# VIII SEMINAR

INVALSI

## "INVALSI data: a tool for teaching and scientific research"

dati invalsi
uno strumento
per la ricerca
e la didattica

**Rome**
**23rd - 26th**
**November 2023**

# INDEX

# INTRODUCTION

The Seminar "INVALSI data: a tool for teaching and scientific research", now at its eighth edition, has become in recent years an opportunity for meeting and discussion on the use of National Surveys INVALSI results and, in general, on the world of evaluation and school. Evaluation does not only mean standardized monitoring of learning levels, but also evaluation and comparison of its functions and potential in relation to the practices and tools through which the school system is able to carry out its educational and formative mission. This year, keynotes of national and international experts were organized on issues related to the evaluation of the education system and the use of data to support school policies.

One of the objectives of the seminar was to bring the world of scientific research and schools closer together in order to build a privileged and concrete space in which enriching the debate by sharing ideas and experiences between education and research stakeholders. The variety of topics improve a multidisciplinary approach to evaluation in the educational and school field, trying to give an account of the contribution that schools can provide to society by making possible the development of knowledge and skills.

The Seminar is organised by the research group of Area 2 - Statistical Service: Patrizia Falzetti (Manager), Paolo Barabanti, Andrea Bendinelli, Leonardo Boulay, Emiliano Campodifiori, Michele Cardone, Federica Collia, Doriana delli Carri, Paola Giangiacomo, Patrizia Giannantoni, Pierangelo Grosso, Jana Kopecna, Fabrizio Lasorsa, Giuseppina Le Rose, Francesca Leggi, Lorenzo Maraviglia, Michele Marsili, Giancarlo Mastrone, Silvia Mazzuca, Carlo Palmiero, Monica Papini, Veronica Pastori, Antonio Severoni, Valeria F. Tortora.

Contacts: uff.statistico@INVALSI.it

# ABSTRACT

## THEME 1. A DIFFICULT EDUCATION: CHALLENGES OF THE SCHOOL SYSTEM TO ENSURE SOCIAL INTEGRATION, SKILLS/KNOWLEDGE AND SOCIAL MOBILITY

**ORGANIZER: INVALSI – ESPANET**
**COORDINATOR: PATRIZIA FALZETTI**
**NOVEMBER 23RD: 2.00 P.M. – 4.00 P.M. {ROOM 1 – RESEARCH SESSION 1}**

---

### Analysing INVALSI data to develop a simplified index of students' socioeconomic and cultural background

**Clelia Cascella**

Introduction. In education research, previous studies have indicated that students' socioeconomic status (SES) is a very strong predictive factor of attainment. Reardon et al. (2013) showed that students from low-SES families enter high school with average literacy skills 5 years behind those of students from high-income families. As with Reardon, Baird (2012) showed that "on average, math scores of students with indicators of high-socioeconomic status (SES) are over one standard deviation above those with low SES indicators". Therefore, measuring students' SES properly is to be considered as a priority for both the (educational) policy and the research agenda.
To this end, a number of different measures have been proposed over time, based on similar but not on the same indicators. The selection of different indicators is due to both theoretical and methodological reasons (e.g., to different conceptualisations of socioeconomic status and thus to different operationalisations of such a concept), but also to the availability (or unavailability) of some indicators in some socio-cultural contexts (e.g., research measuring SES by including family income may be not replicated in studies carried out in contexts where people are less likely to share such a sensitive information).
Nonetheless, previous studies have clearly shown that research based on different measures of SES have produced different and somehow contradictory results thus hindering the possibility to advance knowledge about the relationship between students' SES (and its possible sub-dimensions) and students' attainment.
The current paper aims to develop a simplified measure of students' SES, easy to be replicated in further research, and thus able to foster results comparability.
Among the existing measures of SES, the Economic, Social and Cultural Status index (ESCS) is one of SES indices most frequently used, in secondary analysis, in educational research (Avvisati, 2020). ESCS was developed by the Organization for Economic Cooperation and Development (OECD) (Nardo et al., 2008; OECD, 2012, 2015, 2017) and it is used to measure SES among students involved into the Programme for International Students Assessment (PISA).
Since OECD-PISA survey involves students from around 200 countries, ESCS index has been thought to allow for the comparability of results across different cultures and geographies. As with most of research about the factorial structure of students' socioeconomic status (for a review, among many others, see for example Sirin, 2005), ESCS is based on three main sub-dimensions, that are (i) parental education, (ii) parental occupation, and (iii) family wealth. Both the first and the second dimension are measured via indicators relatively easy to be collected (as they can be provided for example by school secretariats). Moreover, the former (i.e., parental education) is usually measured in years of schooling or in ISCED levels, that have been thought and developed to allow for the comparability of such an information across different sociocultural contexts. The latter (i.e., parental occupation) can be transformed into a comparable measure by using, for example, the framework proposed by the OECD, used in both international and national surveys, such as that carried out every year by the Italian national institute for the evaluation of educational system (Tr. Istituto nazionale per la valutazione del sistema di istruzione e formazione, INVALSI) (Campodifiori et al., 2010). In contrast, the third dimension (i.e., family wealth) is measured via indicators that are not as easy to be collected as those used to measure parental education and occupation. To collect such a kind of information, OECD administers an ad hoc questionnaire, that raises up both organizational and economic costs that may be not affordable to smaller research teams and thus call for the development of alternative SES measures.

Previous research has deeply discussed the contribution of each indicator to the measurement of students' SES. By a side, even though there is an almost unanimous agreement that the higher/the more prestigious the mother's and/or father's education and/or occupation, the better the students' attainment, the extent of such an association may change depending on how these factors are measured and/or combined. On the other side, recent studies have critically reviewed the indicators selected by the OECD to measure "family wealth" and raised up several concerns about their validity. For example, Pokropek et al. (2017) claimed that the variable "homeposs", an index used to measure the possession of some goods that can be considered as a proxy of family wealth, suffer from a lack of comparability over time and across countries: for example, having access to the Internet and possessing software mirrors a completely different wealth in 2000 and in 2012, and/or in industrialised words compared with not industrialised countries.
Research aims. The current paper aims to develop a simplified SES measure, easy to be replicated in further studies. Therefore, the ESCS, that is the SES measure most frequently used in educational research, has been used as a starting point to (i) explore the association between the indicators used to construct it and students' attainment both in mathematics and in text comprehension; and, (ii) develop a simplified version of the ESCS whose predictive power of attainment is as strong as the original measure.
Data. Data collected by INVALSI at grade 5 in 2015 have been analysed for the purposes of the present study. Sample (instead of census) data have been used to work with data net of possible cheating that usually occurs in large-scale assessment (Falzetti, 2013; Longobardi et al., 2018).
Method. The analytical approach is a three-steps strategy.
Step 1. INVALSI data are hierarchical: students are nested into classrooms, classrooms into schools and schools into Provinces and Regions. To account for data hierarchy, INVALSI data were thus analysed via a multilevel analysis (Hox, 2010). Several models have been estimated via an iterative procedure aimed at exploring the relationship between students' attainment and (i) each SES indicator and/or (ii) alternative combinations of SES indicators.
Step 2. Starting from the results from the multilevel analysis, the proposed simplified socio-cultural index (SC-index) has been developed. It is a typological index that combines the highest parental education and occupation.
Step 3. Finally, since the ESCS has been used in the current research as a benchmark to assess SC-index's validity, two further regression models have been estimated to compare their association with students' attainment. The comparability of results based on the employment of the SC-index and of the ESCS has been critically discussed.
Results. Results showed that both mother's and father's education and occupation are strongly (and to the same extent) associated with students' attainment, thus supporting the decision to base ESCS on the highest parental education and parental occupation. Moreover, results showed that the association between the variable 'HOMEPOSS' and students' attainment is statistically significant but very small. Finally results from the multilevel regression showed that the predictive power of the simplified index is as strong as that of the original measure.
The simplified index is not to be considered as "better" than the original one as the latter includes more information about family wealth that can be of help to deepen our understanding about the relationship between students' attainment and their families' background and properly channel both educational policy and practice. Nonetheless, the proposed index (SC-index) shows some important advantages as it is (i) easy to be replicated in different sociocultural context thus allowing for results comparability; (ii) based on parental education and occupation as measured via the ESCS and thus results based on the proposed index may be, to some extent, comparable with results based on ESCS.

***Keywords:*** socioeconomic status, attainment, gap

## Resilient students: what factors behind the success of disadvantaged students?

**Silvia Duranti – Francesco Bogazzi**

The relationship between socio-economic status and school achievement is well documented and extensive literature indicates that students from more advantaged backgrounds perform better at school tests. Despite this relationship, international assessments have highlighted that in practically all OECD countries

there are a relevant number of "resilient students", i.e. students from a disadvantaged socio-economic background who achieve relatively high levels of performance in terms of education (OECD, 2010). Performing analyses over a long time span, Longobardi et al. (2018) also reveal that several countries were able to increase the share of resilient students over time, reflecting improvements in the average performance of students, or a weaker relationship between socio-economic status and performance. Some papers have investigated the phenomenon of resilient students in the Italian school system, focusing both on lower and upper secondary education. For example, Agasisti et al. (2016) use INVALSI micro-data to focus on class and school-level factors that help disadvantaged students become resilient when they switch from primary (grade 5) to lower secondary school (grade 6); using a probit regression model and a propensity score matching, they find that the resilience of disadvantaged students is influenced by some school and class factors, for example peers' outcomes. Agasisti and Longobardi (2014) have a different focus, analyzing the factors favoring the resilience of students in upper secondary education. Using data from OECD-PISA to estimate a multilevel logistic model they find that individual-level characteristics play a role, but also that some school factors are relevant.
Data and objectives. In this paper, we use micro-data provided by INVALSI to focus on individual, class and school-level characteristics that help disadvantaged students to become resilient. To do so, we use data coming from the student, teacher and principal questionnaire.
Methodology. The methodology used consists of three step. As a first step, we construct a database made up of all students attending grade 8 in school year 2018/2019, merged with the information on principals and teachers databases. As a second step we will test different methodologies for the identification of "resilient students", starting from those proposed in the literature. For example, according to Longobardi et al. (2018) resilient students are those among the 25% most socio-economically disadvantaged students in their country but are able to achieve at or above "Level 3" in PISA test, a level that equips them for success later in life. Differently, OECD (2010) identifies resilient students as those with a residual performance (estimated through a regression of PISA test score on the socio socio-economic index) amongst the top quarter of students' residual performance. As a third step, we will analyze the determinants of resilience of students at the end of lower secondary education using a multilevel logistic approach with covariates coming from the student, teacher and principal questionnaire. Will provide information on factors and practices which can help students to succeed notwithstanding a poor socio-economic background.

