A. (2007), "Territorial differences in Italian students' mathematical competencies: evidence from PISA 2003", *IZA DP No. 2603*.

Bronfenbrenner U. (1979), *Ecologia dello sviluppo umano*, il Mulino, Bologna, 1986.

Butler Y. G. (2014), "Parental factors and early English education as a foreign language: A case study in Mainland China", *Research Papers in Education*, 29 (4), pp. 410-437.

Campodifiori E., Figura E., Papini M., Ricci R. (2010), *Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia*, Working Paper, 2/2010, Frascati (RM), INVALSI, retrieved on June 15, 2021, from http://www.invalsi.it/download/wp/wp02_Ricci.pdf.

Cappellari L. (2006), "Background familiare, scelte formative e transizione scuola università", in G. Ballarino, D. Checchi (a cura di), *Sistema scolastico e disuguaglianza sociale*, il Mulino, Bologna, pp. 57-90.

Checchi D. (2010), "Percorsi scolastici e origini sociali nella scuola italiana", *Politica economica*, 26 (3), pp. 359-388.

Checchi D., Peragine V. (2005), "Regional disparities and inequality of opportunity: the case of Italy", *IZA DP No. 187*4.

Conti E., Duranti S., Maitino M.L., Rampichini C., Sciclone N. (2013), "The future has early roots. Learning outcomes and school's effectiveness in Tuscany's primary education system", in *Workshop Youth and their future: Work", Education and Health, October 17th-18th*.

Damiani V. (2016), "Large-scale assessments and educational policies in Italy", *Research Papers in Education*, 31 (5), pp. 529-541.

Devine F., Savage M. (2000), "Conclusion: Renewing class analysis", in R. Crompton, F. Devine, M. Savage, J. Scott (*eds.*), *Renewing class analysis*, Blackwell, Oxford, pp. 99-184.

Donno S., Bagnarol C., Marsili M. (2020), *Analisi spaziale degli apprendimenti scolastici negli istituti del Sud Italia*, Working Papers INVALSI, n. 46.

Easterbrook M.J., Hadden I.R. (2021), "Tackling educational inequalities with social psychology: Identities, contexts, and interventions", *Social Issues and Policy Review*, 15 (1), pp.180-236.

Falorsi P.D., Falzetti P., Ricci R. (2019), Le metodologie di campionamento e scomposizione della devianza nelle rilevazioni nazionali dell'INVALSI, FrancoAngeli, Milano.

Falzetti P. (2019), "Un sistema scolastico in difficoltà e un preoccupante dualismo territoriale: i risultati delle prove INVALSI", *Social Policies*, 6 (3), pp. 527-532.

Fiorio C., Leonardi M. (2010), "Mobilità intergenerazionale nell'istruzione in Italia", in D. Checchi (a cura di), *Immobilità diffusa*, il Mulino, Bologna, pp. 53-85.

Fitzpatrick K.M., Yoels W.C. (1992), "Policy, school structure, and sociodemographic effects on statewide high school dropout rates", *Sociology of Education*, 65 (1), pp. 76-93.

Gottfried M.A. (2014), "Peer effects in urban schools: assessing the impact of classroom composition on student achievement", *Educational Policy*, 28 (5), pp. 607-647.

Grapin S.L., Shriberg D. (2020), "International perspectives on social justice: Introduction to the special issue", *School Psychology International*, 41 (1), pp. 3-12.

INVALSI (2019), *Rapporto nazionale. Rapporto prove INVALSI 2019*, retrieved on June 15, 2021, from https://invalsi-areaprove.cineca.it/docs/2019/Rapporto_prove_INVALSI_2019.pdf.

Lareau A. (2011), *Unequal childhoods: Class, race, and family life*, University of California Press, Berkeley.

Martini A. (2005), "Differenze di genere e organizzazione della scuola secondaria", in C. Marangon, M.T. Siniscalco (a cura di), *Gli studenti quindicenni nel Veneto: quali competenze? Rapporto regionale del Veneto OCSE-PISA 2003*, USR per il Veneto-DG, pp. 177-208.

Martini A. (2010), "Misurare il contesto delle scuole", in Ass. Context (a cura di), *Valutare apprendimenti, valutare contesti*, Edizioni Infantiae.Org, Roma, pp. 137-144.

Martini A. (2020a), *Il divario Nord-Sud nei risultati delle prove INVALSI*, Working Paper N. 52/2020; Working Papers INVALSI, pp. 1-16, retrieved on June 15, 2021, from https://www.invalsi.it/download2/wp/wp52_Martini.pdf.

Martini A. (2020b), "Istruzione e diseguaglianze sociali nel Veneto", *Economia e società regionale*, 38 (1), pp. 60-78.

Murrau L., Scicchitano S. (2008), "Education in Southern Italy: facts and politics", *Ponte*, 64 (10), pp. 25-34.

Naidoo A.V., Van Eeden A., Munch Z. (2014), "Spatial variation in school performance, a local analysis of socio-economic factors in Cape Town", *South African Journal of Geomatics*, 3 (1), pp. 78-94.

Nastasi B.K., Naser S.C. (2020), "Conceptual Foundations for School Psychology and Child Rights Advocacy", in B.K. Nastasi, S.N. Hart, S.C. Naser, *International Handbook on Child Rights and School Psychology*, Springer Nature Switzerland, Cham.

Newell M., Kaundinya P., Gaona M. (2021), "Theoretical Foundations of Diversity in School Psychology", in K.K. Kelly, S.A. Garbacz, C.A.T. Albers (*eds.*), *Theories of School Psychology. Critical Perspectives*, Routledge (Taylor and Francis Group), New York-London.

Oseguera L., Conchas G.Q., Mosqueda E. (2011), "Beyond family and ethnic culture: Understanding the preconditions for the potential realization of social capital", *Youth & Society*, 43 (3), pp. 1136-1166.

Pensiero N., Giancola O., Barone C. (2019) "Socioeconomic inequality and student outcomes in Italy", *Socioeconomic Inequality and Student Outcomes*, Springer, Singapore, pp. 81-94.

Proctor S.L. (2018), "Multicultural Foundations", in S.L. Grapin, J.H. Kranzler, (*eds.*), *School Psychology: Professional Issues and Practices*, Springer, New York, pp. 61-76.

Raimondi E., De Luca S., Barone C. (2013), "Origini sociali, risorse culturali familiari e apprendimenti nelle scuole primarie: un'analisi dei dati PIRLS 2006", *Quaderni di Sociologia*, 61, pp. 34-49.

Rasch G. (1980), *Probabilistic models for some intelligence and attainment tests*, University Press, Chicago.

Ream R.K., Rumberger R.W. (2008), "Student engagement, peer social capital, and school dropout among Mexican American and non-Latino white students", *Sociology of education*, 81 (2), pp. 109-139.

Russo M.C., Nerli Ballati E., Mancini L. (2020), *Tra disuguaglianze sociali e divari territoriali: gli studenti italiani di III secondaria di I grado alla prova INVALSI di Inglese*, paper presented at the XIII Conferenza Espanet, 17-19 settembre, Venezia.

Save the Children (2020), *Nuotare contro corrente. Povertà educativa e resilienza in Italia*, retrieved on June 15, 2021, from https://www.savethechildren.it/sites/default/files/files/uploads/pubblicazioni/nuotare-contro-corrente-poverta-educativa-e-resilienza-italia.pdf.

Schizzerotto A., Barone C. (2006), Sociologia dell'istruzione, il Mulino, Bologna.

Schnepf S.V. (2007), "Immigrants' educational disadvantage: an examination across ten countries and three surveys", *Journal of Population Economics*, 20 (3), pp. 527-545.

Shriberg D., Bonner M., Sarr B.J., Walker A.M., Hyland M., Chester C. (2008), "Social justice through a school psychology lens: Definition and applications", *School Psychology Review*, 37 (4), pp. 453-468.

Shriberg D., Brooks K., Montes de Oca J. (2020), "Child Rights, Social Justice, and Professional Ethics", in B.K. Nastasi, S.N. Hart, S.C. Naser (*eds.*), *International Handbook on Child Rights and School Psychology*, Springer Nature Switzerland, Cham.

Suh S., Suh J., Houston I. (2007), "Predictors of categorical at-risk high school dropouts", *Journal of Counseling & Development*, 85 (2), pp. 196-203.

Thrupp M. (1995), "The school mix effect: the history of an enduring problem in educational research, policy and practice", *British Journal of Sociology of Education*, 16 (2), pp. 183-203.

Thrupp M. (1999), *Schools making a difference. Let's be realistic!*, Open University Press, Buckingham-Philadelphia.

Vallejo C., Dooly M. (2013), "Early School Leavers and Social Disadvantage in Spain: from books to bricks and vice-versa", *European Journal of Education*, 48 (3), pp. 390-404.

5. Digital and informative skills: the differences between students and the role of the school

by Ornella Papa, Rita Marzoli, Sara Manganelli

According to the international survey IEA ICILS 2018 on eighth-grade students, digital and informative skills are not widespread among Italian students and are not adequately developed in schools. ICILS results reveal that most students, although native digital, do not have sophisticated digital skills; the presence of computer equipment in schools doesn't seem to improve these skills if they are not adequately integrated into teaching. The present in-depth study, based on ICILS 2018 Italian data, analyses the students' achievement on Computer information literacy in relation to students' characteristics (gender, migratory origin, and socio-economic-cultural background) and school activities that contribute to developing students' digital and informative skills. In line with international data, Italian female students perform better than their male peers. First-generation immigrant students, like the other students who don't speak Italian at home, perform worse than other students. The major differences are associated to the socio-economic and cultural background of the students; in this context, the presence of learning resources at home, including the number of computers and internet access, also appears to be relevant. However, the students achieve better if they learn and frequently practice ICT tasks at school. These results indicate that the schools have a pivotal role in fostering the development of CIL and in overcoming the digital divide. A contribution to achieving these aims could come from Innovative school libraries, as recognized under Action#24 of the current PNSD (School reform, Low 107/2015).

Le competenze digitali e informative sono poco diffuse tra gli studenti italiani e non adeguatamente sviluppate in ambito scolastico, secondo quanto emerge dall'indagine internazionale IEA ICILS 2018, condotta su studenti all'ottavo grado di scolarità. I risultati ICILS rivelano che gli studenti, pur *essendo nativi digitali, non posseggono competenze digitali sofisticate; la presenza nelle scuole di attrezzature informatiche non è sufficiente a migliorare queste competenze se non vengono adeguatamente integrate nella didattica. Questo approfondimento, condotto sui dati italiani di ICILS 2018, analizza le differenze nelle variabili di outcome, costituite dai* plausibile values *in* Computer information literacy*, in base ad alcune caratteristiche degli studenti. In linea con i dati internazionali le studentesse italiane conseguono risultati migliori dei loro coetanei di genere maschile. Gli studenti immigrati di prima generazione, come gli altri studenti che non parlano a casa la lingua della prova, conseguono risultati peggiori degli altri studenti. Le maggiori differenze sono associate al background socioeconomico e culturale degli studenti; in questo ambito appaiono rilevanti anche la presenza a casa di risorse per l'apprendimento, tra cui il numero di computer e l'accesso ad internet. Tuttavia, i punteggi degli studenti sono migliori se in ambito scolastico hanno appreso e praticano, con frequenza, le competenze digitali e informative. Questi risultati sostengono il ruolo della scuola nello sviluppo di tali competenze e nel superamento del divario digitale. Un contributo per il raggiungimento di questi obiettivi potrebbe giungere dalle Biblioteche Innovative, come previsto dall'Azione#24 dell'attuale PNSD (Riforma della scuola, legge 107/2015).*

1. Introduction

During COVID-19 pandemic period, digital skills have been crucial in the knowledge society. The aim of this paper is to understand with which competences students went through the school changes in this portion of time, with an in-depth analysis regarding Italian students. In Italy, the switch over to Distance Learning determined that digital skills were prerequisite for teaching and learning, without being previously developed. In fact, Italian students were particularly lacking in digital and informative skills according to ICILS[1] 2018 results (Fraillon *et al.*, 2019b). This survey is carried out every 5 years by IEA[2] in collaboration with research national institutes of participating countries, for Italy INVALSI[3]. In Section 2 of this chapter, we give a general overview of the ICILS framework, a detailed description of

[1] International Computer and Information Literacy Study.

[2] International Association for the Evaluation of Educational Achievement.

[3] National Institute for the Evaluation of the Educational System of Education and Training, involved together to other research institutes from participating countries.

the 2018 cycle features and key findings, in addition to an excursus of the existing literature on ICILs data. The overall ICILS conclusion was that digital native students don't have sophisticated digital skills unless they learn and practice ICT[4] tasks in a favourable learning environment. The presence at school of computer equipment is, of course, necessary but not sufficient to develop digital and informative skills, in the dearth of teaching and practice of ICT tasks. The critical results in CIL[5] can be interpreted in relation to these disclosures. At international level 70% of students claimed to use digital devices every day for general purposes but only 20% of students claimed to use them for educational purposes, unfortunately in Italy this percentage is even lower. It seemed useful to deepen and disseminate Italian results in ICILS 2018, just in view of an expected broader national digitalization school plan. In section 3 of this chapter, we present the focus of this paper which is an in-depth analysis on Italian ICILS results by student characteristics and ICT activities. The variety of ICILS data allowed us to analyse Italian students' CIL[5] scores in relation to several characteristics and ICT activities at home and at school, to understand the nature and extent of the digital divide as well as the role of school in reducing it. As already confirmed in the Italian context, the improvement of digital skills has a positive influence on overcoming inequalities by raising academic performance of students from a lower socio-economic background (Pagani *et al.*, 2016). The integration in school of CIL could be implemented by adequate actions, training for teachers, learning resources and programmes. These improvement actions cannot ignore the historical link between Information Literacy and School Library. Authoritative documents, such as IFLA[6]/UNESCO[7] "School Libraries Manifesto" (1999) highlighted the role of school libraries to realize "Media and information skills Programs". In Italy, the current PNSD[8] proclaimed the "Innovative School Library" (Action#24) as a strategic environment to overcome every educational gap including the digital divide. In section 4 of this chapter, we give a global overview of international studies and important insights on this topic, particularly to inform policy and suggesting the inclusion of Innovative School Libraries in every school.

[4] Information and Communication Technology.
[5] Computer and Information Literacy.
[6] International Federation of Library Associations and Institutions.
[7] United Nations Educational, Scientific and Cultural Organization.
[8] National Digital School Plan (Law 107/2015), https://www.istruzione.it/scuola_digitale/allegati/Materiali/pnsd-layout-30.10-WEB.pdf.

2. IEA ICILS

Since 2013 the IEA has been promoting ICILS, a computer-based assessment of eighth-grade students focused on Computer and Information Literacy; the cycles are repeated every five years. As an international survey, ICILS allows comparison between different educational systems in addition to investigate the level of digital and informative skills within countries. The basic aim of ICILS is to provide information and enable understanding of ICT domain to promote the development of school programs suitable for the digital age. The definition of CIL is as follows: "an individual's ability to use computers to investigate, create and communicate in order to participate effectively at home, school, workplace and society" (Fraillon *et al.*, 2013, p. 17). In addition, ICILS detects factors that can influence CIL development; students' characteristic as well as the main learning support contexts are carefully explored by questionnaires to understand how they relate to CIL improvement. ICILS 2013 was attended by 21 educational systems from around the world (Fraillon *et al.*, 2014) while a reduced number of only 13 educational systems participated in ICILS 2018, including Italy which did not participated in the precedent cycle. For countries involved in both cycles, it was possible to make a longitudinal comparison of the results to monitor changes between the two assessments in teaching and learning contexts as well as in their students' scores.

2.1. ICILS 2018

ICILS 2018 involved 12 countries and 2 benchmarking entities[9] from 4 continents, in total 46,651 students and 26,530 teachers from 2,226 schools were assessed. The schools were sampled using the Probability Proportional to Size method, then 20 students and 15 teachers were randomly extracted from the target grade of each sampled school. The stated purpose of ICILS 2018 is "to assess systematically the capacities of students to use ICT productively for a range of different purposes, in ways that go beyond a basic use of ICT" (Fraillon *et al.*, 2019a, p. 1). The CIL construct is articulated in four strands that frame the skills and knowledge: understanding computer use, gathering information, producing information and digital communi-

[9] Territorial realities – provinces or regions – that took part in the survey for their own internal comparison objective, but whose data are not considered in the calculation of the international average.

cation (Fraillon *et al.*, 2019a). The last ICILS cycle included Computation Thinking (CT), an optional component assessed in 9 of the 13 education systems involved; Italy participated only in the main component (CIL) on which the present study is focused.

In addition to the computer-based test on CIL, consisting of two modules, the administration provided a variety of questionnaires that collect a rich amount of information:

- *Questionnaire on National Context* – to describe the structure of the education system, the teaching of digital and informative skills, initiatives and resources associated with ICT;
- *School Principal Questionnaire* – to observe the general characteristics of the school, the guidelines relating to the use of ICT and their perceived value;
- *Teachers Questionnaire* – to explore their experiences, the difficulties and the propensity to use the computer in teaching;
- *Coordinator/Digital Animator Questionnaire* – to evaluate the ICT infrastructure and the assistance available to teachers;
- *Student Questionnaire*, which includes questions about the socio-economic backgrounds, the use of information technology and the attitude to computer use.

From the overall results obtained, the IEA concluded that:

- digital native students do not have sophisticated digital skills;
- providing schools with computer equipment is not enough to improve students' digital and informative skills;
- variations in student scores are greater within national boundaries than between nations;
- there is a gap related to students' socio-economic backgrounds; students from unfavourable backgrounds have significantly worst scores (on average < 30).

In tab. 1 we can see CIL international average and CIL average scores of participating countries and benchmarking entities.

It should be noted that the administration in Italy was carried out at the beginning of the school year, unlike in the other participating countries where the administration took place in the second part of the school year.

Both the lowest average age and school attendance period of Italian students slightly disadvantaged them for ICILS results.

While for the United States, whose ICILS sample was not considered entirely adequate, the issue is different concerning not only comparison with other countries, but also internal data analysis.

Tab. 1 – ICILS 2018 average scores

Country	*CIL average score*	
Denmark	553	(2.0)
Moscow (benchmarking entity in Russian Federation)	549	(2.2)
Korea, Republic of	542	(3.1)
Finland	531	(3.0)
United States (Not meeting sample participation requirements)	519	(1.9)
Germany	518	(2.9)
North Rhine-Westphalia (benchmarking entity in Germany)	515	(2.6)
Portugal	516	(2.6)
France	499	(2.3)
Luxembourg	482	(0.8)
Chile	476	(3.7)
Italy (Testing at the beginning of school year)	461	(2.8)
Uruguay	450	(4.3)
Kazakhstan	395	(5.4)
International CIL average	496	(1.0)

Source: IEA ICILS 2018 International Report (Fraillon *et al.*, 2019b)

As shown in Tab. 1, Denmark ranked 1st with an average score of 553, followed by Moscow (549), Korea (542) and Finland (531). Among the other EU countries, also Germany (518) and Portugal (516) reached average scores significantly higher than the international average, while France (499) did not deviate significantly from the international average. Luxembourg reached an average score (482) lower than the international average but better than the Italian score (461). Italy is followed in the total ranking only by Uruguay (450) and Kazakhstan (395).

Beyond the single national averages, the analyses showed a general problematic situation (Fraillon *et al.*, 2019b) as we can see in detail from the international percentages of students who reached the different skill levels:

- *level 4*: 2% of students were able to perform checks and make judgments while searching for information and while creating information products;
- *level 3*: 19% of students were able to work independently in search of information and with computer management tools;
- *level 2*: 36% of students were able, following the instructions, to use the computer for the simple collection of information and for the use of basic management tools;
- *level 1*: 25% of students had only a functional knowledge of using the computer as a tool;

– *under minimum level*: 18% of students were unable to create digital information products and didn’t have functional knowledge of using the computer as a tool.

So, at an international level only 2% of students were capable to “critically evaluate the information” (level 4) and 21% of students were “independent computer users” (level 3 + level 4). The level 2 of competence – which can be considered the sufficiency threshold – was reached by 36% of students, while 43% of students stopped under level 2, placing themselves at the lower skill levels; they can be referred to as under-performing students because they failed to understand and perform even the most basic ICT operations. The percentages of under-performing students in EU participating countries are shown in fig. 1.

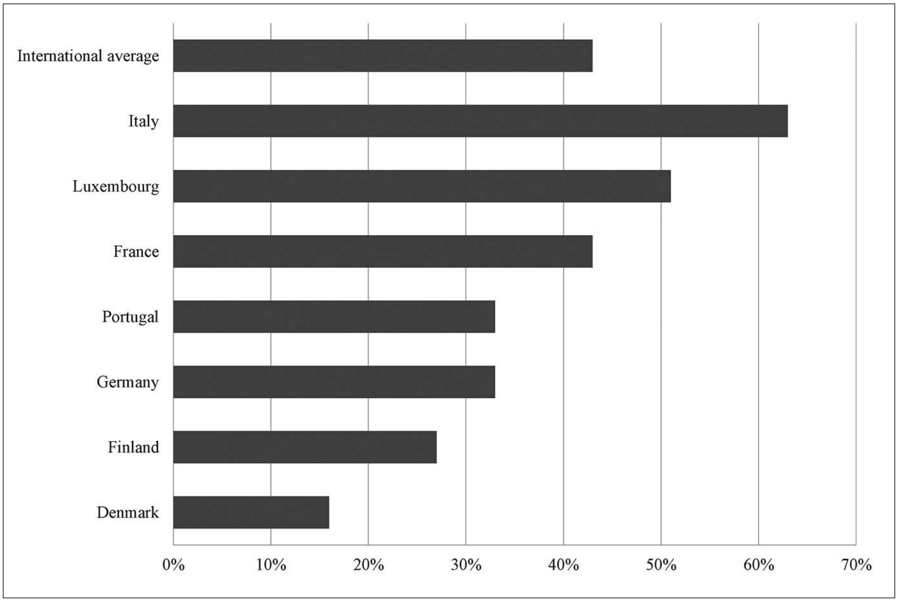

Fig. 1 – Percentages of under-performing students in EU countries

Source: IEA 2019

While wanting to go beyond the logic of ranking, we must take note that Italian students had more difficulties in CIL than students from all participating EU countries: in Italy the percentage of under-performing students was more than 60%. However, except for Denmark, where the percentage of under-performing students was low (17%) and Finland where it was moderate (28%), in other EU countries the proportion of under-performing students ranged from 1 in 3 students to 1 in 2 students. The European Commission, taking into account ICILS data and considering them to be critical, introduced as an EU target for students' digital competence *to have less than 15% of under-performing students in CIL by 2030*. This target was part of the new "Digital Education Action Plan (2021-2027): Resetting Education and Training for the Digital Age". European Commission also dedicated "A focus on digital education" in "Education and Training Monitor 2020". Therefore, we felt it was necessary to disseminate and deepen this data to promote the educational changes that effective school digitalization requires.

Many in-depth studies have been published following the release of ICILS data, to testify the great interest of the whole research community. In the next subsection we report an excursus of in-depth studies on ICILS data, to give an idea of topics and countries involved, also to better contextualize our study. However, as shown in the following subsection, studies on ICILS 2018 are still very few.

2.2. ICILS 2013 and 2018: an excursus of in-depth studies

The aim of this subsection is to draw a map of ICILS data usage and to give a comprehensive overview of the current literature: most discussed topics and countries involved. Due to the heterogeneity of the topics covered, discussing the outcomes of each study would lead us far from the focus of this chapter. The main purpose of this bibliography is:
- to give a state-of-the-art report of the research in this area;
- to show trends and gaps in the research;
- to identify previous studies conducted at national level.

For this literature search the following bibliographic databases were used:
- ERIC (Education Resources Information Centre);
- SCOPUS Elsevier;
- BEI (British Educational Index).

Given that the subject of this research focuses exclusively on studies and research that have ICILS data as the primary descriptor, the search strategy was very focused. Therefore, the search, for all databases, is built using the keyword "ICILS" to identify records. Controlled terms (thesaurus) and exact descriptors are not the focus in this search strategy.

The tables below show the results of the bibliographic search of studies on ICILS 2013 data (tabs. 2-4) and on ICILS 2018 data (tab. 5), except for the IEA ICILS International Reports. The studies on the first cycle are more numerous so they will be further divided by the countries involved.

Tab. 2 – ICILS 2013 in depth studies on all participating countries

Author	*Title*	*Source*
Alkan and Meinck (2016)	The relationship between students' use of ICT for social communication and their computer and information literacy	*Large-Scale Assessments in Education*, 4, p. 15, https://doi.org/10.1186/s40536-016-0029-z
Bundsgaard and Gerick (2017)	Patterns of students' computer use and relations to their computer and information literacy: Results of a latent class analysis and implications for teaching and learning	*Large-Scale Assessments in Education*, 5, p. 16, https://doi.org/10.1186/s40536-017-0052-8
Bundsgaard (2019)	DIF as a pedagogical tool: Analysis of item characteristics in ICILS to understand what students are struggling with	*Large-Scale Assessments in Education*, 7, p. 9, https://doi.org/10.1186/s40536-019-0077-2
Ercikan, Asil, Grover (2018)	Digital divide: A critical context for digitally based assessments	*Education Policy Analysis Archives*, 26 (51), p. 51, https://doi.org/10.14507/epaa.26.3817
Gebhardt, Thomson, Ainley, Hillman, IEA (2019)	Gender differences in computer and information literacy: An in-depth analysis of data from ICILS	*IEA Research for Education*, Volume 8
Ozbasi, Ilgaz (2019)	An examination of educational inputs with the data envelopment analysis: The example of ICILS 2013	*Educational Policy Analysis and Strategic Research*, 14 (3), pp. 129-153

The studies in tab. 2 explore relevant issues on ICILS 2013 international data: the digital divide, gender differences in CIL, the relation of CIL scores with student's patterns of computer use as well as with their ICT use for social communication. An in-depth study focuses on the characteristic of items present in ICILS 2013 test. An interesting study comparatively investigates on how different countries, efficiently use educational inputs to affect information and communication technology literacy.

Tab. 3 – ICILS 2013 in depth studies on all participating European countries

Author	*Title*	*Source*
Gerick, Eickelmann, Bos (2017)	The international computer and information literacy study from a European perspective: Introduction to the special issue	*European Educational Research Journal*, 16 (6), pp. 707-715
Punter, Meelissen, Glas (2017)	Gender differences in computer and information literacy: An exploration of the performances of girls and boys in ICILS 2013	*European Educational Research Journal*, 16 (6), pp. 762-780

Tab. 4 – ICILS 2013 in depth studies on specific countries

Author	*Title*	*Source*	*Country*
Drossel, Eickelmann (2017)	Teachers' participation in professional development concerning the implementation of new technologies in class: A latent class analysis of teachers and the relationship with the use of computers, ICT self-efficacy and emphasis on teaching ICT skills	*Large-Scale Assessments in Education*, 5 (1), p. 19, https://doi.org/10.1186/s40536-017-0053-7	Germany, the Czech Republic
Drossel, Eickelmann, Gerick (2017)	Predictors of teachers' use of ICT in school – the relevance of school characteristics, teachers' attitudes and teacher collaboration	*Education and Information Technologies*, 22 (2), pp. 551-573	The Netherlands, Denmark, Australia, Poland, Germany
Drossel, Eickelmann, Schulz-Zander (2017)	Determinants of teachers' collaborative use of information and communications technology for teaching and learning: A European perspective	*European Educational Research Journal*, 16 (6), pp. 781-799	Slovak Republic, Lithuania, Czech Republic, Poland, The Netherlands, Germany
Elstad (*ed.*) (2016)	Digital expectations and experiences in education	Sense Publishers, Rotterdam	Norway
Gerick, Eickelmann, Bos (2017)	School-level predictors for the use of ICT in schools and students' CIL in international comparison	*Large-Scale Assessments in Education*, 5 (1), p. 5, https://doi.org/10.1186/s40536-017-0037-7	Australia, Germany, Norway, the Czech Republic
Mažgon, Šebart, Štefanc (2015)	The role and use of e-materials in vocational education and training: The case of Slovenia	*Turkish Online Journal of Educational Technology*, 14 (4), pp. 157-164	Slovenia
Senkbeil, Ihme (2020)	Assessment of ICT Literacy: Do Multiple-Choice Tasks and Simulation-Based Tasks Measure Comparable Constructs? Comparison of the Test Results of Two Large-Scale Instruments: ICILS 2013 Versus NEPS	*Diagnostica*, 66 (3), 147-157	Germany

Tab. 5 – ICILS 2018 in depth studies

Author	*Title*	*Source*	*Country*
Drossel, Eickelmann, Vennemann (2020)	Schools overcoming the digital divide: In depth analyses towards organizational resilience in the computer and information literacy domain	*Large-Scale Assessments in Education*, 8 (1), p. 9, https://doi.org/10.1186/s40536-020-00087-w	All educational systems
Eickelmann, Labusch, Vennemann (2019)	Computational Thinking and Problem-Solving in the Context of IEA-ICILS 2018	Passey, Bottino, Lewin, and Sanchez (*eds.*), "Empowering Learners for Life in the Digital Age. OCCE 2018", in *IFIP Advances in Information and Communication Technology*, vol. 524, Springer, Cham, https://doi.org/10.1007/978-3-030-23513-0_2	All educational systems, Germany
Kameshwara, Nurullah, Meng, and Sandoval-Hernandez (2020)	Teachers' pedagogical autonomy, professional development and students' digital skills: New evidence from Italy	"Autonomie locali e servizi sociali", *Quadrimestrale di studi e ricerche sul welfare*, 2 (20), pp. 421-439	Italy

A study on gender differences was carried out on the 12 European countries involved in ICILS 2013; the other study provides various secondary comparative in-depth analyses at a European level (tab. 3). In tab. 4 and tab. 5, the "country" column shows the geographical areas that the studies explore.