***Keywords:*** educational equality, students' resilience, school factors, student factors

## Implicit school drop-out: personal and contextual features of a complex phaenomenon

**Maraviglia Lorenzo – Ornella Papa – Giuseppina Le Rose**

In recent years, also as a consequence of goals and benchmarks set at UE level, the issue of early school drop-out has drawn the general attention of Italian researchers and decision makers, in addition to schools. Systematically associated with a disadvantaged socio-economic background of students, early school leaving is a problem of worrying dimensions in our country. Unfortunately, just as a high percentage of students is involved in the phenomenon of explicit dropout, an equal percentage of students, despite completing their school curriculum, does not possess the minimum skills necessary to participate adequately in the society, to undertake university studies or to enter into the labour market. The implicit dropout, on which this study is focused, is a less evident and therefore more subtle phenomenon that has been recently coupled with explicit dropout thanks especially to INVALSI contribution and data.The drop out has been analysed in our country especially from the point of view of the social, cultural, and economic features of the students most exposed to the risk of dropping out. Less attention has been paid to the analysis of the characteristics of the schools and of the territorial contexts where dropout occurs with greater intensity. In terms of geographical variability, for example, much emphasis has been placed on the difference between macro-areas of the country (North-West, North-East, Centre, South, South and Islands) and between regions; however, other potentially relevant aspects, such as those relating to the differences in the incidence of early school leaving between central and peripheral areas of the same city, or between plan and mountain areas of the same province or a region, have been explored to a lesser extent. Implicit

dispersion is a phenomenon conceptually related to explicit dispersion, but in ways that, however, need to be in depth analysed. At the level of schools or territories, for example, the incidence of implicit dropout could go hand in hand with that of explicit dropout, or it could constitute a channel through which, in some way, complex territorial systems try to compensate the problem of early school leaving. These aspects must first of all be analysed through a reconstruction of the profiles of the students who, at the various levels of the educational path, lack the minimum skills required by the system; secondly, the characteristics of these students must be compared with those that in the literature are defined as "precursors" of explicit dispersion, in order to verify the consistency or, conversely, the discrepancy between these potential orders of causes; finally, the incidence of implicit dropout phenomena are compared in different types of schools at territorial level – in order to understand in more detail the interaction between them.
The aim of this research is to deepen the understanding of the phenomenon of implicit dropout: how it manifests itself at the level of individual students (personal, social, cultural, economic characteristics), of different school contexts (spaces, resources and internal organization, different type of school for the 13th year) as well as of different local contexts, as specified above. The characterization of implicit dropout is, as mentioned, one of the prerequisites for the composition of a broader interpretative framework of the functioning and dysfunctionality of the Italian school system – especially where it is most called upon to improve the results and resilience of students most vulnerable to school failure.
The main data used are the INVALSI microdata (8th grade and 13th grade) for the entire students' population of the 2021/2022 school year (for all available subjects: Italian, Mathematics, English Reading and Listening) and the data from Student Questionnaire. In order to better characterize the school and territorial contexts, in addition to the data of the School Principal Questionnaire, various additional statistical data are used from source ISTAT, including the very recent ones of the last population census (Censimento Continuo) and the last census of industry and services.
The analysis starts with a detailed reconstruction of the existing literature on the precursors of explicit early school leaving and of that - less wide but constantly expanding - on the precursors of implicit early school leaving. Moving from this basis, pertinent variables and proxies are identified within the available databases and analysis is carried out with the support of multilevel regression models. The multilevel structure is needed in order to be able to appreciate the impact of the territorial context on the dynamics that form the object of investigation. The analyses carried out highlight a very articulated picture that confirms the importance of both individual and context factors, especially with regard to the socio-economic and cultural background (ESCS Index) but identify in some characteristics of schools (for example adequate spaces, resources and, for the thirteenth grade also the different types of schools) the possibility of increasing resilience and improving student results. The distribution of implicit dispersion partly follows that of explicit dispersion but, at the same time, highlights peculiarities – for example from a geographical, citizenship and socio-cultural point of view – that require further investigation. In this perspective, detailed indications and suggestions are provided on the paths to follow and on the needs that emerged from this first exploration, underlining the need for even greater integration between different data, in order to reconstruct processes that are intrinsically very complex.

***Keywords:*** implicit drop out, explicit drop out, contextual features, ESCS

## Teacher insecurity and inequalities in student learning. New evidence from INVALSI data, 2017-2022

**Giovanni Antonini – Gianluca Argentin**

Schools are assigned the task of promoting equal opportunities and merit, allowing all students to pursue the paths that society offers, without distinctions of social and economic origin, let alone territorial; it is even hoped that schools will be able to give more opportunities to those who have less, for example because of a more difficult family or geographical background. The aim of this paper is to reflect on whether the potential equalising function of the Italian school is actualised by focusing on the 'teachers' factor. More precisely, we want to investigate a possible hidden mechanism of reproduction of inequality by the education system, i.e. the allocation of students to teachers with temporary contracts. One of the historical difficulties of our education system is in fact the widespread precariousness of teachers and a consequent

turnover of teachers during the course of their studies. This results in changing school staffing levels and educational discontinuity for students. The situation has its origin in the paradoxical coexistence of: i. the precarious status of a large proportion of teachers, employed on annual contracts or on shorter substitute contracts directly employed by schools; ii. the difficulty in filling certain types of teaching positions, especially for STEM subjects; iii. the seat transfers requested by tenured teachers (Abbiati et al., 2021). Several empirical studies have assessed how teachers' preferences are strongly influenced by non-pecuniary factors such as student characteristics, and from this it follows that teachers demand to move from more problematic schools to easier settings (Barbieri et al., 2011) and how a higher propensity to leave has a negative impact on student learning (Barbieri et al., 2017). A possible mechanism of perpetuation of inequality in the functioning of the Italian school system is thus configured, with the best teachers concentrated in the most favourable contexts and the others in the most disadvantaged ones (Abbiati et al., 2017). The aim here is to enrich this line of research, expanding the scope of previous studies both on the methodological level and in terms of understanding the processes at work.
Research object and hypothesis. The object we analysed is the assignment of unstable teachers to students and the consequences this condition has on students' learning. Starting from the present literature, and using the available data, we test the following research hypotheses with much larger sample sizes than those employed to date (see below):
H1: teachers who are not stable (temporary contractually or in the school) at work level are more likely to be assigned to socially disadvantaged students, thus with lower values of the ESCS indicator of socio-economic status.
H2: We expect this allocation to operate at both school and classroom level, within schools; this is an issue that has not been adequately investigated previously, despite the relevance of breaking down the two processes.
The question arises as to whether the two phenomena investigated vary in intensity according to school grades, without having strong a priori assumptions.
Data. The data used in this work come from the combination of data referring to teachers with that referring to students, both originating from INVALSI. In the analysis implemented, advantage is taken of the fact that students' data are present with an identifier of their school class and that, associated with this, are the answers provided on a voluntary basis to in-depth questionnaires by teachers of Italian and Mathematics. The database used is original, since data from the national sample for the school years 2017/2018, 2018/2019, 2020/2021 and 2021/2022 are jointly analysed, in relation to all the tests administered in compulsory schooling, i.e. the second and fifth grades of primary school (grades 2 and 5) and the third of secondary school (grade 8). The analysis, therefore, has the merit of combining teachers' data with students' data for four years, generating an extremely large sample to study the phenomenon. The multilevel structure of the data (students, classes, schools) makes it possible to estimate the level of concentration of students with socio-economic backgrounds in classes and schools by associating it with the temporary status of teachers.
Methods. We use two analytical approaches:
1) Descriptive analysis: all data from the sample population were analysed, correlating the characteristics of students (in particular their ESCS), classes and schools with the temporary status of teachers, in order to understand whether there are any matches between specific characteristics. Both bivariate analyses and regression models are used here. These analyses are also conducted separately by school grade and discipline taught (Italian and mathematics).
2) Estimation of the effects of having a temporary teacher: we carried out both correlational analyses using multivariate regression models, separately for Italian and mathematics, and analyses using fixed effects models on the students. We exploit the variation within students between subjects in the condition of having a temporary teacher, studying the consequences this has on test scores. These analyses are also conducted separately for school grades.
Results. Our analyses also show that the teacher-student pairing is not as neutral as one would be led to believe, especially in secondary school. The evidence therefore suggests that in Italy, too, the unequal assignment of students to teachers may be a hidden form of educational inequality. This process seems to be at work not only at school level but also at class level. The situation seems to have negative effects on learning more for those from less well-off families. Overall, teacher insecurity and its mismanagement seem to be a piece in the processes of school reproduction of educational inequalities.