Few studies focus on ICILS 2018 data, with only one in-depth study on Italy (tab. 5).

As we can see, compared to the studies on ICILS 2013 data, there are fewer in-depth studies on ICILS 2018 data, including studies that focus on all educational systems. Exploring the literature, we realised that our research interest had not yet been the subject of any study involving Italy. The only previous study focused on ICILS 2018 Italian data aimed to investigate the relationship between teachers' pedagogical autonomy and students' CIL achievement. The results showed that teachers' pedagogical autonomy, moderated by teachers' reciprocal professional development in the school context, had a positive correlation to students' CIL achievement. Based on these results, the authors recommended "education authorities in Italy to shift more resources to local schools in addition to providing centralised professional development training to teachers" (Kameshwara *et al.*, 2020).

In the following section, we present our study in which the Italian ICILS results are analysed by student characteristics and ICT activities.

3. Italian results in CIL by student characteristics and activities

The subject of this study is the digital and informative skills of Italian students in relation to their characteristics and the activities carried out. In particular, student results are studied in relation to:

- students' characteristics such as gender, migratory origin and socio-economic-cultural background;
- school and extracurricular activities that contribute to developing students' digital and informative skills.

The data has been collected by INVALSI in collaboration with IEA, as part of the ICILS 2018 Project. In Italy the administration took place at the beginning of the school year, so the average age of Italian students is the lowest, 13.3 years old, since in the other participating countries the administration was conducted in the second part of the school year. The Italian students' sample includes 2,810 students of whom 53.3% are males and 46.7% females; 10.3% of students share an immigrant background, for 3.1% first generation and for 7.2% second generation. As already mentioned, 20 students were randomly extracted from 150 schools sampled using the Probability Proportional to Size method. This study focuses on results of the CIL test and on information collected through the Student Questionnaire. The hypothesis investigated is that the data converge to confirm the useful and necessary inclusion of digital and informative skills at school.

3.1. Analyses and results

Following the IEA recommendations for analysing ICILS data (Fraillon *et al.*, 2020; Mikheeva and Meyer, 2020), the statistical analyses were conducted using IEA's International Database (IDB) Analyzer (IEA, 2020) with SPSS (IBM Corporation, 2016) on the ICILS Italian database. This approach allowed us to take into consideration the complex sample design (students nested within schools) and to use the appropriate sample weights and variance estimation techniques (Fraillon *et al.*, 2020; Mikheeva and Meyer, 2020). Differences between students with contrasting characteristics (e.g., boys vs. girls; immigrant students vs. native students) were examined by means of mean difference tests. Differences in ICT resources and activities between Italy and the ICILS countries were examined by means of differences in percentages. Finally, correlation coefficients and linear regression analyses were computed in order to examine the relationships between Italian results in ICILS and students' characteristics or ICT activities.

3.1.1. Variation in students' CIL in relation to their characteristics

The results of our analysis showed (fig. 2) that the average CIL scale scores of female students were significantly higher than those of male students ($p < .05$).

In the ICILS Italian sample, 10.3% of the students had an immigrant background, 3.1% of them were first generation students (i.e., foreign-born students with foreign-born parents), while 7.2% of them were second generation students (i.e., students born in Italy and with foreign-born parents). Students with an immigrant background (fig. 2) achieved significantly lower scores than native students. However, there were important differences between first generation and second-generation students (fig. 3). Our analyses showed that first generation students had scores that were significantly lower than native students, while there were no significant differences between second generation and native students. The CIL scores of first-generation students were also significantly lower than those of second-generation students.

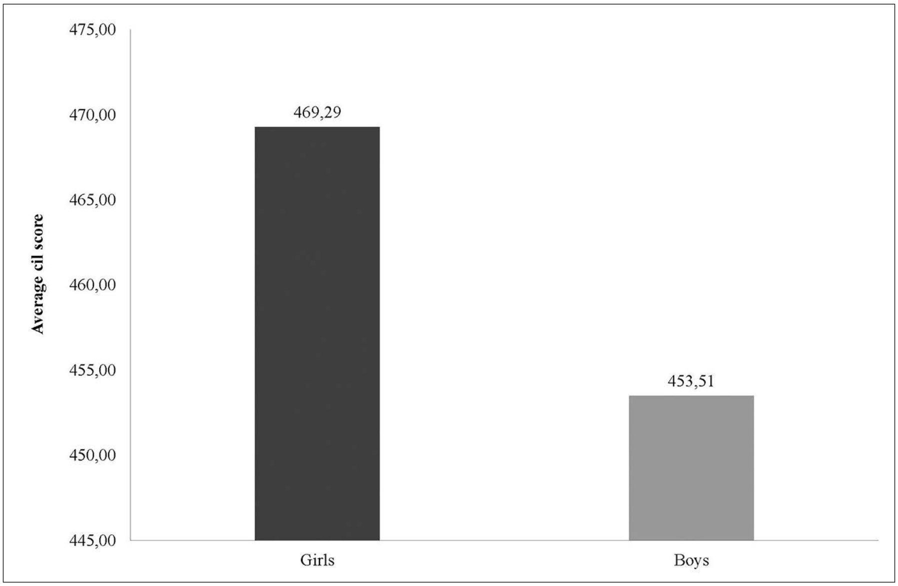

Fig. 2 – Average CIL scale scores for boys and girls

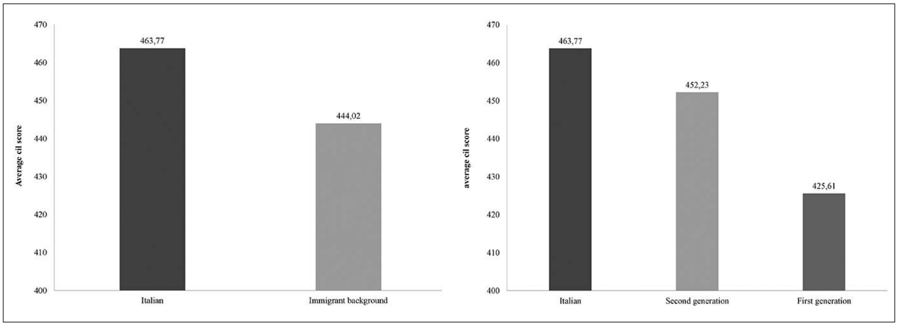

Fig. 3 – Average CIL scale scores for Italian students and students with an immigrant background

In the Italian ICILS sample, most of the students (78.6%) speak Italian at home most of the time, while 15.5% speak a dialect and 5.8% speak another language. Our analysis showed that there were significant variations in CIL scores between students speaking Italian and those speaking dialect or another language (fig. 4). More specifically, students speaking dialect, or another language obtained significantly lower scores ($p < .05$) than students speaking Italian. There were no statistically significant differences between students speaking a dialect and students speaking another language.

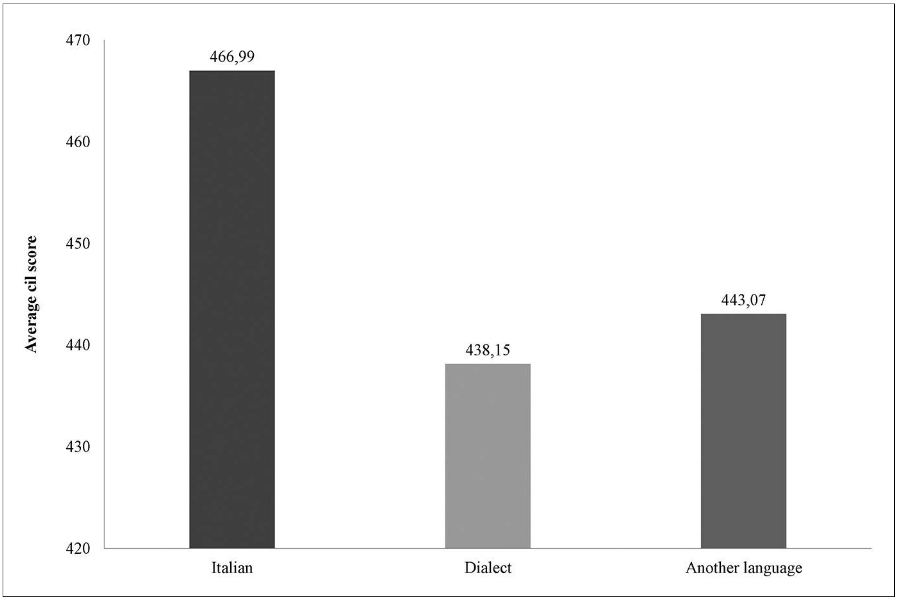

Fig. 4 – Average CIL scale score by language spoken at home

In ICILS 2018, the socioeconomic status (SES) of the students' families was measured based on the National Index of Students' Socioeconomic Background (NISB; Fraillon *et al.*, 2020). The NISB was derived from the following three indices: highest occupational status of parents, highest educational level of parents, and the number of books at home. The NISB scores consists of factor scores from the principal component analysis conducted on these indices, with an average of 0 and standard deviations of 1. Our analysis showed that the students' SES had a positive impact on their CIL scale scores ($B = 26.21$, $p < .05$). Therefore, higher SES students achieved significantly better CIL performances. We computed also the tertiles of the SES scores, in order to distinguish three groups of students: lower SES (comprising students with SES scores in the first tertile), central SES (comprising students with SES score in the second tertile), and higher SES (comprising students with SES score in the third tertile). Then, we compared the CIL scale scores of these three groups of students (fig. 5). The results of this analysis showed that the average CIL scores of students in the "higher" groups were statistically significantly higher than that of students in the "lower" groups ($p < .05$).

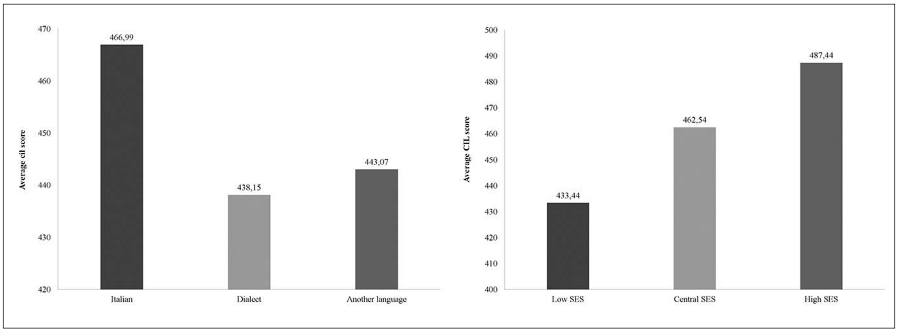

Fig. 5 – Average CIL scale scores by socioeconomic status

3.1.2. Variations in students' CIL in relation to ICT at school

In ICILS 2018, school ICT coordinators were asked about the extent to which they perceived that in their school the use of ICT for teaching was hindered by different factors, such as insufficient internet bandwidth or speed or not enough computers for instruction. The results were reported at the level of students: that is, in percentages of students who were enrolled at schools where each of the factors was reported as a hindrance to ICT use for teaching a lot or to some extent. Tab. 6 shows these results for Italy and the ICILS 2018 average.

In Italy, the percentage of students enrolled at schools where the use of ICT for teaching and learning was hindered by insufficient ICT resources were significantly higher than the ICLS average in all the factors taken into consideration (tab. 6). Therefore, Italian schools seem to suffer more than other ICILS countries for the lack of ICT resources.

In ICILS 2018, students were asked to indicate the extent to which they had learned how to do various ICT tasks at school, which included, for example: provide references to internet resources, work out whether to trust information from the internet, present information for a given audience or purpose using ICT. Tab. 7 shows, for each ICT task, the percentages of Italian students recording that they learned at school about it to a large or moderate extent, together with the ICILS average.

Tab. 6 – Percentages of students enrolled at schools where ICT coordinators reported that the use of ICT for teaching and learning was hindered a lot or to some extent by insufficient computer resources

	Percentages of students enrolled at schools with					
	Too few computers with an internet connection	*Insufficient internet bandwidth or speed*	*Not enough computers for instruction*	*Lack of sufficiently powerful computers*	*Problems in maintaining ICT equipment*	*Not enough computer software*
Italy	40 (4.0) ▲	63 (4.3) ▲	61 (3.9) ▲	65 (3.7) ▲	72 (3.5) ▲	56 (4.5) ▲
ICILS average	28 (1.1)	49 (1.2)	48 (1.3)	47 (1.3)	44 (1.3)	38 (1.3)

▲ More than 10 percentage points above the ICILS 2018 average

Source: IEA ICILS 2018 International Report (Fraillon *et al.*, 2019b)

Tab. 7 – Percentages of students who reported having learned to a large or moderate extent about CIL at school

	Percentages of students who reported having learned at school to a large or moderate extent to							
	Provide references to internet resources	*Search for information using ICT*	*Present information for given audience or purpose using ICT*	*Work out whethe to trust information from the internet*	*Decide what information obtained from the internet is relevant to include in schoolwork*	*Organize information obtained from internet resources*	*Decide where to look for information on the internet about an unfamiliar topic*	*Use ICT to collaborate with others*
Italy	50 (1.3) ▽	70 (1.4) ▽	51 (1.2) ▼	64 (1.2)	66 (1.1)	66 (1.1) ▽	67 (1.0)	58 (1.3) ▽
ICILS average	68 (0.3)	74 (0.3)	66 (0.3)	65 (0.3)	68 (0.3)	68 (0.3)	67 (0.3)	60 (0.4)

▽ significantly below the ICILS 2018 average ▼ more than 10 percentage points below the average

Source: IEA ICILS 2018 International Report (Fraillon *et al.*, 2019b)

The percentage of Italian students who learned the ICT tasks at school to a large or moderate extent was significantly lower than the ICILS average in most of the tasks. In three tasks (i.e., work out whether to trust information from the internet, decide what information obtained from the internet is relevant to include in schoolwork, and decide where to look for information on the internet about an unfamiliar topic) the percentages of Italian students were similar to the ICILS average. Therefore, compared to other ICILS countries, fewer Italian students seem to learn about ICT at school.