***Keywords:*** precariousness, teacher-student matching, inequalities in learning

# THEME 2. DIFFERENT DATA SOURCES AND THEIR INTEGRATION: POLICY INDICATORS FOR THE EDUCATION AND TRAINING SYSTEM

**ORGANIZER: INVALSI – ISTAT**

**COORDINATOR: BARBARA BALDAZZI**

**NOVEMBER 23RD: 2.00 P.M. – 4.00 P.M. {ROOM 2 – RESEARCH SESSION 2}**

---

## Measuring early school leaving: how to get institutional data and INVALSI data to talk to each other

**Massimo Armenise – Barbara Baldazzi**

Introduction. Reducing early school leaving is one of the European Union priorities. In fact, this phenomenon generates serious repercussions on young people and society in general: greater difficulty in finding a job, limited employment prospects, lower participation in social, political and cultural activities; increased risk of poverty and poor health (Istat 2020).
School dropout, in fact, is an obstacle to economic growth and employment, holding back productivity and competitiveness and fueling poverty and social exclusion.
According to ISTAT, the percentage of young people between the ages of 18 and 24 who have dropped out of education and training prematurely (achieving only a lower secondary school diploma) turns out to be 11.5 percent in Italy, or about 465,000 young people, in 2022. Early school leavers involves more boys (13.6 percent) than girls (9.1 percent); regions in the South (15.1 percent) than those in the Center (8.2 percent) and North (9.9 percent).
At the European level, Italy, despite showing a marked improvement, remains among the countries with the highest incidence of early school leavers, followed by Germany, Hungary, Spain and Romania. This indicator is one of the benchmarks (9 percent) in the new Strategic Framework for European Cooperation in Education and Training. See Council Resolution on a strategic framework for European cooperation in education and training towards a European educational area and beyond (2021-2030), 2021/C 66/01 (OJ C, C/66, 26.02.2021, p. 1, CELEX:
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32021G0226%2801%29).
The strong disparities in dropout rates that emerge in certain geographic areas of the country could indicate specific structural problems, and if this evidence is combined with the need to implement policies aimed at mitigating territorial differences, it is therefore explained how relevant is the need to refine the level of territorial analysis as much as possible, so as to be able to distinguish and initiate specific measures for those territories and schools most affected by this phenomenon.
The objective of this work is to continue measuring school dropout with an easily constructed indicator, capable of going down to a smaller territorial detail. In addition, it is possible to investigate the different dropout rates affecting different types of schools. In addition we will make a descriptive comparison between the indicators produced is the INVALSI data.
Data and Method. To try to quantify the phenomenon of early school leavers at the territorial level, open data from the national student registry (source Italian Minister of Education - MI) were used. The information from the open data registry of the Italian Minister of Education is currently available for the school years 2015/2016, 2016/2017, 2017/2018, 2018/2019, 2019/2020, 2020/2021 and 2021/2022 and provides census information on:

- the number of students enrolled in each Italian school;
- student distribution by age and grade;
- the type of school;
- and the exact geographical location of the school.

In this way, each school is located in a particular territory and provides information on how many students are enrolled there, for each age. In this way, by tracking the school path of students in a given territory (e.g. the province) for the period between the end of compulsory schooling (i.e. between 15 and 18 years), it is possible to calculate how many drop-outs there are (in that territory or for type of school). The early school leavers rate has been calculated at the provincial level, but an important aspect of this indicator is that it could also be calculated at the municipal areas or at the any other territorial level obtained from the sum of

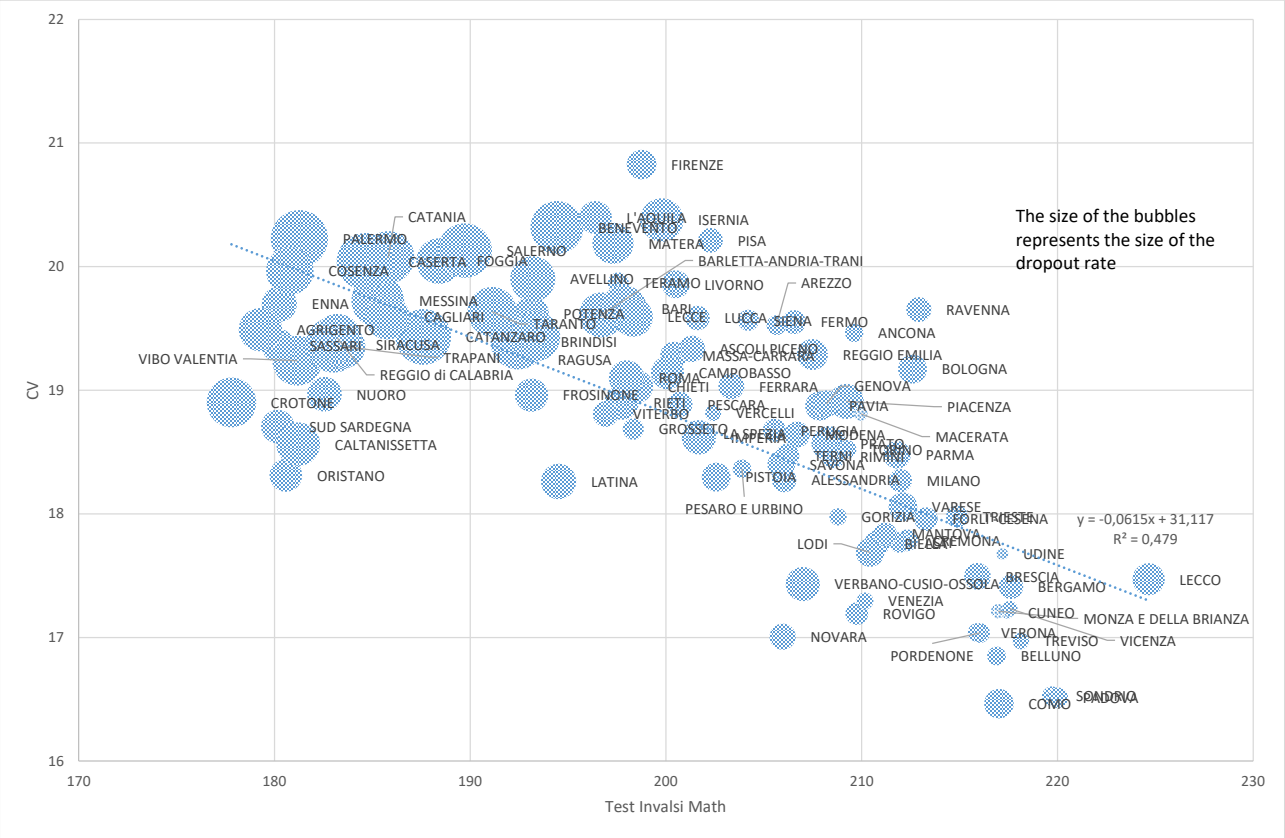
the municipal areas, as for example Local Labor System or Internal areas or Functional Urban Area, or for type of school and so on.
In practice, what we do is follow the schooling of the cohort of 15-year-olds enrolled in all schools in a given territory, in the 2015-2016 school year, until they turn 18 (in the 2018-2019 school year). By tracking this cohort of students, it is possible to see how many of these 15-year-olds continued their studies until the age of 18.
In a more formal way, the dropout rate elaborated (called OUT from Education, OUT-E) is the percentage change of 18-year-olds enrolled in all schools (ISCR18) in a given territory in 2018-2019 compared to the 15-year-olds who were in schools in 2015-2016 (ISCR15):

$$OUT\text{-}E_{,t} = (ISCR18_{p,t} - ISCR15_{p,t-3}) / (ISCR15_{p,t-3})*100$$

where *t* is the school year and *i* is the area over which the rate is calculated (e.g. regions, provinces or other level).
The variation of students may result not from actual school drop-outs, but also from other possible factors (all very limited in their impact). First of all: the cohort of students is not a closed system: new students may enter it from outside or students may leave it as a result of e.g. internal or external migration flows. This factor may alter the value of the dropout rate by overestimating or underestimating it, even if

1) internal or external migration processes are more likely in young individuals, and to have a child at least 15 years old to change school in another country or another region (or another province, municipality) means to be at least over 35: for this reason it seems to us presumable that this factor may have little effect on 15-18 year olds.
2) by the cancellation of a student in a certain school due to traumatic events (death); but, mortality rates are very low in that age group;
3) by a student's decision to change school, choosing one outside the territory taken into consideration.

It is also important to remember that not leaving school at 18 does not necessarily mean completing it. Lastly, some students may choose other educational paths such as IFP (Istruzione e formazione professionale – VET, "vocational education and training").
An overall territorial consistency seems to emerge from the preliminary results even at provincial level. The southern Italian provinces are characterized by higher early school leavers rates: twenty provinces have one students on five drops out of school early; this rate grows up in the Napoli and Caserta province where it is close to thirty percent; but among the worst provinces, some central northern territories (such as Prato, Piacenza, Florence, Livorno, Reggio nell'Emilia, and Imperia) also stand out. This last result also highlights how there is also a good deal of variability within the regions themselves, and how it is important to analyse this phenomen at a more detailed territitorial level.
The geographical areas where there are more job opportunities are those with the lowest school drop-out rates. Thus, the idea that people drop out of school because they want to look for a job does not seem to be predominant in determining a high drop-out rate.
Another interesting result is how in the "post" pandemic the dropout rate (calculated through our indicator) has decreased almost everywhere. This trend seems to be confirmed in the South as well, albeit with some exceptions: Crotone, Siracusa, Matera, Potenza, Trapani.
By treating the OUT-E indicator and data from the INVALSI tests simultaneously, additional considerations can be made.
Viewing INVALSI test scores as a measure of the cognitive and mathematical skills of 15- and 18-year-olds is also a way of recognising that they are one of the key determinants of future individual economic outcomes, and that they depend fundamentally on family background and often on the inequality of opportunities faced by children. The graph shows the inequality of scores on the INVALSI Tests in mathematics (2018-2019, degree 10) as measured by the coefficient of variation, relative to the average student and with the size of the bubbles representing the dropout rate, for Italian provinces.
Interestingly, there is a clear negative relationship between inequality and average scores (less in provinces where the percentage of dropouts is higher. To be analyzed is the substantial difference in score inequality for provinces that are in the same range of mean scores. For example For example, the inequality is higher for Florence than for Viterbo or Potenza.




Further analysis will focus on how inequality, or more precisely the entire distribution of scores, changes over time in a given province using the ratios between deciles precisely to emphasize the distributional characteristics of the results.

***Keywords:*** early school leavers, geographic inequality

## Classroom rank in mathematics and career choices

**Maddalena Davoli – Maurizio Strazzeri – Enzo Brox**

Introduction and Research questions. Social scientists have long been interested in understanding how individuals are affected by their socioeconomic environment. A large literature has applied this idea to education and investigates how peers affect educational choices and production. Several studies focus on the effects of either high-performing peers (Balestra et al., 2021) or disruptive peers (Carrell et al., 2018) and link their results to optimal classroom composition. A recently growing literature investigates the effect of a naturally occurring feature of classroom composition, students' ordinal rank (Murphy and Weinhardt, 2020).

The recent rank literature has found strong effects of classroom rank on many educational outcomes including performance, high school graduation, college enrollment, major choices risk behavior and crime (e.g., Elsner and Isphording, 2018, 2018; Comi et al., 2021). However, there is little evidence on the relationship between classroom rank and labor market outcomes apart from a study by (Denning et al., 2018). We contribute to this literature and investigate the education choice and labor market effects of ordinal classroom rank in secondary school. While previous studies widely focus on college enrollment and major choices, we focus on students in vocational education programs.

In this paper, we study the impact of students' ordinal rank in math in the last year of compulsory school on subsequent occupational choices, earnings, and human capital investments after entering the labor market. Previous literature on rank effects suggest that classroom rank change students' beliefs and behavior—which in turn can influence students occupational choices. Students do not now their ability and use their peers to infer their own ability. As actual or perceived ability in a particular subject gives students incentives

to select into occupation that require skills in this subject, we would expect that classroom rank in math is related to occupational choices and, in particular, increases the probability to select into STEM occupations. We perform our analysis on Swiss PISA data, matched with administrative schools and earnings records.
We follow recent literature and exploit differences in math ability distributions across classrooms to identify the effect of rank in math on various students' outcomes. To identify the causal effect of classroom rank we use an identification strategy that leverages idiosyncratic variation in rank across classrooms. Regression models with classroom fixed effects and flexible controls for absolute math ability show that ranking in the top of the class compared to the bottom increases the likelihood to select into an occupation with STEM skill requirements. Next, we test if these initial occupational choices have consequences on students earnings streams in the years after graduating from the vocational education program. We further show that this effect might be driven by students' self-perception of their math ability as well as effort provision in this subject. Moreover, we exploit detailed information on students' enrollment status in Swiss educational institutions to assess the hypothesis that, if students' occupational choices are based on perceived ability rather than actual math ability, they might not be efficient and students might be more likely to change occupational or educational field after selecting their initial occupation.
Data: For the empirical analysis, we use student-level data from the extended version of the Swiss section of the PISA-2012 wave, additionally linked to :

- student register data (information on students' education status, school track, classroom identifier, VET training occupation) from 2012-2020.
- administrative earning records from tax data from 2012-2020
- information on skill requirements of training occupations (math, natural science, language, foreign language).

We include in our final sample only students for whom we have at least one other student observation in the same classroom andthat could be successfully linked to our administrative data sources. Our final sample consists of 11 684 9th-grader student observations from 1 470 classes of 492 schools.
Methodology: To estimate the effect of students' ordinal rank in the math ability distribution in the classroom on subsequent educational or occupational choices, we follow recent literature on rank effects (e.g., Murphy and Weinhardt, 2020) and compare students who have the same absolute ability but differ with respect to their ordinal rank in the classroom due to different ability distributions of their peers in in the classroom.
In our empirical analysis, we use students' percentile rank in the classroom to measure students' ability rank. To compute the percentile rank Ric of student i in classroom c, we first determine student i's absolute rank in the classroom, nic, by sorting students in accordance with their position in the within-classroom ability distribution. Students' absolute rank nic is a number between 1 and the overall number of students in the classroom (Nic). We assign the absolute rank value of 1 to the student with the lowest ability in the classroom and the highest number (i.e., Nic) to the student with the highest ability in the classroom. Next, we transform the absolute rank in the classroom to the percentile rank using the equation:
$Ric = (nic - 1)/(Nic - 1)$
Independent of class size, Ric assigns value 0 to lowest ability students and value 1 to highest ability students.
We then estimate the equation:
$yic = \beta Ric + f(Aic) + \gamma t Xic + \delta c + \varepsilon ic$, (2) 3
where yic is a measure of educational or occupational choice or yearly income in a given year of student i in classroom c, Ric a students' rank in classroom c, Aic student i's math ability, Xic a vector of student i's background characteristics (sex, age, parental education, nationality, migration status, language spoken at home), and εic represents an error term. We control for students' math ability, Aic, with a second-order polynomial. Additionally, we add a set of class fixed-effects, δc, that capture any mean differences in the outcome variable between classroom (e.g., due to classroom-specific teacher effects). To allow for within-school correlation of the error term, we cluster standard errors at the school level.
Results: The analysis is still ongoing, but first results are already available and make it clear that the school environment—and students' rank in the classroom—are linked to students' occupational choices and subsequent income flows. Further, our estimates suggest that students' changes in behavior and beliefs are a likely channel for the observed effect and that rank is not associated withoccupational mismatch.

1. We observe a positive and slightly increasing impact of our rank measure on yearly income starting in the year 2015. In 2020—the year for which we obtain the highest estimate—ranking in the top of the classroom compared to the bottom, conditional on math ability, is associated with an increase in yearly

income by 3 221 CHF (9.4 % relative to the sample mean). Students entering a vocational education program are the main factor explaining the positive association between classroom rank and yearly income.
2. Students who start a vocational education program and have a higher rank in math conditional on absolute ability might select into training occupations that provide higher-paid job opportunities after graduation, such as occupations which place higher importance into skills related to math/science (STEM) dimension. We indeed find evidence of this fact, as classroom rank in math is positively associated with the STEM intensity of the training occupation,that potentially provide higher-paid job opportunities.
3. Besides its effect through changes in teacher and parental investments, changes in students' beliefs and behavior are the main mechanism that explains students' outcomes due to classroom rank. We exploit students' responses to 8 PISA questions regarding their attitude towards math and find that students with higher classroom rank in math are more interested in math, have a higher self-perception of their math ability, and provide higher effort studying math. Hence, higher effort provision might explain the link between math rank and subsequent labor market choices.
4. Finally, we assess whether classroom rank in math also affects students' likelihood to change the field of study after the initial choice of their occupation. This would suggest that student with a higher rank in math might over-assess their ability and select into occupations for which they are not well suited. We do not find evidence of an association between classroom rank in math and increased investments in education fields different from the initial one. Instead, we find an increased investments in education programs in the same education field, which we interpret as evidence that classroom rank does not lead to occupational mismatch.

***Keywords:*** ordinal rank, peer effects, occupational choices

## The heuristic value of data for school evaluation as a data-driven decision-making process

**Michela Freddano – Miriam Mariani**

Introduction. The paper aims to reflect on the potential of using data from different sources for the implementation of data-driven evaluation and decision-making processes at school, through a summary of the theoretical framework in which the topic fits: data literacy, participants in the evaluation process (Ackoff, 1989; Mandinach, Honey & Light, 2006), characteristics of data-driven decision-making processes (Mandinach et al., 2006; Kaufman et al., 2014), system approach to evaluation (Senge, 1990; Bocchi & Ceruti, 2007).

In school evaluation, one of the greatest challenges facing different governance systems nowadays is the strategic use of data produced by the different instruments of the Italian National Evaluation System of Instruction and Training: the Italian national testing on student achievement, supplementary surveys, the Self- Evaluation Report (SER) and the related School Questionnaire, the Three-Year School Educational Plan (SEP), the School Social Report (SSE), etc.

Research object and hypothesis

The literature shows that data-driven decision-making processes increase productivity and operational effectiveness, especially in the business field (Brynjolfsson & McElheran, 2017). In education, models based on data analysis and interpretation are useful in prioritizing the most urgent needs and the identifying the appropriate solutions based on available resources (Reeves, 2010 in Kaufman et al. 2017).

The purpose of this paper is to explore the value of data made available as part of school self-evaluation and school external evaluation, with particular regard to the potential of using INVALSI testing data in relation to data from other sources.

In school evaluation, the synergy between self-evaluation and external evaluation shows the need of the correct reading and interpretation of data, not only of different type (e.g., quantitative or qualitative), different source (e.g. data provided and data collected) and level (e.g., individual, class, school, district, region, macro-area, nation).

Data used

The paper shows a simulation of reading and interpreting process of INVALSI data contained within the SER, focusing, therefore, only on the analytical phases of the evaluation process, in order to provide useful indications for those who carry out the activities for self-evaluation or external evaluation and to outline strategies for identifying at-risk situations on which to design an intervention. In particular, the Map of Indicators of the format of SER 2022-2025 (INVALSI, 2022) had been used as a guiding tool for setting the analytical process. Data from two dimensions had been taken into account:

1. Context: areas 1.1 "School Population" and 1.2 "Territory and Social Capital," where the variety of data sources (Ministry of Instruction and Merit, INVALSI, ISTAT, schools) provides a "biographical registry" picture of the school reality and of its context (student population composition, socio- economic and cultural status of families, ESCS index);
2. Outcomes: areas 2.1 "School Outcomes," 2.2 "Results in National Standardized Tests," and 2.4 "Distance Outcomes" where data on school outcomes, INVALSI testing on student achievement, including distance learning, and continuity of studies and job placement.

Method. The internal evaluation team and the external evaluation team allow triangulation between different profiles, useful for evaluating the educational, pedagogical, psychological, social, managerial and organizational aspects from a systemic point of view. Theoretical framework by Mandinach, Honey & Light (2006) named "Theoretical Framework for Data Driven Decision Making", and later updated (Mandinach & Gummer, 2016), which, starting from Ackoff's model (1989), involve an iterative process of steps (from data to information, from information to knowledge, from knowledge to decision) and a series of necessary cognitive skills such as collecting data, organizing them, analyzing and selecting information, synthesizing and proposing priorities. In the various stages of the evaluation process, data-driven dimension is constant, so data reading from the SER and the other strategic documents (SEP, SSR and improvement plan), their contextualization, the integration of quantitative datasets with qualitative data (proposed by schools in self-evaluation or collected through interviews in the external evaluation), the formulation of judgments and priorities for action, and the return of the outcomes of the evaluation visit to the visited schools contribute to the evaluation process. The process of triangulation of viewpoints, first for knowledge and then for decision-making, takes place through feedback and learning mechanisms, as the simulation (Sterman, 2000) that takes place in observers (Di Battista et al., 2006; Minati, 2008) constitutes the tool for the analysis of evidence but also of serendipity as unexpected behavior (interpretive approach) and for the orientation of strategic choices in the organization (design approach) (De Toni, 2021). Starting from the variety and interdisciplinary nature of the theoretical and the multiple perspectives, this contribution intends to propose a reflection on the selected data by highlighting the salient aspects of the process of reading data and interpreting them, both procedurally and instrumentally.