A scale was derived from the items of the questionnaires that asked students about learning the different tasks of ICT at school (Fraillon *et al.*, 2020). Higher scores on this scale indicate greater attribution to school-based CIL learning. We analyzed the correlation between students' score on this scale and their CIL scale score. Students' learning of ICT at school was positively correlated with their CIL scores ($r = .15$; $p < .05$). Therefore, students had better CIL performances when they have learned about ICT at school to a greater extent. This suggests that schools can foster students' computer information literacy by enhancing their learning of ICT at school.

Students were asked to report how often they used ICT for school-related purposes that ranged from the conventional to less conventional and included for example: prepare reports or essays; use coding software to complete assignments, work online with other students.

Tab. 8 shows the percentages of Italian students that used ICT on a weekly basis for each school-related purpose, together with the ICILS average.

The percentages of Italian students who reported a weekly use of ICT for school-related purposes were significantly lower than the ICILS average on most of the tasks. However, the percentages of students were higher than the ICILS average on two tasks: use the internet to do research and make video or audio productions.

All the items in this scale were used to derive a scale of students' use of ICT for study purposes, where higher scale scores reflect more frequent use of ICT. We analyzed the correlation between students' score on this scale and their CIL scale score. The correlation between the students' use of ICT for school related purposes and their CIL scores was not statistically significant. Therefore, this kind of use of ICT did not seem to be related to students' CIL. This result was consistent with findings from other ICILS' countries (Fraillon *et al.*, 2019b).

Tab. 8 – Percentages of students using ICT on a weekly basis for specified school-related purposes

	Percentages of students who reported at least weekly use of ICT to									
	Prepare reports or essays	*Prepare presentations*	*Work online with other students*	*Complete worksheets or exercises*	*Organize your time and work*	*Take tests*	*Use software or applications to learn skills or a subject*	*Use the internet to do research*	*Use coding software to complete assignments*	*Make video or audio productions*
Italy	20 (0.9) ▽	14 (0.8) ▽	15 (0.7) ▼	18 (0.9) ▼	24 (1.0) ▽	14 (0.6) ▽	22 (0.8) ▽	62 (1.2) △	13 (0.7)	22 (1.0) △
ICILS average	26 (0.3)	22 (0.3)	25 (0.3)	30 (0.3)	28 (0.3)	20 (0.3)	24 (0.3)	59 (0.4)	14 (0.3)	18 (0.3)

▽ significantly below the ICILS average; ▼ more than 10 percentage points below the ICILS average; △ significantly above the ICILS 2018 average

Source: IEA ICILS 2018 International Report (Fraillon *et al.*, 2019b)

Finally, we focused on the Italian students' use ICT during lessons. In ICILS 2018 students were asked how often they used each of 11 listed ICT tools during lessons. Some of these tools were general applications (productivity, word processing, and presentation software and computer-based information resources), while others were specialist applications (multimedia production, concept mapping, real-world data capture, simulations and modeling software, computer-based information resources, interactive digital learning resources, and graphing or drawing software). Tabb. 9 and 10 show the percentages of Italian students who used each of these tools in "most" or in "every or almost every lesson", together with the ICILS average.

Tab. 9 – Percentages of students using general ICT applications during most or all lessons

	Percentages of students who reported at least weekly use of ICT (general applications)			
	Tutorial or practice software	*Word-processing software*	*Presentation software*	*Spreadsheets*
Italy	8 (0.6) ▽	14 (0.7) ▼	15 (0.7) ▼	10 (0.6) ▽
ICILS average	13 (0.3)	28 (0.4)	26 (0.4)	16 (0.3)

▽ significantly below the ICILS 2018 average; ▼ more than 10 percentage points below the ICILS 2018 average

Source: IEA ICILS 2018 International Report (Fraillon *et al.*, 2019b)

Tab. 10 – Percentages of students using specialist ICT applications during most or all lessons

Percentages of students who reported at least weekly use of ICT (specialist applications)							
	Multimedia production tools	*Concept mapping software*	*Tools that capture real-world data*	*Simulations and modeling software*	*Computer based information resources*	*Interactive digital learning resources*	*Graphing or drawing software*
Italy	12 (0.8)	15 (0.8) △	10 (0.6)	8 (0.6)	29 (1.1)	14 (0.7)	14 (07)
ICILS average	11 (0.2)	9 (0.2)	10 (0.2)	8 (0.2)	29 (0.4)	15 (0.3)	14 (0.3)

△ significantly above the ICILS 2018 average

Source: IEA ICILS 2018 International Report (Fraillon *et al.*, 2019b)

The percentages of Italian students who reported using general applications during most or all lessons (tab. 9) were significantly lower than the ICILS average of all the different tools taken into consideration. However, the percentages of Italian students using specialist applications (tab. 10) were not statistically different from the ICILS average and were higher for using concept mapping software.

Two scales were derived from these items (Fraillon *et al.*, 2020), one represented the extent to which general applications were used in class, while the second represented the extent to which specialist applications were used in class. We analyzed the correlations between Italian students' CIL scores and each of these two scales. The use of general applications in class was positively correlated with students' CIL scores (r = .20; p < .05), while the use of specialistic applications in class was negatively correlated with students' CIL scores (r = -.09; p < .05). Therefore, students achieved better CIL performances when general applications were used more frequently during lessons, but they achieved worse CIL performances when specialistic applications were used more frequently during lessons. These results are consistent with findings from other ICILS countries and further studies are needed to understand them (Fraillon *et al.*, 2019b).

4. CIL and Innovative school library

The spreading of the Innovative School Libraries would make a great contribution to enhance Italian students' competences in CIL. In Italy, the development of the Innovative School Libraries was introduced with Law 107/2015, and specifically with Action #24 of the PNSD. The Plan outlined the guidelines of the future school towards a path of innovation and digitalization. However, investments in learning environments to develop and practice digital and informative skills do not yet seem adequate; in terms of policies, it is necessary to overcome a fragmented system whereby investments are concentrated in a few schools (Fornasari, 2019). Particularly, the Innovative School Libraries have not been implemented in every school; the funds allocated for their activation were able to satisfy only a limited number of schools, participating in a dedicated Project Call; schools without a library could not participate in such a Call, thus increasing the differences between schools. Last but not least, the lack of dedicated personnel, i.e., school librarians, adversely interferes with the functioning of school libraries.

The PNSD assigned a key role to Innovative school libraries in overcoming all forms of disadvantages, including the digital divide. In fact, these advanced school libraries combine the digital aspect with the historical link between libraries and information literacy.

The expression information literacy was coined in USA within the Information Industry Association during 1974; in the same year the concept was transposed and applied in library field which remained for decades the most fertile environment for its development. Fifteen years later, following recommendations from the ALA[10] on information literacy, the NFIL[11] was created and currently consists of more than 90 international organizations; it engages to introduce and develop information literacy in educational and work fields. Since almost 50 years ago, IASL[12] has been organising an annual conference comparing school librarians, teachers, and researchers on how school libraries can contribute to the development of lifelong learning and information skills. So, the current concept of information literacy was developed with the main contribution of library science and partly of psychology; it relates to the processes of identifying information needs, researching and locating information, evaluating their quality. Twenty-first century literacy passes through the ability to be "computer information literate" that is be literate in the search for information even in a digital environment. In ICILS 2018 (Fraillon *et al.*, 2019a) Computer Information Literacy is conceptualized as the set of skills that an individual must possess to meet the requirements imposed by the communication society. The strong link between information literacy and school libraries is supported by influential institutions at international level as IFLA and UNESCO. Translated into many languages[13], "School Library Manifesto – Teaching and Learning for all" (IFLA and UNESCO, 1999)[14] promotes the school library as part of the educational process and strategic environment for information literacy; in particular, it recognizes the key role of school libraries in the development of media and information skill programs (Grizzle *et al.*, 2013). Equal emphasis on the importance of school libraries for information literacy can be found in IFLA "Guidelines for School Libraries", published in 2002[15] and reviewed in 2015[16]. In addition to these historical authoritative Institutions, the most recent ENSIL14 documents and shares experiences of using the school library from an information literacy perspective; Italy also adheres to this important network. Actually, there is no shortage of virtuous examples at national level; mainly university libraries have long included, in their annual work plan, information literacy activities;

[10] American Library Association.

[11] National Forum on Information Literacy.

[12] International Association of School Librarianship.

[13] Italian edition by the Italian Library Association, 1999: Translation by Luisa Marquardt revised in 2003.

[14] Updated edition expected in 2021.

[15] Italian edition by the National Commission of School Libraries of the Italian Library Association. Coordination and review by Luisa Marquardt and Paolo Odasso, Rome, AIB, 2004.

[16] Italian translation by Luisa Marquardt.

contributions have been given to suggest methods and content of information education courses in school library (Ballestra, 2011). But the didactic function assigned to the school librarian in European countries such as Denmark, Germany, France, England, Norway, and Poland (Bolletti, Lombello and Marquardt, 2000) reveals a gap for Italy that should be recovered in order to develop media and information literacy. The Italian education system took the first steps in this direction during the last school reform (Law 107/2015) by establishing the "Innovative school libraries". A subsequent project has been set up to which a limited number of schools have been able to join and to develop these advanced digital libraries. The Innovative school libraries could play an important role in the development of digital and informative skills, therefore a broader plan for the digitalization of the school should provide for its diffusion, enhancement, and dedicated library staff. Although not deepened in ICILS, Denmark – which ranked first in CIL – is the country where the school libraries are compulsory in all school orders.

The introduction of information literacy as a separate subject raises the same concerns as the single teaching of digital technologies, which would have led to a partial failure (Calvani, 2013). Rather, a review of teaching/learning patterns may be necessary, including the development of these skills in relation to other areas of competence. For example, the teaching of digital citizenship (introduced from September 2020) would find a complement in information literacy and a privileged place of development in Innovative School Library.

5. Conclusions

Active and conscious participation in the society of the 21st century requires sophisticated digital and informative skills for research, evaluation, and use of information. It has long been argued that the lack of political attention and adequate training in digital skills is associated with increasing levels of inequality (Van Dijk, 2006). ICILS results support the central role of school in the development of CIL, stating the need to integrate these skills into teaching and learning. Countries with the best CIL scores have the highest percentages of students attending schools where they frequently learn and practice ICT tasks; in these schools there are adequate infrastructure, learning materials and professional learning support for teachers. The results of the study on Italian ICILS data confirm the usefulness and suggest the urgency of implementing CIL at school to improve achievement and reduce the digital divide. Actually, Italian students who frequently learn and use

ICT tasks at school achieve better in CIL than those who do not, unfortunately the latter prevail. In Italy, the under-performing students in CIL are 63% of students, many more than on the international average (43%) and in the other EU countries. The scores of Italian students in CIL give cause for concern, because of the digital divide within the country and in relation to other countries. The analyses on CIL in relation to characteristics of Italian students such as gender, origin and socio-economic-cultural background confirm the differences found at international level. The analyses conducted dividing the students into three groups by SES scores show a difference of more than 50 points between CIL average score of students with high SES and that of students with low SES. These differences should be better investigated and faced in relation to the territorial disadvantaged background and to the "segregation" in schools with few resources and low achievement. Looking forward to a major national school digitalization plan, it should be stressed that a change in traditional teaching and learning patterns is just as necessary. The mere presence at school of computer equipment is not sufficient to develop digital and informative skills in the dearth of teaching and practice of ICT tasks. The European Commission in "Education and Training Monitor 2020 – Italy" dedicates a section focused on teaching and learning in the digital age, highlighting current critical issues, and recommending the proper changes for improvement. The EU "Digital Education Action Plan (2021-2027): Resetting Education and Training for the Digital Age", with the target *of less than 15% of under-performing students in CIL by 2030*, is a major challenge for EU countries but especially for Italy.

The Innovative school libraries have been recognized (PNSD, School Reform, 2015) as information and documentation centres with a key role in learning and overcoming any disadvantage, including the digital divide; they could be a landmark for the development of CIL programs as well as a learning environment for these skills. The link between library and information literacy is recalled by the achievement of the best ICILS results in Denmark, the country where the school libraries are compulsory in all schools. Finland also has excellent results in CIL and is known for its inclusive digital libraries that support the education system, probably the best in the world and certainly a virtuous example of innovation. Collaborative teaching, since 2016 introduced in Finnish schools, aims to provide students with the appropriate skills for life in the knowledge society. A subject is addressed from many points of view and by various means, including the use of technologies. Particular attention is paid to the conscious and critical use of technologies that allows to identify fake news and avoid cyber bullying (Repubblica, 2017).

References

AIB – Commissione nazionale biblioteche scolastiche, Marquardt L. (trad. a cura di) (2000, rev. 2003), *IFLA/UNESCO Manifesto sulla biblioteca scolastica*, AIB, Roma, retrieved on June 21, 2021, from http://www.aib.it/aib/commiss/cnbse/manif.htm.

AIB – Commissione nazionale biblioteche scolastiche, Marquardt L., Odasso P. (trad. a cura di) (2004), *Linee guida IFLA/UNESCO per le biblioteche scolastiche*, AIB, Roma.

AIB – Commissione nazionale biblioteche scolastiche, Marquardt L. (trad. a cura di) (2016), *Linee guida IFLA/UNESCO per le biblioteche scolastiche. 2ª edizione riveduta*, AIB, Roma, retrieved on June 21, 2021, from http://www.retebibliomediascuole.org/wp-content/uploads/2020/06/ifla-school-library-guidelines-it.pdf.

Ballestra L. (2011), *Information literacy in biblioteca. Teoria e pratica*, Editrice Bibliografica, Milano.

Bolletti M., Lombello D., Marquardt L. (a cura di) (2000), *Senza confini: formazione e azione educativa del bibliotecario scolastico*, CLEUP, Padova.

Cadalanu G. (2017, 30 maggio), "Finlandia, nasce la scuola senza materie: la rivoluzione dei più bravi del mondo", *la Repubblica*, retrieved on June 21, 2021, from https://www.repubblica.it/esteri/2017/05/30/news/finlandia_nasce_la_scuola_senza_materie_la_rivoluzione_dei_piu_bravi_del_mondo-166760382/.

Calvani A. (2013), "Qual è il senso delle tecnologie nella scuola? Una "road map" per decisori ed educatori", *TD Tecnologie Didattiche*, 21 (1), pp. 52-57, retrieved on June 21, 2021, from https://ijet.itd.cnr.it/article/view/123/65.

Directorate General for Education, Youth, Sport and Culture – European Commission (2020), *Education and Training monitor 2020 – Italy*, Publications Office of the European Union, Luxembourg.

Fornasari A. (2019), *Digital Education Policies in Italy: A Recognition on the Actions Realized. Proceedings of the 6th International Conference on Educational Technologies 2019*.

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 June 21, 2021, 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 June 21, 2021, from https://research.acer.edu.au/ict_literacy/23.

Fraillon J., Ainley J., Schulz W., Friedman T., Duckworth D. (2020), *IEA International Computer and Information Literacy Study 2018 Technical Report*, IEA, Amsterdam, retrieved on June 21, 2021, from https://www.iea.nl/sites/default/files/2020-05/ICILS%202018%20Technical%20Report-FINAL_0.pdf.

Fraillon J., Ainley J., Schulz W., Friedman T., Gebhardt E. (2014), *Preparing for Life in a Digital World. IEA International Computer and Information Literacy Study*

International Report, IEA, Amsterdam, retrieved on June 21, 2021, from https://research.acer.edu.au/cgi/viewcontent.cgi?article=1009&context=ict_literacy.
Fraillon J., Schulz W., Ainley J. (2013), *International Computer and Information Literacy Study: Assessment Framework*, IEA, Amsterdam, retrieved on June 21, 2021, from (PDF) ICILS 2013 Framework | me za – Academia.edu.
Grizzle A., Moore P., Dezuanni M., Asthana S., Wilson C., Banda F., Onumah C. (2013), *Media and Information Literacy: policy and strategy guidelines*, UNESCO, Paris, retrieved on June 21, 2021, from https://unesdoc.unesco.org/ark:/48223/pf0000225606.
IFLA School Libraries Section Standing Committee, Oberg D., Schultz-Jones B. (*eds.*) (2015), *IFLA school library guidelines*, 2nd revised edition, IFLA, The Hague, retrieved on June 21, 2021, from https://www.ifla.org/files/assets/school-libraries-resource-centers/publications/ifla-school-library-guidelines.pdf.
IFLA, UNESCO (1999), *IFLA/UNESCO School Library Manifesto*, IFLA, The Hague, retrieved on June 21, 2021, from www.ifla.org/publications/iflaunesco-school-library-manifesto-1999.
Kameshwara K.K., Eryilmaz N., Tian M., Sandoval-Hernández A. (2020), "Teachers' Pedagogical Autonomy, Professional Development and Students' Digital Skills: New Evidence from Italy", *Autonomie locali e servizi sociali*, 2, pp. 421-439, retrieved on June 21, 2021, from https://doi.org/10.1447/98729.
Mikheeva E., Meyer S. (2020), *IEA International Computer and Information Literacy Study 2018 User Guide for the International Database*, IEA, Amsterdam, retrieved on June 21, 2021, from https://www.iea.nl/sites/default/files/2020-03/ICILS2018_IDB_User_Guide.pdf.
Pagani L., Argentin G., Gui M., Stanca L. (2016), "The impact of digital skills on educational outcomes: evidence from performance tests", *Educational Studies*, 42 (2), pp. 137-162.
Van Dijk J.A. (2006), "Digital divide research, achievements and shortcomings", *Poetics*, 34 (4-5), pp. 221-235.

6. Difficulty perception in answering argumentative INVALSI tests: a qualitative study

by Camilla Spagnolo, Marta Saccoletto

The paper shows preliminary results of a qualitative investigation focused on students' perceived difficulties after solving INVALSI Math tasks (grade 9 and 10 students). This survey is part of a larger study made up of a first qualitative phase and a second quantitative phase. The purpose of the first qualitative phase is to identify the factors that influence students' perceived difficulties.

Despite factors that contribute to increasing or decreasing the difficulties of a task (in an absolute sense) are widely studied in the literature, the student's perceived difficulty relative to a Math task is not.

In order to investigate students' metacognitive factors, mathematical competencies (through the resolution of INVALSI tasks), as well as the explication of difficulties faced in the comprehension and resolution of tasks, a questionnaire was administered during the qualitative phase. The aim of this research is to highlight some of the factors that contribute to building students' perceived difficulties in relation to Mathematical tasks.

Future phases of the research will allow us to compare students' perceived difficulty to that of teachers and to the skill levels constructed by INVALSI.

Mostriamo i primi risultati di un'indagine qualitativa incentrata sulla percezione di difficoltà di studenti (di grado 9 e 10) dopo aver risolto dei task INVALSI di Matematica. Questa indagine si inserisce in uno studio più ampio composto da una prima fase qualitativa e una seconda fase quantitativa. Scopo della prima fase qualitativa è quello di delineare i fattori che influiscono sulle percezioni di difficoltà degli studenti.

In letteratura sono ampiamente studiati i fattori che contribuiscono ad aumentare o diminuire le difficoltà di un task (in senso assoluto), ma non la percezione di difficoltà dello studente relativamente a un task.

Durante la fase qualitativa è stato somministrato un questionario con lo scopo di indagare fattori metacognitivi degli studenti, competenze matematiche (attraverso la risoluzione di task INVALSI), nonché l'esplicitazione delle difficoltà incontrate nella comprensione e risoluzione dei task. Obiettivo di questa ricerca è mettere in luce alcuni dei fattori che contribuiscono a costruire la percezione di difficoltà degli studenti relativamente a task matematici.

Le fasi successive della ricerca permetteranno di mettere in relazione la percezione di difficoltà degli studenti con quella degli insegnanti e con i livelli di abilità costruiti da INVALSI.

1. Introduction

For research in Mathematics Education, the topic of *difficulties* is exceptionally extensive and central. Several researchers have considered many of the aspects related to difficulties (Capozio *et al.*, 2018).

Some studies aim to figure out the causes, to find out different factors that impact on the issue and to develop – starting from this knowledge – paths of prevention and getting over difficulties (see for example Baccaglini-Frank *et al.*, 2018; Zan, 2007).

The main factors that literature has highlighted relate to i) students' abilities or skills; ii) difficulty of the task in terms of abilities or skills needed to solve it; iii) affective factors. The factors that are most closely related to the aim of our research are affective factors, such as beliefs, emotions, and attitudes (McLeod, 1992).

The aim of our research is to highlight some of the factors related to the difficulties students *perceive* when faced with a mathematical task.

In mathematics Education, there is no agreed upon definition of "perceived difficulty" and few surveys have explored factors related to this topic. Therefore, we would try to clarify what we mean here by considering a specific situation: when a student is solving a math task. In the process of solving a Mathematics task, a student may stumble across multiple difficulties that may depend both on the student's own peculiarities (such as his/her skills and knowledge, beliefs and attitudes) or on peculiarities of the task (such as the text or the mathematical content involved). These characteristics of the Mathematical task are the same ones that can also influence his idea of the task and consequently can influence his perceived difficulties. Therefore, while closely related, difficulty and perceived difficulty are two different issues. We believe that one of the main differences is that perceived difficulty influences students' behaviour in approaching a Math task.

The perceived difficulty of a student does not necessarily coincide with the "objective" difficulty of the task (which, for instance, is attributed through the levels of ability built by INVALSI), or with the teacher's perceived difficulty of the task. Consequently we think that identifying some factors related with students' perceived difficulties could help us understand why some INVALSI tasks have blank answers. In order to do so, we will have to distinguish between the perceived difficulty before and after solving a Math task.

In this paper we deal only with the factors that influence the perception of difficulty *after* solving the task.

2. Background of the research and statement of the problem

In recent years, several authors have investigated how the variation of some elements within the questions can impact the students' behavior and their perception of difficulty (Vicente *et al.*, 2007).

Among the factors that influence students' approach to answering a written test, the text phrasing is vital: one formulation is not necessarily better or worse than another, but changing the wording actually changes the problem (Bagni and D'Amore, 2005). As pointed out by D'Amore (2000) and Bolondi, Branchetti, and Giberti (2018), text changes, even if minimal, can affect students' approach to the problem. The text factors include, for example, editing, punctuation, syntactic complexity, word density, information order, explicit declaration of intermediate objects necessary for the solution (Laborde, 1995). However, students' solution strategies can also be influenced by numerical factors, such as numerical magnitude (Thevenot and Oakhill, 2005). This result is confirmed and deeply analyzed by De Corte, Verschaffel and Van Coillie (1988), who studied how the kind of number – such as integer, decimal greater or smaller than 1 – can affect students' *difficulties*.

According with the literature results, we hypothesized and analyzed that, among the perceived-difficulties-influencing-factors, there may be the "three components of the difficulty of the WP [Word Problem]" (Daroczy *et al.*, 2015): the linguistic complexity of the question text, the numerical complexity of the arithmetic problem and the relationship between the linguistic complexity and the problem numerical complexity. Moreover, we investigated other fundamental aspects, such as students' attitudes and beliefs in mathematics.

3. Inquiring design

As we anticipated before, the study is work in progress. The inquiring design concerns two phases: a first qualitative phase will be followed by a second quantitative one. This paper concerns the preliminary results of the qualitative study, which involved upper secondary students. These results (emerged from the first qualitative phase) will be used in the large-scale analysis (phase 2) to better understand their impact.

In the following we present some examples of factors, which emerged from the preliminary results of the qualitative phase.

This first qualitative phase has been led by quantitative data on a large scale: the research is qualitative, but the INVALSI tasks that we submit to students and the questions that we construct for the questionnaire are guided by the quantitative INVALSI results.

The experimentation involved 79 students: two grade 9 classes and two grade 10 classes of an Italian High School. Students completed an online questionnaire which was followed by whole-class focus groups.

3.1. Questionnaire and Focus group

The questionnaire and the focus group allowed us to investigate some of the factors that contribute to a student's personal perceived difficulty with argumentative questions and in particular how much of the perception of difficulty is related to "constructing an argument" and "recognizing an argument".

The questionnaire was composed of four sections:

- **Section 1**: the goal was to investigate what metacognitive factors and factors related to students' attitudes and beliefs might be;
- **Sections 2, 3**: the goal was to investigate perceived difficulty regarding two specific INVALSI tasks. Students had to first solve the task and then answer some questions related to the task and their perceived difficulty;
- **Section 4**: the goal was to let the student identify the most difficult task, justifying the answer.

Particularly, the questions in the first section aimed to examine diagnoses of difficulties related to negative or positive attitudes toward mathematics, but also about responsibility for learning. Specifically, four questions in this section were taken from a previously validated questionnaire (Zan and Baccaglini-Frank, 2017).

The questionnaire was administered in-person to students and was filled out through Google forms. The questionnaire was followed by whole-class focus groups conducted remotely through Google Classroom (applet Meet). The focus groups were conducted with the idea of helping categorize some of the answers given by the students. For this reason, the answers given by the students were commented on and a discussion with the whole class followed.

3.2. Choice of the tasks

The questionnaire concerned two INVALSI tasks. The criteria we followed to choose these two items are the following. Firstly, we decided to investigate the perceived difficulties relating to the INVALSI argumentative questions. Secondly, we paid attention to argumentative questions relating to the Numbers area. Thirdly, with the help of the teachers of the classes involved in the experimentation, we selected tasks whose content had already been dealt with. This decision made it possible to exclude that the perception of difficulty was influenced by the fact that the students did not know the topic. Fourthly, we chose INVALSI tasks closer to the grades of our students' sample. Due to INVALSI tests being administered to Italian students of grade 2, 5, 8, 10, 13, we decided to consider grade 8 and 10 tasks. Finally, we selected tasks with correct response percentages less than 50%, because we wanted to select tasks with evident difficulties.

The two tasks chosen had both common and different features from a mathematical perspective. On the one hand, the task chosen for Section 2 was a multiple-choice item that required recognizing a correct argumentation, while the task chosen for Section 3 was an open-ended item that required production of argumentation. On the other hand, for both items, focus of the content involved literal calculus and both tasks could be solved using the same strategy: proving the falsity of a statement through a counterexample. In other words, the mathematical content and the solution strategy were the same, while the requested process was different. Below we report the selected tasks.

Task 1 belongs to Section 2 of the questionnaire and was submitted to grade 8 Italian students by INVALSI in 2017.

n è un numero naturale.

Antonio afferma che "4*n*–1 è sempre un multiplo di 3".
Antonio ha ragione?
Nella tabella che segue indica <u>la sola</u> argomentazione che giustifica la risposta corretta.

Antonio ha ragione...	Antonio non ha ragione...
A. ☐ perché $4n-1=3n$	C. ☐ perché $4n-1$ è sempre dispari
B. ☐ perché se $n=4$ allora $4n-1=15$	D. ☐ perché se $n=3$ allora $4n-1=11$

Fig. 1 – Item D25-B, grade 8, Mathematics INVALSI test, 2017[1]

As mentioned above, the task requires recognizing the correct argumentation. It also requires being able to identify multiples and divisors of a natural number and to be able to deal with literal calculation.

National percentages show that 40.3% of students answer correctly (and thus choose answer D), 50.2% answer incorrectly, and 9.5% do not answer. Among students who get it wrong, 20.6% choose option A, 8.5% choose option B, and 21.1% choose option C. The results are summarized in the graph below.

[1] The authors translated Task 1 text as follows: "n is a natural number. Anthony affirms that '4n-1 is always a multiple of 3'. Is Anthony right? In the table below, mark *the only* argument that justifies the correct answer. A. Anthony is right because 4n-1=3n. B. Anthony is right because if n=4 then 4n-1=15. C. Anthony is not right because 4n-1 is always odd. D. Anthony is not right because if n=3 then 4n-1=11".

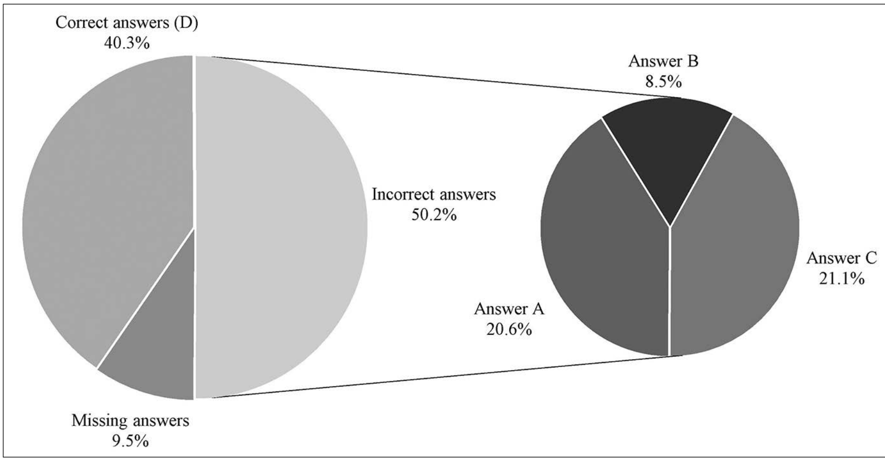

Fig. 2 – Results of item D25-B, grade 8, Mathematics INVALSI test, 2017

Task 2 belongs to Section 3 of the questionnaire and was submitted to grade 10 Italian students by INVALSI in 2014:

D6. **Marco afferma che, per ogni numero naturale *n* maggiore di 0, n^2+n+1 umero primo.**
Marco ha ragione?