Results. The identification of the data considered most relevant, interpreted and contextualized allows to draw a summary picture and to bring out correlations between the description of the context and specific problems, thanks to the contribution made by the experience and professionality of the people involved.

The integration between the various data allows a systemic reading (Senge, 1990; Morin, 1985 in Bocchi & Ceruti, 2007) of the complexity of reality and to exploit the heuristic value of transforming individual disordered data into an intelligible and instrumentally strategic knowledge framework for decision-making purposes. However, this translational process is not immediate, making it difficult to carry out evaluation activities with regard to the approach to data, both in reading the data provided to the school by INVALSI on test results and in the interpretation of SER data by external evaluators.

In this regard, there is the need of promoting specific training courses structured at different learning levels to improve data literacy and data-driven decision-making skills that support evaluators in reading, systemizing and interpreting data, recognizing qualitative, quantitative, provided and collected data and integrating them into a single dataset, formulating priorities and judgments by looking at data from a systemic perspective, and using process models and tools as well as digital platforms and data visualization systems. Educational institutions are complex realities, for the evaluation (internal and external) of which, therefore, requires an approach that takes into account the non-linearity of internal dynamics and promotes a reflective and participatory evaluation process based on the evidence found in data (data-driven processes) and includes ad hoc training paths.

The aim to consider, in fact, training as an investment and strength for the effectiveness of the evaluation process, stems from the need to focus on key competencies to strengthen the self-reliance of developing action plans aimed at improvement from the perspective of the school-system as a learning organization.

***Keywords:*** data literacy, data driven decision making, reflective process, learning organization, INVALSI

# Tackling school dropout through the school climate: a proposal of a process indicator from the Experimentation of Self-Evaluation Report for Vocational Education and Training (SER VET)

**Daniela Torti – Michela Freddano**

The term “school dropout” is broad and does not only include early school leaving, but repetitions, irregular attendance, educational failure, implicit school dropout (Ricci, 2019). Early school dropout has been increasingly considered an indicator of the quality of school systems, in terms of equity, inclusion, opportunities and educational success (Pandolfi, 2016). The main literature (Alistair & Leathwood, 2013; Batini, 2014; Batini & Bartolucci, 2016; Colombo, 2015; Lamb, Markussen, Teese, Sandberg & Polesel, 2011; Scales, 2015; Pandolfi, 2016) uses systemic and multidimensional models and interpretations of the phenomenon.

In this perspective, recent studies highlight the incidence of various factors divided into three macro-categories (Eurydice & Cedefop, 2014):

- factors regarding students' psychological characteristics (learning difficulties and disabilities, special educational needs, refusal to go to school, disengagement and lack of motivation, self-perception of inadequacy, etc.);
- family and socio-economic factors (belonging to a socially and economically disadvantaged category, parents' low level of education and parents' involvement in their children's school life, family instability, etc.);
- factors internal to the school system (school organization and climate, methods of carrying out teaching, quality of interpersonal relationships, repetitions, etc.).

Specifically, the term “school climate” refers to the "shared beliefs, values, and attitudes that shape the interactions between students, teachers, and principals" (Mitchell et al., 2010: 272). Due to its complexity, the main reference literature show school climate construct as a multidimensional concept composed by four main dimensions: a) teaching and training environment, b) relational; c) security and d) institutional levels (Cohen, 2013; Wang and Degol, 2016). The education and training dimension focuses on the overall quality of the atmosphere of the educational and training institution, including curricula, instruction, teacher education and professional development. The report highlights the quality of interpersonal relationships within the school. Safety represents the degree of physical and emotional safety provided by the school, as well as the presence of effective, consistent and fair practices to ensure discipline. Finally, the institutional environment reflects the organizational or structural characteristics of school environment. Collectively, these four dimensions encompass nearly all the essential characteristics of the school environment that greatly affect the cognitive, behavioural and psychological development of students, teachers and school staff (Wang & Degol, 2016). The school climate, therefore, reflects the nature and quality of school life, as well as the perception that individuals have of their school starting from the experience gained within the school itself (Biasi, Castellana and De Vincenzo, 2023) .

Despite the complexity and multidimensionality of the construct, research studies highlight the effect that a positive climate can have in determining effective education (Scheerens, 2018). As a fact, many studies affirm that a positive climate effects academic performance, commitment and motivation to learn students, as well as on reduces problematic episodes and dropout rates ( Radshaw et al., 2014; Cohen et al., 2009; Cornell, Shukla & Konold, 2016; Gottfredson et al., 2005; Shukla, Konold & Cornell, 2016; Thapa et al., 2013; Way, Reddy & Rhodes, 2007). Therefore, evaluating school climate is essential for activating improvement processes aimed at supporting students in developing the necessary skills for both personal and professional fulfilment necessary for effective participation in democratic society and countering complex phenomena such as school dropout.

Starting from these premises, the article shows the indicator of the school climate for self-evaluation used in the experimentation of the Self-Evaluation Report (SER, Italian acronym RAV) for Vocational Education and Training (VET) with the aim to investigate the possibility of relate this process indicator to the indicators concerning student dropout and learning assessment in the INVALSI tests.

This research study is part of the wider Experimentation of the SER VET carried out in the training years 2020/2021 and 2021/2022 (INVALSI, 2022). The introduction of VET (in Italy, of regional competence) into the evaluation procedure managed by the Regulation on the national evaluation system in the field of education and training (Presidential Decree 28 March 2013, n. 80), involved the first phase of the procedure concerning the self-evaluation and compilation of SER by each VET schools. The purpose of the

Experimentation was to allow the VET schools in relation to the segment dedicated to the fulfilment of the right and duty to education and training, to use a method for self-evaluation similar to that of schools that refer to the National Evaluation System. The research study carried out in the context of the Experimentation has made possible to harmonize the tools for the self-evaluation of VET, on the one hand matching them to those used in the system for schools, on the other guaranteeing the specificity of this particular segment of instruction. Through the self-evaluation, the VET schools had the opportunity to analyze and deepen the training process that had the greatest impact on students' outcomes, in order to identify strengths, weaknesses and potential levers for the improvement training and professional community of each location. In line with what is present in the SER of the other orders, also for VET the indicators from SER are, together with the guiding questions, stimulus to reflect on the results achieved in each specific area. Most of the SER VET indicators had been calculated from the data collected through two tools: the VET School Questionnaire and the Teachers Questionnaire. The Teachers Questionnaire had the purpose of defining two indicators of the dimension of the Processes, Educational and Didactic Practices present in the Learning Environment area of the SER VET: Methodologies that promote students' learning and Relational Climate. This article focuses on this last indicator, inherent to the perception of the school climate by teachers. As part of the SER VET Experimentation, for the study of the relational climate, following an accurate review of the literature, the School Level Environment Questionnaire (SLEQ) had been used (Johnson, Stevens and Zvoch, 2007) as a widespread in the assessment of teachers' perception of the school climate in international school and training contexts (Johnson, Stevens and Zvoch, 2007). Compared to the original version, 17 items referring to the four subscales were selected: collaboration between trainers, relationship with students, decision-making process and innovation. Items had been translated from English to Italian adapting the format to the Italian educational and VET context. Therefore, a process of revision of the items was carried out, also in discussions with VET experts. The Teacher Questionnaire, implemented on the LimeSurvey platform in the period May-July 2021, had been intended for all VET teachers in the VET schools involved into the experimental project, and provided for the completion in total anonymity. Overall, 3,413 trainers from 172 of the 173 VET schools participated at the survey.
Once processed, the data relating to the Relational Climate indicator collected with the Teacher Questionnaire were returned to the VET school on the SEF, in the "Learning Environment" process area, both at school level and as territorial reference on the Self-Assessment Report.
The validation of the school climate scale and the study of the relationship between school climate, school dropout and, where present, student learning outcomes in INVALSI testing are currently underway. On the occasion of the Seminar, the main results from this study will be presented with the aim of probing the hypothesis of applicability of this tool for self-evaluation also to other segments of instruction. In our opinion, having a standardized school climate evaluation tool can put educational and instruction institutions in the position to intercept the quality of the school climate and to plan interventions to improve school environment also aimed at preventing complex phenomena such as early dropout.

***Keywords:*** school self-evaluation, school climate, early school leaving, teacher questionnaire, indicators

# THEME 3. INTERNATIONAL LARGE-SCALE ASSESSMENTS (ILSAS)

## METHODS AND RESULTS

**ORGANIZER: INVALSI**

**COORDINATOR: MAGDALENA ISAC**

**NOVEMBER 23RD: 2.00 P.M. – 4.00 P.M. {ROOM 3 – RESEARCH SESSION 3}**

---

## The relationship between school working conditions and teacher job satisfaction in Italy: Evidence from TIMSS 2019

**Nurullah Eryilmaz**

The Buona Scuola reform, also known as The Good School reform (La Buona Scuola, 2015, Law 107/15), was launched in 2015 in an effort to address the current issues in education and create an efficient educational system. The legislation that supports the reform calls for a number of initiatives in the areas of autonomy, hiring new teachers, adding a merit-based component to teachers' pay, requiring teacher development methods, and fostering digital innovation and skills in schools. These steps have the potential to improve students' achievement in school.