Scegli una delle due risposte e completa la frase.

☐ **Marco ha ragione, perché** ……………………………………………………………………

……………………………………………………………………………………………

☐ **Marco non ha ragione, perché** ………………………………………………………………

……………………………………………………………………………………………

Fig. 3 – Item D6, grade 10, Mathematics INVALSI test, 2014[2]

As mentioned above, the task requires validating results through an argumentation. It also requires using the techniques and procedures of arithmetic and algebraic calculation, and being able to use letters as symbols and variables.

National percentages show that 17.8% of students answer correctly, 55.3% answer incorrectly, and 23.5% do not answer. The results are summarized in the graph below (fig. 4).

[2] The authors translated Task 2 text as follows: "Mark states that, for every natural number n greater than 0, n2+n+1 is a prime number. Is Mark right? Choose one of the two answers and complete the sentence. Mark is right, because… Mark is not right, because…".

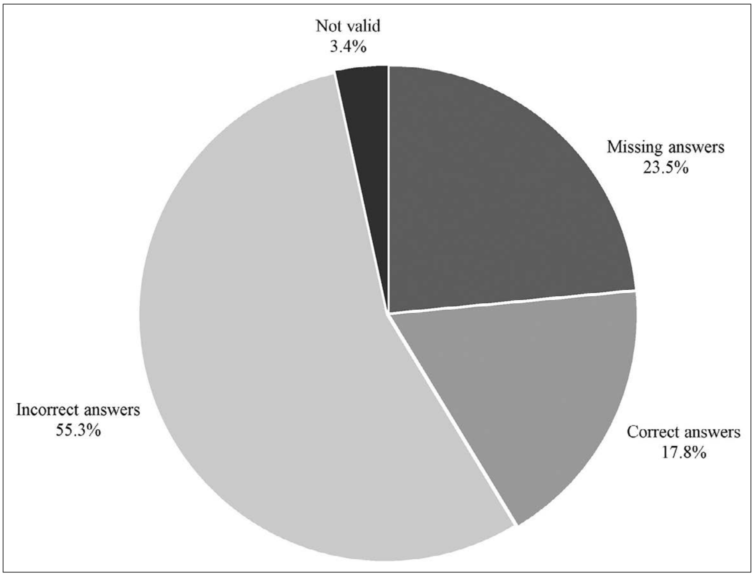

Fig. 4 – Results of item D6, grade 10, Mathematics INVALSI test, 2014

The final questions in Section 2 and 3 respectively are mostly open-ended questions and are the same for both sections. The purpose of these additional questions was to inquire students' ideas and to link them – in a strictly qualitative way – with students' attitudes, beliefs or peculiar INVALSI item elements. Particularly, we investigated the relation between the perceived difficulty and the request to construct or recognize an argument.

4. Data analysis[3]

4.1. First section questionnaire answers (metacognitive section)

At the beginning of the questionnaire, we asked students: "Write three adjectives that you think describe you properly as a math student". In order to obtain an immediate and easy to read data representation, we split the adjectives in positive (black words) and negative (italic grey words) (Capozio, Di Martino and Passaro, 2018). The words' size is directly proportional to the words' frequencies and we combined the synonyms in order to avoid repetitions.

[3] The questionnaire was administered in Italian and the students' answers were also collected in Italian. The texts (provided in English) were translated by the authors.

Fig. 5 – Word cloud with positive and negative adjectives

The word-cloud shows that negative adjectives are more frequent than positive ones. Furthermore, most of the students specified if they are “Attentive” or “Inattentive” during mathematics class. The more frequent positive adjectives are “Attentive”, “Interested” and “Diligent” (frequency higher than 10), followed by “Skilled” (frequency equal to 7) and by other adjectives with frequencies lower than 6. Instead, as far as negative adjectives are concerned, “Inattentive”, “Unable”, “Lazy”, “Ignorant”, “Anxious” and “Bored” occur more frequently. It is interesting that all the adjectives that are linked with emotions (as anxious, nervous, etc.) appear in grey and are linked to negative emotions. In conclusion, the choice of the adjectives splits classrooms between attentive, diligent students and inattentive, lazy or bored students. It seems anyway that most of the students are experiencing some trouble with mathematics, as only negative emotions appear and the adjectives describing students’ skills and knowledge are mostly negative. With the aim to deeply inquire about the reasons for students’ mathematics difficulties, from the students’ point of view, we proposed the question: “What do you think are the reasons for your difficulties in Mathematics?”. This question was followed by some statements, such as “I am not intelligent enough”, “The teacher is too demanding” or “My study method is wrong”. For each statement students were asked to choose an answer from “Not at all” to “Very much”. The answers are collected in the graph below. Students describe themselves as interested in Mathematics and state that it is useful. In fact, only three students claim that their difficulty is linked with the fact that they make little effort because they are not interested in Mathematics, or they think that Mathematics is useless. On the other hand, 57 students, 71% of the sample, totally disagree with the sentence “I make little effort because I believe that Mathematics is useless”.

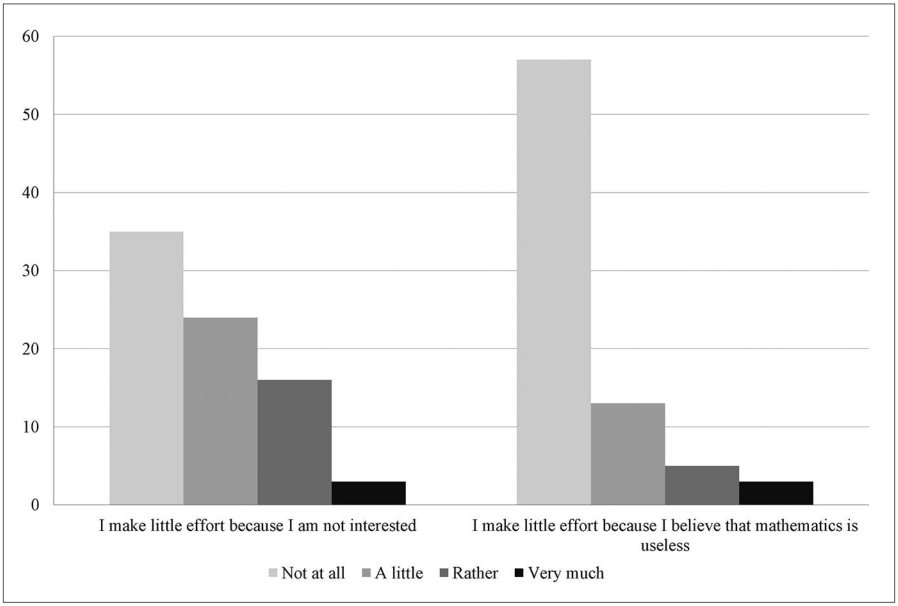

Fig. 6 – Some of the results of the question “What do you think are the reasons for your difficulties in Maths?”

These results are in tune with the broad social conviction that Mathematics is useful, as although most students are unable to explain why and to recognize whether they are effectively using Mathematical related knowledge (Niss, 1984). Moreover, 67 students do not attribute their difficulties to a lack of intelligence, rather to a difficult kind of intelligence.

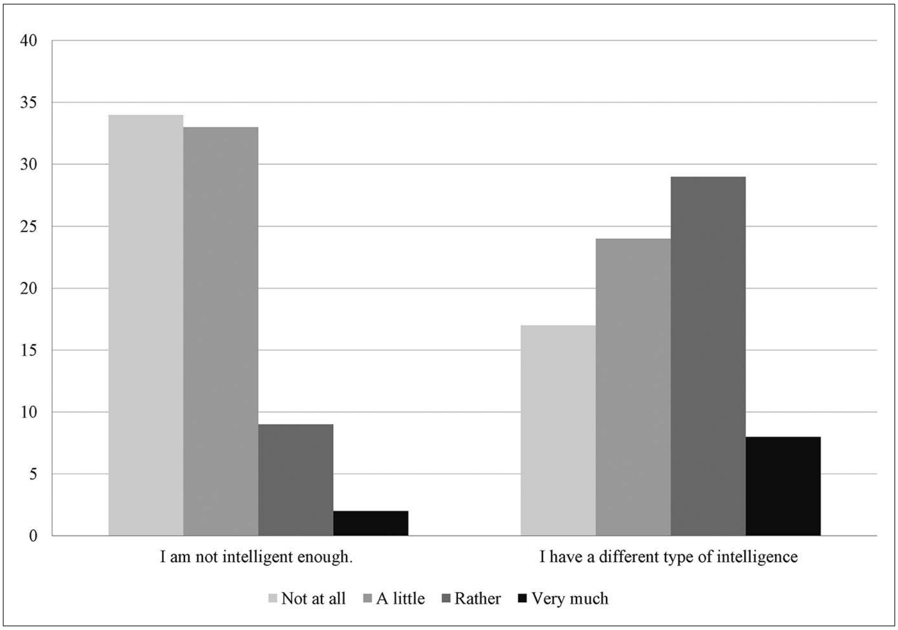

Fig. 7 – Some of the results of the question "What do you think are the reasons for your difficulties in Maths?"

The statement with the higher frequencies of students answering "Rather" or "Very much" is "I get anxious" ("Mi faccio prendere dall'ansia"). In fact, 46 students, 58.22% of the sample, answered 31 "Rather" (15 students) or "Very much" (31 students). This confirms the importance of the metacognitive factors (e.g. Capozio, Di Martino and Passaro, 2018) and supports the word-cloud inferences.

4.2. Remarks on the results of INVALSI tasks

4.2.1. First INVALSI task

The following graph (see fig. 8) represents how students answered the first proposed INVALSI task (which is represented in fig. 1).

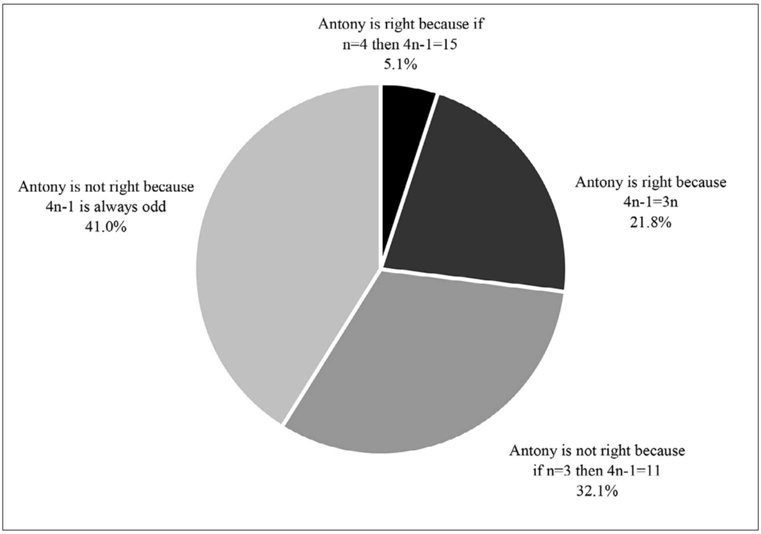

Fig. 8 – Sample students answer at the first INVALSI task

Sample students' answers are quite different from the answers at the national level, this is probably related with the sample choice. In fact, as stated before, the classes selected are from a specific High School and share the same curriculum.

In our case, the percentage of the correct answer is 32.1%, lower than the correct answer at a national level. Moreover, we can see that 21.8% of the students answer "Antonio ha ragione perchè 4n-1=3n" ("Anthony is right, because 4n-1=3n"), showing a difficulty with literal calculations. As an example, we report a student justification for this answer:

Anthony is right, because n=x, hence 4x-1=3x.

This previous response highlights a lack of control, both in the literal calculation and in the justification production. We can also interpret this answer as an example of a ritual justification (Harel and Sowder, 1998).

Above all, most of the students, 41.0% of the sample, choose the answer "Antonio non ha ragione perchè 4n-1 è sempre un numero dispari" ("Anthony is not right, because 4n-1 is always odd"). During the focus groups, we inquired about the motivation for this high percentage, and we identified two main reasons related to this kind of answer. The first reason, more widespread and expected, was the fact that students were focused on the statement's truth and they did not pay attention to the fact that the statement was not a valid justification for the answer. The following answer is an example of this kind of behavior:

> I tried to substitute the n with any natural number other than 0. For example, (4x6)-1 = 24-1 which is an odd number. Done with any other natural number, the explanation is always the same. (4x8)-1=32-1=31, odd number[4].

Though the student chose values for the *n,* from which follow numbers that are not multiple of three, his conclusion shows a focus on the parity of the numbers.

Secondly, some of the students showed difficulty in the interpretation of "4n-1", which they interpreted as "any odd numbers". The incorrect interpretation of the literal expression leads them to think "4n-1 is any odd number, and not all the odd numbers are multiple of three" and therefore to choose the related answer. In other words, this time we can identify problems related to the control of algebra expression, rather than difficulty with deductions.

4.2.2. Second INVALSI task

The following graph collects the students' answers to the second INVALSI task, the open-ended question represented in fig. 3.

We can observe that the percentage of uncorrected answers is extremely high (91.0% of the sample). Although approximately half of the students understand that Mark is not right, they are not able to produce a valid justification. We recognized two typical incorrect justifications to the statement "Mark is not right". The first one is of the kind "No, because no", as for example the answer "non ha ragione perchè non sempre è un numero primo" ("He is not right because it is not always a prime number"). The second argumentation is linked with the fact that n could be any natural number. A typical response in this case is "dipende che numeri ci sono al posto di n" ("it depends on which numbers are in substitution of n").

[4] "Ho provato a scambiare la n con un qualsiasi numero naturale diverso da 0. Per esempio (4x6)-1=24-1 che è un numero dispari. Fatto con qualsiasi altro numero naturale la spiegazione è sempre uguale. (4x8)-1= 32-1= 31, numero dispari" (authors' translation).