Moreover, the "La Buona Scuola" (The Good School) reform in Italy, implemented in 2015, aimed to bring significant changes to the education system, including improvements in teacher working conditions. Ultimately, this reform might lead to higher teacher job satisfaction by improving teacher working conditions.

This study allows us to see how the reform has been implemented effectively in the context of teacher working conditions and teacher job satisfaction. In doing so, in this study, we aimed to determine the relationship between school working conditions and teacher job satisfaction by analyzing the Trends in International Mathematics and Science Study (TIMSS) 2019 Italy data. The findings of this study add to the body of knowledge on teacher job satisfaction and have practical implications for policymakers, school leaders, and other stakeholders in education who are working to improve teaching practices and foster supportive and conducive working environments for teachers in Italy.

In the following section, we first present some background information on the factors related to school working conditions that were taken into consideration for this study, teacher job satisfaction, as well as relevant previous studies in this field of study. Later on, we discuss the Italy TIMSS data, our samples, variables, analytical strategy and present our findings. Lastly, we mention the implications, limitations, and future directions of the research.

Literature Review. Teacher Job Satisfaction. Due to its importance in predicting employee performance, job satisfaction has been a research issue over the seven decades (Judge et al., 2001). Despite much discussion about job satisfaction, there is no agreed-upon definition of job satisfaction. Commonly, the level of satisfaction a person gets from their working environment is referred to as job satisfaction (Pepe et al., 2017).

School working conditions. The term "school working conditions" refers to the variety of factors in an educational setting that may have an impact on teachers' experiences and wellbeing (Burkhauser, 2017; Ladd, 2011). The working conditions in schools have a significant impact on the motivation, effectiveness, and job satisfaction of teachers (Bascia & Rottmann, 2011).

School Leadership. Effective leadership is defined by cooperative and supportive behaviors, clear communication, and the accessibility of resources and opportunities for advancement in your career. Teachers are more likely to perceive greater levels of job satisfaction if they believe their school leaders to be competent, approachable, and supportive.

Student discipline. In the context of the classroom and school, student behavior, conduct, and adherence to standards are referred to as student discipline. It includes components like disruptive behavior, disobedience, negativity, and other improper conduct and misbehaviors from students might lead to teacher stress and burnout.

Teacher workload. The term "teacher workload" describes the quantity of tasks, responsibilities, and scheduled requirements placed on teachers in their professions. It includes many components like preparing lessons, classroom management, evaluations, administrative duties, extracurricular activities, and professional development.

Current Study Based on the definitions and empirical evidence mentioned above, this current study lays the foundation for researchers and policymakers to further understand how school working conditions and teacher characteristics influence teachers' job satisfaction. Specifically, this study aims to investigate the impact of leadership support, student discipline, teacher workload, as well as individual factors such as gender, teaching years, having a mathematics or mathematics education background, and participation in professional development programs, on teacher job satisfaction in Italy. By examining these factors, the study seeks to provide valuable insights into the factors that contribute to teacher job satisfaction in the Italian context.
Methodology. Data. This research uses secondary data analysis of data from the Trends in International Mathematics and Science Study (TIMSS) 2019 conducted by the International Association for the Evaluation of Educational Achievement (IEA). In this study, we used data from the eighth-grade mathematics teacher questionnaires across. In Italy, 209 teachers participated in TIMSS 2019.
Variables. Dependent variable: The outcome variable for this study is teacher job satisfaction which consists of five items in the data set (BTBG08A, BTBG08B, BTBG08C, BTBG08D, BTBG08E). The items denoting teacher job satisfaction were recoded so that higher values represent higher levels of teacher job satisfaction. Cronbach's alpha for teacher job satisfaction scale was 0.907.
Independent variables: The independent variable of student discipline consists of five items in the data set (BTBG07D, BTBG07E, BTBG07F, BTBG07G, BTBG07H). The items denoting student discipline were recoded so that higher values represent better student behavior. Cronbach's alpha for student discipline scale was 0.882.
The independent variable of teacher workload consists of five items in the data set (BTBG09B, BTBG09C, BTBG09D, BTBG09E, BTBG09H). The items denoting teacher workload were recoded so that higher values represent higher levels of teacher workload. Cronbach's alpha for teacher workload scale was 0.729.
The independent variable of school leadership support was captured only using single item (BTBG06L). This variable has the highest mean score (M=3.84).
Moreover, teacher demographics information and professional characteristics, such as gender, teaching experience, mathematics and mathematics education as a major, number of hours of professional development.
Lastly, in this study, to answer the research question, we estimated one regression model. In this model, we look at the relationship between school-working conditions and teacher characteristics and teacher job satisfaction.
All estimations were performed using Mplus 8.4 (Muthen & Muthen, 2019) and executed in the R statistical software (R Core Team, 2019).
Result. Leadership support was found to be positively related to indicators of teacher job satisfaction. This relationship is statistically significant. As a result, we can conclude that teachers who have more leadership support from their principals tend to have a higher level of job satisfaction than teachers who have less leadership support from their principals.
A similar pattern to the one found for leadership support can be seen in the relationship between student discipline and teacher job satisfaction. The student discipline scale was a positive, statistically significant predictor ($p<0.001$) of teacher job satisfaction.
The other variables did not show any significant effect on teacher job satisfaction.
Conclusion. The results contribute to the existing body of knowledge on teacher job satisfaction and provide practical implications for policymakers, school administrators, and educational stakeholders in creating supportive and conducive working environments for teachers. By unraveling the intricate link between school working conditions and teacher job satisfaction, this research aims to improve educational practices and contribute to the professional well-being of teachers in Italy.

***Keywords:*** Italy, teacher job satisfaction, school working conditions, Buona Scuola reform

## School Belonging and (Forced) Migration

**Daniel Shephard**

Globally there more than 280 million migrants and over 100 million forcibly displaced individuals (UNDESA 2021; UNHCR 2023). Young people of school-going age are disproportionately represented in the latter population (UNHCR 2023). Unfortunately, very little is known about how school systems have created an enabling environment for both migrants and local students to feel that they belong in their (new) schools (Dryden-Peterson et al. 2022). There is evidence that favors local students as well as evidence that favors students who are migrants by choice or necessity (Andersson et al. 2010; Smith et al. 2016; Wicht 2016).
This study will use data from TIMSS to interrogate differences in students' sense of school belonging between (forced) migrants and local students and between schools with high and low concentrations of migrants. These differentials will then be compared between countries organized into three regions: the Middle East, Continental Europe (including Italy), and high income Anglophone countries. The goal of this study will be to elucidate how data from International Large Scale Assessments (ILSAs) can inform debates around how well school systems are supporting the inclusion and integration of foreign-born students--especially those faced with forced displacement. The analysis will be done using appropriately weighted descriptive statistics and regressions within the R package RALSA v1.3.0.
While there has been recent research using ILSAs to elucidate the relationship between forced migration concentration and teachers' characteristics (Cook & Kim, 2023) and numerous individual case studies of school belonging (Russell & Mantilla-Blanco 2022), there has been limited work using ILSAs to understand the relationship between students' sense of school belonging as it relates to (forced) migration status and concentration. This is an important gap to fill, because students' sense of school belonging provides an important test of whether inclusion policies merit their name. In this sense, students' sense of school belonging is intrinsically important (Dryden-Peterson et al. 2019). However, extensive literature also demonstrates how school belonging is important instrumentally, because it is linked to positive academic, psychological, and social outcomes for children (Allen et al. 2018, Tan et al. 2022).

***Keywords:*** migration, belonging, composition, TIMSS

## Investigating the impact of inquiry-based teaching on science learning: a further analysis in Italy

**Zhijun Chen – Andres Sandoval Hernandez**

Inquiry-based teaching is a promising approach to developing students' reasoning and thinking skills (Teig et al., 2019). This approach encourages students to ask questions, gather evidence, and draw conclusions. Inquiry-based teaching is based on the scientific method, and it has been shown to be effective in helping students learn. The relationship between teaching methods and student achievement is one of the most important indicators of teaching effectiveness (Teig et al., 2019). In recent years, there has been a growing trend toward inquiry-based teaching.
An increasing number of studies have explored the relationship between inquiry teaching and student science achievement, but they have failed to draw a firm conclusion about the effectiveness of inquiry-based teaching. Some studies have found that inquiry-based instructional practices are associated with positive outcomes (e.g., Schroeder et al., 2016), while more recent studies using ILSA data have shown that these practices are actually negatively associated with science achievement (e.g., Cairns & Areepattamannil, 2017). One explanation for these inconsistent findings could be due to how inquiry-based teaching is measured and how the data is analyzed based on its measurement.
More specifically, many studies that explore the effectiveness of inquiry-based teaching primarily assess the frequency of inquiry activities, emphasizing how often they are implemented (e.g., Urdanivia Alarcon, 2023). Moreover, these studies commonly assume a linear correlation between the occurrence of inquiry activities and students' academic achievement. In contrast, Creemers and Kyriakides (2008) suggest studies on teaching effectiveness should explore the potential of a nonlinear relationship between the occurrence of teaching behaviors or activities and students' academic achievement.

For example, on the one hand, teachers should implement inquiry-based teaching to a significant extent in science classes to develop students' scientific literacy (Teig et al., 2018). On the other hand, the effective implementation of this approach demands a substantial amount of class time, as highlighted by Guskey (2000). If teachers spend too much time on inquiry activities, they may not have enough time for other important teaching and learning practices. Hence, there might be an optimal amount of inquiry-based teaching that leads to better student achievement, which indicates a curvilinear relationship.

In 2018, Teig and her colleagues find the curvilinear relationship between inquiry-based teaching and science achievement using TIMSS 2015 Norway data. However, Teig and her colleagues' study (2018) is limited to one educational system, leaving the question of whether the curvilinear relationship between the variables in investigation would exist in other educational systems.

Object and research hypothesis. To fill in this gap, in this study, we aim to investigate the relationship between inquiry-based teaching and student science achievement in science using TIMSS 2019, following Teig et al. (2018). We also aim to determine whether the relationship is linear or curvilinear across different education systems. Additionally, we choose countries with significant results for further analysis to compare their educational systems with other countries.