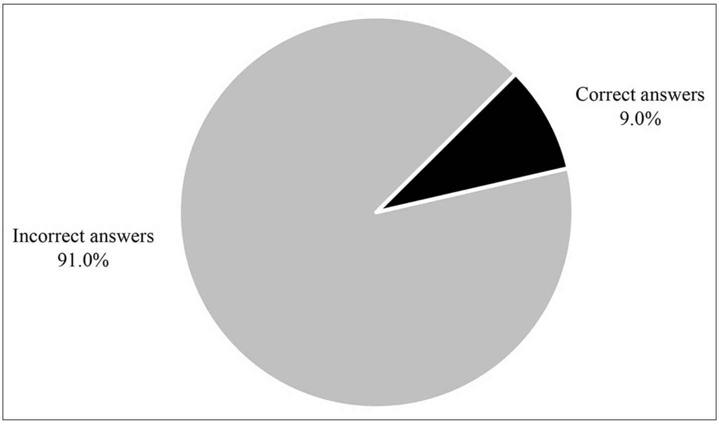

Fig. 9 – Students' answer to the second INVALSI task (open-ended question)

4.3. Observations related to the difficulty perception

Following the general presentation of the response at the metacognitive questions and at the INVALSI tasks, we can discuss the answers to those questions related to students' difficulty perception. The following is a three-step analysis. Firstly, we investigated the difficulty level that students attributed to the single INVALSI task and the difficult comparison of the two items. Secondly, we compared the difficulty perception levels and the fact that students answered correctly or incorrectly to the INVALSI tasks. Thirdly, we investigated the possible relations between the task difficulty perception and the preparation perception in addressing the specific task.

As far as the first step is concerned, we asked students to assign a difficult level at the INVALSI tasks after they solved it. The scale was 1 ("Molto facile" = "Very easy") to 10 ("Molto difficile" = "Very difficult"). The two following graphs show students' answers for the first INVALSI task and the second one respectively.

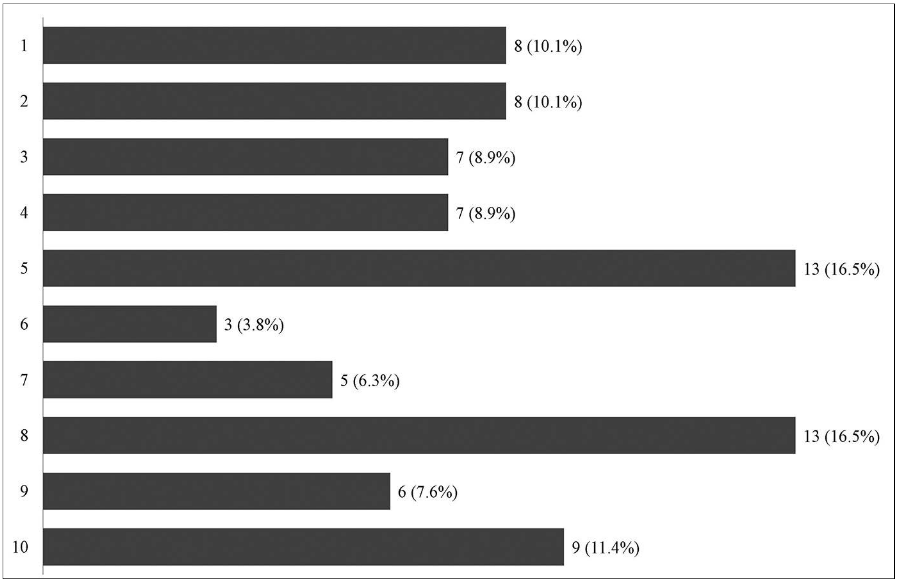

Fig. 10 – Students' answer to the question "On a scale 1 (very easy) to 10 (very difficult) how difficult did you find this task?" related to the first INVALSI task

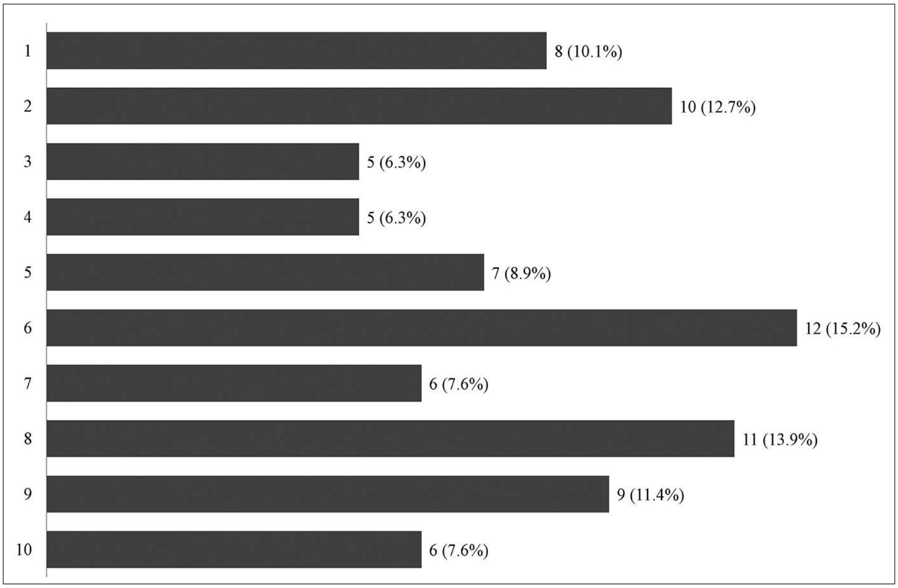

Fig. 11 – Students' answer to the question "On a scale 1 (very easy) to 10 (very difficult) how difficult did you find this task?" related to the second INVALSI task

We can observe that there is not such a difference between the two graphs: students are distributed among all choice options and the answer average is equal to 5.56 for the first item and 5.58 for the second one. So, although students' answers to the two INVALSI tasks were different in percentage, from these graphs does not emerge a difference with respect to the difficulty perception. In other words, the students' difficulty evaluation for each question (considered separately) do not seem to particularly differ.

Furthermore, at the end of the questionnaire, we asked students to compare the difficulty of the two INVALSI tasks they addressed, asking them directly which of the INVALSI tasks they had found more difficult. The figure 12 shows the responses to this question.

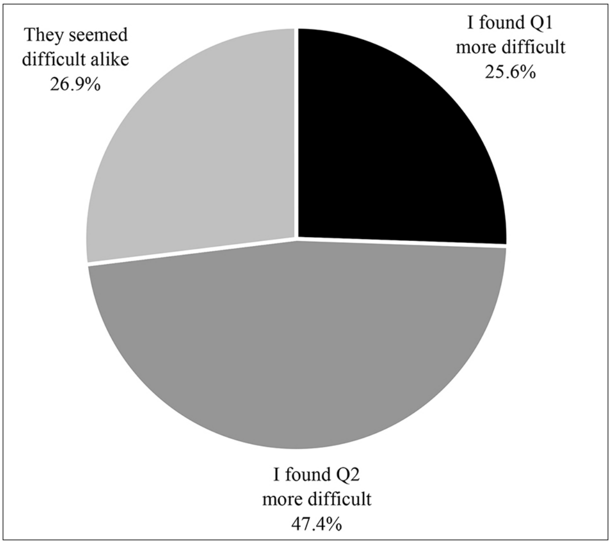

Fig. 12 – Students' answer to the question "Which of the two [INVALSI] tasks did you find more difficult?"

In this case, almost fifty percent of the students stated that they had found the second INVALSI task more difficult than the first one. Approximately one student out of four stated that he/she had found the first task more difficult and approximately one student out of four stated that he/she had found the two tasks difficult alike. Hence, in this case we reckon a difference between the difficulty perception of the two INVALSI tasks.

4.4. Remarks on the difficulty levels assigned to the tasks and comparison between the previous questions

Finally, we compared the answers of the three questions:

- (D1) On a scale 1 to 10, how difficult did you find this [first] task?
- (D2) On a scale 1 to 10, how difficult did you find this [second] task?
- (D3) Compare the two tasks you addressed during this test. Which of the two tasks did you find more difficult?

Investigating the consistency between the higher difficulty level attributed to the INVALSI tasks answering (D1) and (D2), and the (D3) response.

Figure 13 represents the result of the comparison. In the graph, students' responses are represented by bubbles. Each bubble is characterized by four factors: the x and y coordinates in the Cartesian plane, the texture, and the size. The x-coordinate represents the answer to the question (D1), in other words it indicates the level of difficulty attributed by a student to the first INVALSI task (Q1). The y-coordinate represents the answer to question (D2), i.e., the level a student has attributed to the second INVALSI task (Q2). The texture of the bubble indicates the answer to the question (Q3). The white colour is for students who responded that they found the first INVALSI task more difficult; the grey is for students who responded that they found the second INVALSI task more difficult; the strip-texture for those who responded "They [the two tasks] seemed difficult alike". Finally, the size of the bubbles represents the number of students who gave the same response. The smallest bubbles correspond to a frequency equal to 1, the largest bubbles to a frequency equal to 4. For example, the bubble with coordinates (10; 3) and the strip-texture indicates a student who assigned a difficulty level equal to 10 to the first INVALSI task (Q1), equal to 3 to the second INVALSI task (Q2) and stated that two tasks were equally difficult.

This graph highlights a behaviour that affects some students: a mismatch between the level of difficulty chosen during the evaluation of the single task (answers to D1 and D2) and during their difficulty comparison (answer to D3). If all the students answered consistently to these three questions, in fact, all the white bubbles would belong to the $x > y$ half-plane and all the grey dots to the $y > x$ half-plane, moreover all the strip-texture bubbles would be situated on the bisector. Looking at the graph we can easily see that this is not the case. Specifically, the behaviour concerns about half of the students (47.4%). In our opinion, this could be an indicator of the existence of students' issues in assessing the difficulty of one or more mathematical tasks. It could also indicate that the evaluation of tasks changes, depending also on the fact that the Task is evaluate individually or in comparison with others. Further investigations are surely needed to inquire deeper about this phenomenon.

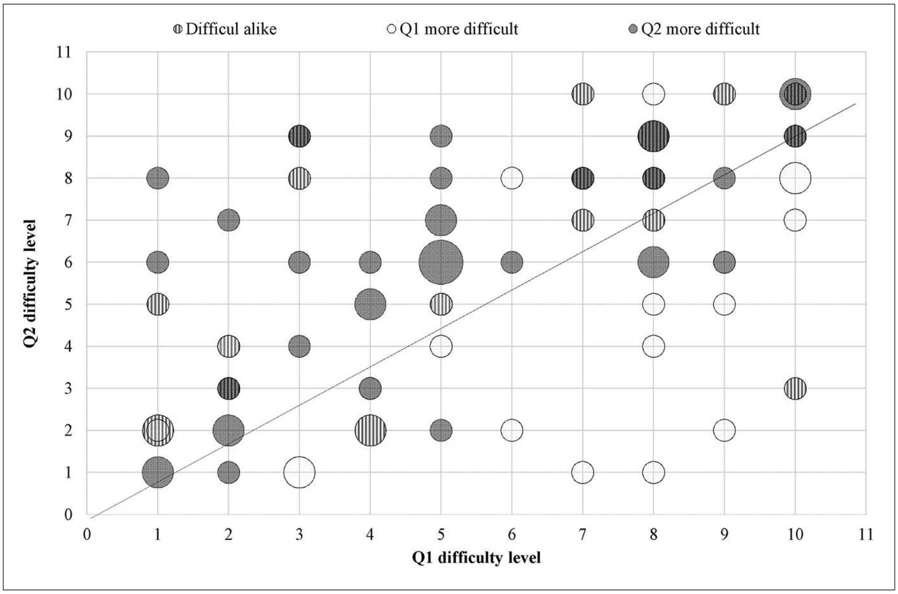

Fig. 13 – Comparison students' answers to D1, D2 and D3

4.5. Remarks on the possible relationship between correct/incorrect answers and difficulty perception level

Furthermore, we investigated the relationship between students' difficulty perception and the fact that they answered correctly or incorrectly to the two INVALSI tasks. In other words, we tried to inquire (investigate) if students tend to perceive a task easier if they answer correctly or vice-versa. In the following graphs, abscissa values are students' difficulty levels for the considered INVALSI item (Q1 or Q2). The light grey bars represent the correct answers for each difficulty level and the dark grey ones the uncorrected answers.

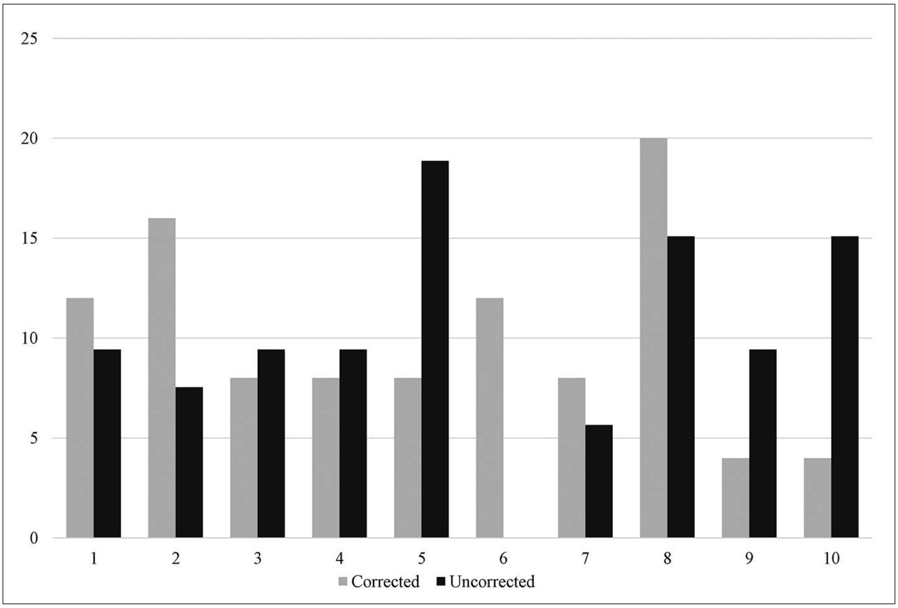

Fig. 14 – Comparison between correct/uncorrected answers and difficulty perception levels for the first INVALSI task

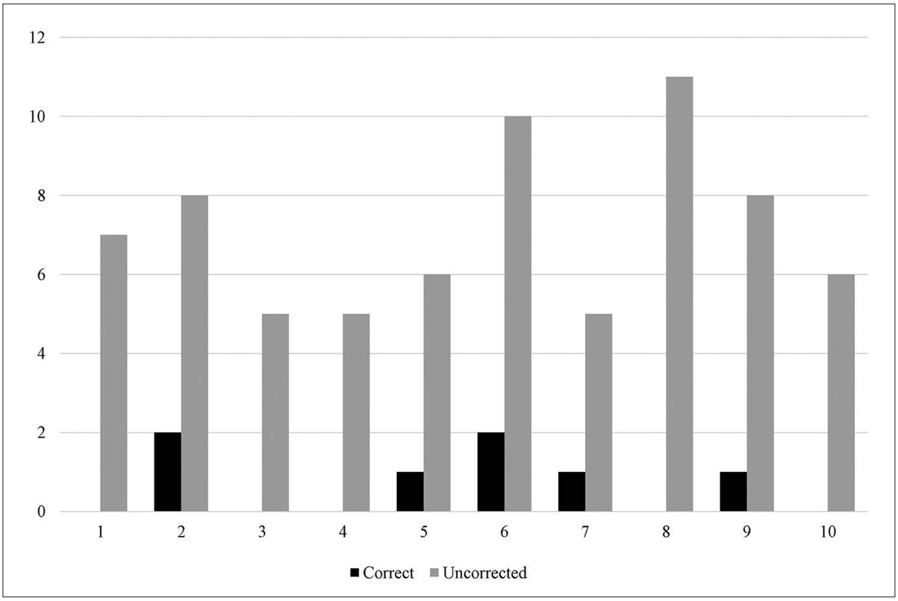

Fig. 15 – Comparison between correct/uncorrected answers and difficulty perception levels for the second INVALSI task

Examining this graph it seems that there is no correlation between the difficulty perception level and the fact that the student's answer is correct or uncorrected. In other words, from an exclusive qualitative point of view, answering correctly or incorrectly does not influence the difficulty perception of the student.