Data. This study utilized data from the Trends in International Mathematics and Science Study (TIMSS) in 2019. TIMSS is a widespread assessment that occurs every four years and aims to gather information about the academic performance of students in mathematics and science at the fourth and eighth-grade levels. The study encompassed numerous countries and involved the collection of contextual data through questionnaires administered to students, teachers, and school leaders. For this particular study, the analysis focused on a subset of 37 countries that took part in the eighth-grade assessment in 2019.

Measures. Science achievement. In this study, we used eighth-grade science achievement as the outcome measure. The assessment was in four content domains: earth science, chemistry, physics, and biology. For each student, five plausible values of their science achievement were reported. We used all five plausible values in our models.

Inquiry-based teaching

TIMSS 2019 collected data on how often teachers use different teaching practices in the classroom. Following Teig et al.'s (2018) study, we selected five items out of 15 teaching and learning activities to build a scale of inquiry-based teaching, such as “design or plan experiments or investigations.”

Socioeconomic status (SES)

Three items were selected from the student background questionnaire to create a measure of socioeconomic status (SES) at the student level. These items were students’ ratings of the number of books at home and their mothers’ and fathers’ highest education levels. We also created a measure of SES at the classroom level using the same three items under the doubly-latent variable framework. Both individual and classroom SES were added into the model as controlling variables.

Method. As we were investigating the effectiveness of inquiry teaching in the classroom, we conducted multilevel structural equation modeling (MSEM) with both student and classroom levels. The data analysis was conducted in two steps using Mplus with the XWITH command (Muthén & Muthén, 2017).

First, we conducted multilevel confirmatory factor analysis (MCFA) to establish and test measurement models of inquiry-based teaching across 37 different educational systems. This allowed us to create a model that accurately represented the relationships between different aspects of inquiry-based teaching.

Second, we performed MSEM on the basis of the previously examined factor structures in four models. These models examined the relationships between inquiry-based teaching, SES, and achievement and whether the relationship between inquiry-based teaching and achievement was linear or curvilinear.

Both models were tested for the pooled datasets and the datasets collected by each educational system. This allowed us to compare the results across different educational systems and to get a more comprehensive understanding of the relationship between inquiry-based teaching and achievement.

Results achieved. Initial findings from the pooled dataset of TIMSS 2019 suggest a curvilinear relationship between inquiry-based teaching and science achievement. These findings indicate that achievement increases with a higher frequency of inquiry-based teaching up to the optimal frequency. However, an excessive amount of inquiry activities may have an adverse effect on student performance. Similar findings are also confirmed by the Italian dataset. Further analysis will be performed based on the comparison between Italy and other European education systems.

The outcomes of our study will help generate valuable information that will enable us to propose theoretical mechanisms for understanding the connection between inquiry-based teaching and educational achievement. By taking a comprehensive and internationally comparable approach, we gain a deeper

understanding of the factors that influence the development and trajectories of educational systems worldwide.

***Keywords:*** inquiry-based teaching, science learning, ILSA

## What makes a good reader? Evidence from PIRLS 2021

**Paola Giangiacomo – Valeria F. Tortora – Michele Cardone**

Introduction. Reading is a fundamental skill for the cognitive, social and emotional development of boys and girls. It allows access to knowledge, to understand the world, to express one's identity and to participate in democratic life. However, reading is not a natural skill, but the result of a complex and articulated learning process, which requires the intervention of various individual, family, scholastic and social factors. To monitor the level of preparation of students in reading and identify good educational and teaching practices in this area, the International Association for the Evaluation of Scholastic Achievement (IEA) has carried out the PIRLS survey (Progress in International Reading Literacy Study ), which has been repeated every five years since 2001. The objective of the survey is to provide internationally comparable data on the performance and contextual factors influencing the learning and teaching of reading of fourth grade students. The survey is based on two main purposes of reading: literary experience and acquisition and use of information. Furthermore, the survey evaluates four comprehension processes within each of the two purposes: i) focusing and retrieving explicitly stated information, ii) making simple inferences, iii) interpreting and integrating ideas and information, iv) evaluating and criticizing contents and text elements. The survey uses tested tools and methodologies validated thanks to a vast cooperation of international research institutions, under the supervision of the PIRLS and TIMSS International Center of Boston College. The survey involves over 60 education systems (countries and benchmarking entities ) around the world.
This article presents the main results of the IEA PIRLS 2021 survey for Italy and compares them with those of the other participating countries, highlighting the strengths and areas for improvement of the Italian education system with particular reference to students' reading skills at fourth year of schooling. Furthermore, the PIRLS survey, in addition to measuring the pupils' performance levels, allows us to collect and analyse, through context questionnaires, a great deal of information on the family environment and on numerous factors concerning the student, closely related, in the literature of reference, with learning by reading. In fact, children learn to read in different contexts and through a great variety of activities and experiences and each environment (home, school and national context) can support the other by creating a connection between home and school which is fundamental for learning.
Object and research hypothesis. The aim is to provide internationally comparable data on the reading literacy levels of boys and girls after four years of primary school and to survey the variables associated with reading achievement. This article will present the main results of the PIRLS 2021 survey, with particular attention to the positioning of Italy compared to other European countries. We will analyze the differences between geographical areas and the variables that most influence students' reading ability. Finally, the educational and didactic implications of the survey results will be discussed, proposing possible actions to improve and develop students' linguistic skills. We will compare these data with the previous 2016 survey to analyze the impact that the pandemic may have had on the results achieved.
This article describes and analyzes information relating to the family context and the characteristics and attitudes of students towards reading; this was done both by comparing the results relating to the indices and the average return associated with them in a comparative perspective with respect to the other countries participating in the survey, and by investigating in more detail the situation of our country, at the level of the various geographical areas, analyzing the indices and variables contained within the context questionnaires. The hypothesis of the research is that the results of the survey offer food for thought and action for the improvement and development of the linguistic skills of Italian students.
Method. The IEA PIRLS 2021 survey took place between 2020 and 2021, with data collection facing the challenges posed by the COVID-19 pandemic. To ensure the safety of the students, teachers and staff involved, prevention and protection measures have been adopted that are appropriate to national regulations and local conditions. In some countries, including Italy, the possibility has been offered to administer the survey in digital mode ( ePIRLS ), with interactive materials and online project simulations.

The Italian sample is made up of 4,500 grade 4 students from 150 primary schools distributed throughout the country. The participation rate of the schools was 93%, while that of the students was 95%. 51% of students took the survey digitally, while 49% took the survey on paper.
In addition to the reading test, students completed a questionnaire about their reading habits and preferences, their school and family experiences, and their opinions about reading. The parents of the students also filled out a questionnaire on the socio-economic characteristics of the family, on reading activities with their children and on the resources available at home for reading. The teachers of the sample classes filled in a questionnaire on their professional characteristics, teaching practices, class climate and reading curriculum. Finally, the school principals filled out a questionnaire on the characteristics of the school, the resources available, the school climate and the management of the health emergency.
Results. The average score of Italian students in the IEA PIRLS 2021 survey is 525, down 9 points compared to 2016. This result places Italy in 23rd place among the 57 participating countries, with a statistically significant difference compared to the international average (500). Among the countries of the European Union, Italy ranks 14th out of 24, with a non-significant difference compared to the European average (524). The countries with the highest scores in the survey are Singapore (602), Hong Kong SAR (598), Ireland (582), Finland (578) and Poland (577). The countries with the lowest scores are Morocco (323), South Africa (323), Egypt (330), Oman (340) and Qatar (341).
The analysis of the results by reading purposes and processes shows that Italian students perform better in reading "for literary experience" (531) than in reading "to acquire and use information" (519). Furthermore, Italian students perform better in “focusing and retrieving explicitly stated information” (538) and in “making simple inferences” (533) than in “interpreting and integrating ideas and information” (518) and “evaluating and critiquing contents and textual elements” (511).
The analysis of the results by gender shows that Italian female students have an advantage of 20 points over their male peers (536 vs. 516), in line with the international average. This gender gap is present in all reading purposes and processes, but is wider in reading for literary experience (24 points) and in evaluating and criticizing textual content and elements (23 points).
The analysis of the results by socio-economic context shows that Italian students from families with a higher level of education than their parents have an advantage of 67 points compared to students from families with a lower level of education than their parents (559 vs. 492). This difference is higher than the international average (48 points) and the European average (45 points). Furthermore, Italian students from families with a higher level of employment than their parents have an advantage of 55 points compared to students from families with a lower level of employment than their parents (548 vs. 493). This difference is also higher than the international average (41 points) and the European average (38 points).
The analysis of the results by school context shows that Italian students who attend schools with a positive climate have an advantage of 42 points compared to students who attend schools with a negative climate (539 vs. 497). This difference is lower than the international average (51 points) and the European average (53 points). Furthermore, Italian students who attend schools with good health emergency management have an advantage of 28 points over students who attend schools with poor health emergency management (532 vs. 504). This difference is similar to the international average (27 points) and the European average (28 points).
The analysis of the results by family context shows that Italian students living in families with a positive climate have an advantage of 44 points compared to students living in families with a negative climate (537 vs. 493). This difference is lower than the international average (54 points) and the European average (56 points). Furthermore, Italian students living in families with a good availability of reading resources have an advantage of 63 points compared to students living in families with a low availability of reading resources (551 vs. 488). This difference is higher than the international average (46 points) and the European average (47 points).

***Keywords:*** PIRLS 2021, print and digital reading, fourth grade students, educational and didactic implications, large-scale international assessments, methods, comparative research

# THEME 8. LEARNING TO LEARN IN SCHOOLS, IN ADULT EDUCATION AND IN VOCATIONAL EDUCATION AND TRAINING

**ORGANIZER: INVALSI**

**COORDINATOR: CRISTINA STRINGHER**

**NOVEMBER 23RD: 2.00 P.M. – 4.00 P.M. {ROOM 4 – RESEARCH SESSION 4}**

---

## Learning to Learn: teachers' conceptions 'in medias res'

**Lucia Scipioni – Roberta Cardarello**

Introduction. Learning to learn (L2L) is a universally recognised as an important 'hyper-competence' (Demetriou, 2015). It is also recognized as a value and a necessity and has thus led to several contributions in the political, cultural and educational field (European Council, 2006; 2018). Despite the plurality of possible interpretations, there is unanimous agreement on the centrality of teachers' beliefs and conceptions regarding this competence (Hounsell, 1979; Waeytens et al., 2002; Stringher, 2021; Gentile, 2022). Therefore, teachers' authentic conceptions must still be investigated, especially since research on competence issues in general has shown frequent gaps between declaration of principles or goals/objectives and effective teaching practices or actions.