4.6. Remarks on the possible relationship between perceived difficulty and perceived preparedness (task 1 and task 2)

On the other hand, we also asked students to attribute a level to their preparation. That is, we asked them "Do you feel prepared to answer this question?" from a scale 1 (I am not at all prepared) to 5 (I am fully prepared).

We hence compared the perceived preparation level with the difficulty perception for each student and for each INVALSI task. The following graph shows the results of such comparison. On the x-axis, we find the difficulty perception level (one to ten). Each color represents the answer of the preparation level, the lighter color stands for lower preparation level (1) and the darker color stands for the higher preparation level (5), as shown in the legenda.

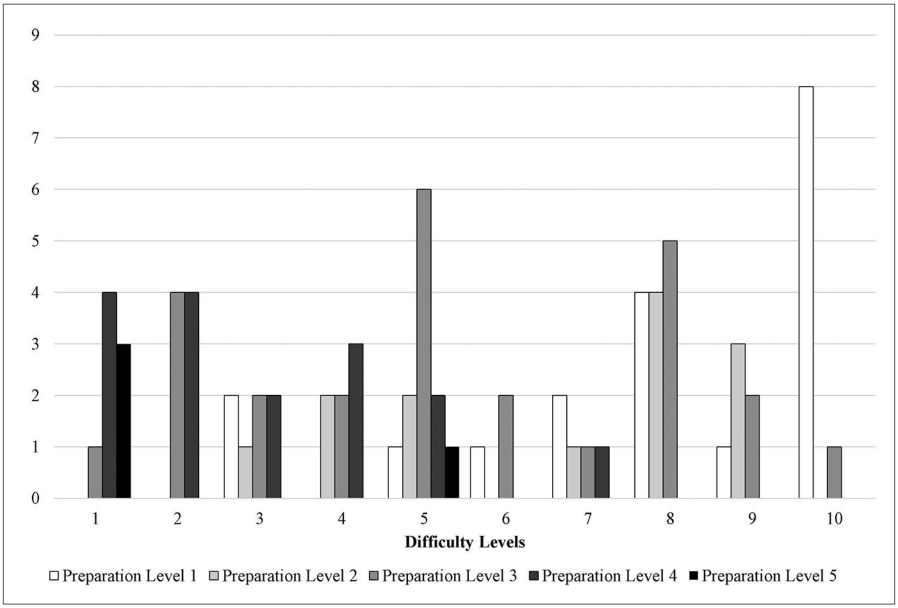

Fig. 16 – Comparison between difficulty levels and preparation levels for the first INVALSI task

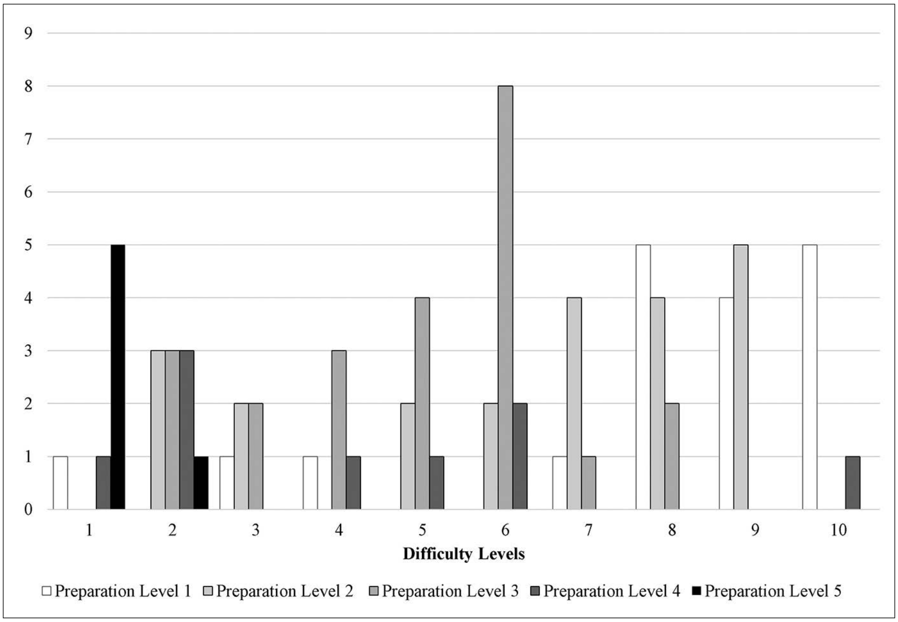

Fig. 17 – Comparison between difficulty levels and preparation levels for the second INVALSI task

We can hence notice that preparation levels one or two appear more likely in correspondence to high levels of difficulty perception. On the contrary, preparation levels equal to four and five decrease for difficulty perception levels increasing. Finally, we can notice that a preparation perception equal to three is more or less stable for all the difficulty perception levels.

It seems that, except for some exceptions, students who think not to be prepared to answer the question find the INVALSI task more difficult, vice-versa, students who state that they are prepared find the question easier.

5. Discussion and concluding remarks

A first partial, but remarkable conclusion is related to the fact that students' perceived difficulty seems to depend more from metacognitive factors, rather than being related with the (un) correctness of the answer. We hypothesize that the perceived difficulty is more related to having faced similar questions in class, and similar to something I've seen before doesn't mean the student can handle it.

In assigning a level of perceived difficulty to a question, students may be influenced by:

- **metacognitive aspects**: lack of capacity to judge their own skills, knowledge and abilities;
- **factors linked more strictly to the task**: elements of the text, elements of the images, etc.

Specifically, our study revealed more the first ones, and less the second ones (which there are already results in the literature). We believe, however, that it may be interesting to investigate the perceived difficulty even before solving the task and to relate it to the perceived difficulty after solving it.

Next step (always related to the qualitative phase), will consist in identifying categories of factors that influence the perceived difficulty (dependence on school practice, dependence on the importance of the topic, dependence on content, etc.). These indicators will be fundamental in the large-scale analysis (phase 2) to understand the weight of every factor. In fact, starting from the qualitative phase results, we will build an adaptive questionnaire. This second phase will take place during this school year. In particular, we want to investigate how and to what extent the difficulty perceived by students is connected to the difficulty that is attributed to the task by the teacher and any relationship with the ability INVALSI levels. Furthermore, we will inquire whether it is possible to arrange the question in order of difficulty, varying some question elements.

In conclusion, phase 2 of our study will clarify some of the open problems we discussed above:

- to study student perceived difficulty before and after addressing a question;
- to quantitatively study the perceived difficulty;
- to change the type of questions and choose simpler difficulties;
- to inquire the possibility of sorting questions by perceived difficulty;
- to compare the student's perceived difficulty level with the INVALSI difficulty level;
- to compare the student's perceived difficulty level with the teachers' perceived difficulty level;
- to investigate whether the construct of students' perceived difficulty can give us information about students' enacted resolution processes (or vice versa).

References

Baccaglini-Frank A., Di Martino P., Natalini R., Rosolini G. (2018), *Didattica della matematica*, Mondadori, Milano.

Bagni G.T., D'Amore B. (2005), "Epistemologia, sociologia, semiotica: la prospettiva socio-culturale", *La matematica e la sua didattica*, 1, pp. 73-89.

Bolondi G., Branchetti L., Giberti C. (2018), "A quantitative methodology for analyzing the impact of the formulation of a mathematical item on students learning assessment", *Studies in Educational Evaluation*, 58, pp. 37-50.

Capozio A., Di Martino P., Passaro D. (2018), "'Io e la matematica': un'indagine sull'esperienza matematica", *Didattica della matematica. Dalla ricerca alle pratiche d'aula*, 4, pp. 9-26.

D'Amore B. (2000), "Lingua, matematica e didattica", *La matematica e la sua didattica*, 1, pp. 28-47.

De Corte E., Verschaffel L., Van Coillie V. (1988), "Influence of number size, problem structure, and response mode on children's solutions of multiplication word problems", *J. Math. Behav.*, 7, pp. 197-216.

Harel G., Sowder L. (1998), "Students' proof schemes: Results from exploratory studies", *American Mathematical Society*, 7, pp. 234-283.

Laborde C. (1995), "Occorre apprendere a leggere e scrivere in matematica", *La matematica e la sua didattica*, 9 (2), pp. 121-135.

Mc Leod D. (1992), "Research on affect in mathematics education: a reconceptualization", in D. Grows (*ed.*), *Handbook of Research on Mathematics Teaching and Learning*, McMillan Publishing Company, New York, pp. 575-596.

Niss M. (1994), "Mathematics in society", *Didactics of mathematics as a scientific discipline*, 13, pp. 367-378.

Thevenot C., Oakhill J. (2005), "The strategic use of alternative representation in arithmetic word problem solving", *Quarterly Journal of Experimental Psychology*, 58, pp. 1311-1323.

Vicente S., Orrantia J., Verschaffel L. (2007), "Influence of situational and conceptual rewording on word problem solving", *British Journal of Educational Psychology*, 77 (4), pp. 829-848.

Zan R. (2007), *Difficoltà in matematica. Osservare, interpretare, intervenire*, Springer, Milano.

Zan R., Baccaglini-Frank A.E. (2017), *Avere successo in matematica. Strategie per l'inclusione e il recupero*, UTET Università, Torino.

The authors

Francesco Avvisati is Senior Survey Specialist (Chief Methodologist) in the Directorate for Education and Skills at the OECD.

Barbara Baldazzi, researcher in Socio-Demographic Statistics at ISTAT since 1997. Researcher in SDGs project: "Sustainable Development Goals" with the responsibility of Goal 3 (health and well-being), Goal 4 (education). Researcher in the BES project "Measurement of equitable and sustainable well-being in Italy" with coordination of Education and Training sectors. Project Manager of "Adult Education Survey" (2017 and 2012).

Caterina Balenzano, PhD in Psychology. She is Assistant Professor in General Sociology at the University of Bari – Department of Political Science and studies social protection and support policies, programs and 53 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.

Sergio Di Sano, PhD in Developmental Psychology. He is an Assistant Professor of School Psychology at the University of Chieti-Pescara, Department of Neuroscience, Imaging, and Clinical Sciences. His research focuses on learning and adaptation processes in the school environment, with a specific emphasis on reading skills, school climate and digital wellbeing. He collaborates with schools on improvement projects based on participatory research-action.

Sara Manganelli, Phd, is a researcher at INVALSI. She has worked in IEA PIRLS, TIMSS, ICCS international comparative surveys, and in national projects about school evaluation. She has conducted several studies for developing psychometric instruments and analysing large-scale datasets with multilevel techniques and structural equation modelling.

Rita Marzoli is director of the INVALSI Library and Documentation Centre in Rome. She is editorial board coordinator for "INVALSI per la ricerca", a double-blind peer-reviewed open access series published by FrancoAngeli. She is research specialist librarian in education and social sciences. Her research interests include information retrieval in the social sciences, systematic review methods, research evaluation, information literacy, bibliometrics, open access.

Mariagiulia Matteucci is an Associate Professor at the Department of Statistical Sciences, University of Bologna. Her research interests are item response theory models, Bayesian psychometrics, and educational measurement.

Stefania Mignani is a Full Professor at the Department of Statistical Sciences, University of Bologna. Her research interests are in the field of latent variable models and statistical methods for social and educational phenomena, with a particular attention to competence assessment.

Ornella Papa is a specialist in Psychological Evaluation and researcher at the INVALSI Library. She dealt with the OECD PISA Projects, Evaluation and Improvement, INVALSI Tests about Special Educational Needs. She deepened the role of School Libraries for student achievement and developing information literacy. Her current research interests are in the field of educational equity, the digital divide and learning environments.

Marta Saccoletto, Bachelor in Mathematics, she is a doctoral student in Mathematics Education at the University of Torino. Her research interests include argumentative processes in and for mathematics education, and (formative) assessment activities.

Giada Spaccapanico Proietti, Digital Transformation Specialist and Data Scientist at Emilia-Romagna Region. Former postdoc fellow at the University of Bologna, her research interests were computational statistics, automated test assembly, and computerized adaptive testing.

Camilla Spagnolo is a researcher in Mathematics Education at the Free University of Bozen/Bolzano. Her main research interests concern argumentation processes in mathematics, teaching based on competence and teachers' training. She is an Adjunct professor at the University of Bologna and member of the Nucleo di Ricerca in Didattica della Matematica (NRD). She is involved in teacher training courses for all school levels.

Nathan Viltard is a PhD student at the Paris School of Economics (EHESS) and former intern in the Directorate for Education and Skills at the OECD.

For several years, in Italian school there are fewer and fewer students; this trend is offset neither by the increase of schooling nor by the rise of students with an immigrant background. In addition to this issue, there is also the heavy problem of school dropout. Studies and researches on students achievement are therefore highly relevant because they may focus on their difficulties and strengths and could provide ideas and suggestions to help them.

This book collects six papers submitted during the V edition of the Seminar "INVALSI data: a tool for teaching and scientific research" held in Rome from 25th to 28th February 2021, with the aim of detecting which variables affect are determinant for a proper everybody's educational achievement stressing national INVALSI survey and other international ones.

As Statistical Service we hope this partnership could last and may generate many other research works.

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.

FrancoAngeli
La passione per le conoscenze