Research object and hypothesis. The survey here presented intends to sound out some teachers' representations of L2L, and most importantly, it tests particularly genuine representation, as contributions from the teachers involved were not provided beforehand, but rather they emerged from comments after the projects and activities had come to an end. The hypothesis is that, in this way, conceptions also differing from those to be found in the literature, or Italian articulations of those already consolidated at an international level - narrow vs. broad conceptions (Hounsell, 1979; Waeytens et al., 2002; Stringher, 2021) or central activities with potential to promote competence vs. peripheral activities (Brito Rivera et. al., 2021) - may emerge, in addition to a variability that depends on the degree of schooling, which must necessarily be considered. It has been said that what teachers think is central to the promotion of competence, because their beliefs and theories feed into the day-to-day action of doing school, much more intensively than any ad hoc project, or any activity. In the literature, teachers' conceptions are often investigated through questionnaires or self-perception scales that also examine the L2L construct (Marcuccio, 2009; Lucisano, Botta, 2023). As shown by Stringher and colleagues' research in 2021, when approaching and investigating teaching and educational practices through the words of teachers, one obtains accounts which have the advantage of enriching concepts and are certainly more relevant for measuring the distance and consistency of those conceptualisations. Along similar lines, the present study focuses on a sample of teachers from the Italian province of Reggio Emilia and has promising characteristics, as teachers' conceptions are not investigated in abstract terms; instead, they emerge from the teaching practices themselves, that is, the research examines these opinions in close connection with practical implementations. Thus, the main objective is both methodological and substantive: to test a method of investigating conceptions, but also to verify whether more articulate, naive or precise conceptions emerge in this way.

Data. The starting data are represented by INVALSI studies (Deakin Crick, Stringher, Ren, 2014; Stringher 2016; Stringher, 2021), with a focus on cognitive data on the opinions of Italian teachers. In this sense, we consider the results of a deductive and interpretative analysis of a number of interviews with Italian preschool, primary and secondary school teachers, which aimed to analyse the L2L representations of the interviewees, distinguishing their conceptions and functions, and to investigate the presence or absence of specific criteria (activities and orientations) that exemplify how L2L is promoted, in accordance with theory (Brito Rivera et. al, 2021).

Method. The survey was carried out through the collection of self-selected projects from a small group of teachers who had declared that they attended an L2l course. In order to capture salient aspects of the L2L concept, an interview was carried out to find details of the individual projects. The interview (Scipione, 2021) focused on the analytical and circumstantial description of the project, and only after gathering information on the project did the interviewer move on to a question that fostered the in-depth exploration and explication of the interviewee's idea of learning to learn. which aimed to analyse the L2L representations of the interviewees, distinguishing their conceptions and functions, and to investigate the

presence or absence of specific criteria (activities and orientations) that exemplify how L2L is promoted, in accordance with theory (Brito Rivera et. al, 2021). The question is: “In your opinion, how is this project related to/promotes learning to learn?". Other questions and subsequent interview prompts were instead intended to trace the presence of certain components or dimensions of L2L in the projects, selected on the basis of their presence in various frameworks and their practicability in school contexts (Hoskins and Fredriksson, 2008; Caena, 2019; Caena and Stringher, 2020). In particular, the examination focused on the tasks that had been proposed to the learners and on the form that certain elements of L2L took, namely self-management, self-control, self-reflection and self-evaluation of learning (Scipione, 2021). We will limit ourselves here to an examination of the teachers' conceptions manifested as a result of the above-mentioned question.
Results. The study took into account 16 projects (pre-school, primary, junior and senior secondary school), varying in duration and consistency. They covered different subjects, from Science Education to Art to Technology, and also included projects focusing on the acquisition of the study method. The researchers recognise aspects that can be traced back to the classic areas that make up L2L, namely cognitive, metacognitive and socio-affective. From the central question of commenting on the project from the point of view of L2L competence, calibrated on the project itself and whose answer had to find argumentation and feedback, at least three main aspects emerge, which are both recurrent, relevant from a pedagogical point of view, and partly de-centred with respect to the classic dimensions of the construct. The first element is the method, in particular the study method but not only: hence the methods, i.e. of instructional design or problem solving, to be equipped with and spent in multiple contexts. Secondly, a methodological novelty factor (represented by the project) of the 'activist' type, which triggers a concrete operationalisation of the pupils (projects, artefacts) and determines their strong motivation to learn. The third factor is the group, and specifically group and team work as the main opportunity for learning, and above all for awareness. As can be seen, these are conceptually inhomogeneous concepts. Method-centred projects imply a view of learning in perspective, i.e. they make L2L coincide with being equipped with tools that enable one to master texts or deal with problem-solving situations in the future. The reference to motivation linked to a didactic novelty makes L2L coincide with the current experience of interest, of satisfaction in learning, i.e. having a satisfying experience on an affective and cognitive level. The reference to the group, finally, makes the L2L coincide with a 'new' way of learning, one linked to the sharing and collaboration practices that a given (challenging) task facilitates and promotes, and which elicits both cognitive and social and affective acquisitions. The data that have emerged may represent only a first partial typology of conceptions of competence on the teachers' side, which subsequent research could extend, but may prove useful in imagining training that keeps away from the risks of abstractness.

***Keywords:*** learning to learn, teachers’ conceptions, teaching practices

## Learning to Learn: training future teachers

**Lucia Scipione – Annamaria Gentile – Agnese Vezzani – Pierina Giovanna Bertoglio**

Introduction. Learning to learn (L2L) is considered a core competence for living and working in the 21st century. It is a multifaceted competence that helps citizens to cope in a volatile, uncertain, complex and ambiguous environment (Laukkonen et al., 2019; Stringher, 2021). As reiterated in the LifeComp framework (Sala et al., 2020), L2L is a composite competence that encompasses within it cognitive and metacognitive skills (e.g. study strategies, reflexivity, awareness of one's own learning, and self-regulation) and social-affective and motivational skills (such as adaptability, collaboration, self-confidence, and self-control) (Caena, 2019). From a policy and pedagogical perspective, the discussion on learning to learn, which combines cognitive and noncognitive aspects, is aimed to define the construct and identify strategies to be able to promote and evaluate it. How to support teaching that (intentionally) promotes the development of L2L competence in different school grades and also teacher training in this direction? Studies on learning to learn that have contributed to this line of research, particularly on teacher education, have worked on teachers' conceptions (Hounsell, 1979; Waeytens et al., 2002; Stringher, 2021; Gentile, 2022) and on teaching and evaluative practices to foster L2L (Stringher, 2021), but they have also focused on the documentation of good practices (Scipione, 2022) and on teachers' professional development

(Marcuccio, 2009; 2015; 2016). In addition, some experiments on competence-based teaching give useful indicators in terms of research and application, including to the curriculum, of resources for the design and assessment of L2L competence (Marcuccio, 2009; Batini, 2012; Kloosterman, Taylor, 2012; Mannion, Mercer, 2016; Capperucci, 2020). The work presented here starts from the recognition of the need for teacher training regarding L2L competence giving future teachers tools to promote this competence.

Research object and hypothesis. Hence, the idea of a training initiative, characterized by a specific modality, the effectiveness of which was intended to be tested. Specifically, a training methodology for teacher education was set up to combine theoretical knowledge and teaching practice. This study sought the most suitable ways to direct future teachers to work in favor of L2L teaching in preschool and primary school. Then, the organizational structure of the internship projects, which provides constant control over instructional design and scientific supervision, ensured a controlled context of intervention. As for the L2L construct, it was established to present a conceptual framework (Scipione, 2021) with indicators considered a priority, in addition to some contextual elements and practicable tools. The hypothesis of working on overcoming a "narrow view" of learning and implementing meaningful classroom experiences on distinctive elements of L2L competence is sustained by the idea of the effectiveness of exercising the reflective dimension, guaranteed in the multiple stages of monitoring instructional design.

Data. INVALSI has been promoting a debate for years on how to promote the L2L key competence in school and training contexts (Deakin Crick et al., 2014; Sala et al., 2020; Stringher, 2021). Considering that INVALSI does not produce national data on L2L, the data from which it is feasible to start is precisely the outcomes of that research that has investigated competence from the perspective of teachers and their conceptions as well as teaching and evaluative practices (Hounsell, 1979; Waeytens et al., 2002; Marcuccio, 2009; 2015; 2016; Capperucci, 2020; Gentile, 2022; Scipione, 2022). This is framed by: the presence of "narrow views" of L2L, focused on study techniques and strategies (Stringher, 2021); traditional teaching practices; and few meaningful classroom experiences on the connotative elements of the construct (Stringher, 2021), very often present as extra-ordinary experiences (Scipione, 2022).

Method. The training initiative presented here is situated within the project matrix defined in the Bachelor of Science in Primary Education program of the University of Modena and Reggio Emilia (Unimore). Unimore internship projects engage students in a general training part and a specific training part.

The L2L internship project was developed in 4 phases (initial training, intermediate training, intermediate review of individual project, and final sharing of the experience with presentation of final papers) and involved 14 students from the 4th and 5th year of Primary Education Sciences, as well as 2 internship-organizing supervisors and 3 researchers, for a total of six meetings between November 2022 and May 2023.

In the first phase, a project-specific training phase, the topics covered the different L2L definitions found in the literature both in national (MIUR, 2012) and international (European Council 2006; 2018; Sala et al., 2020) policy documents. On the side of didactic and design transposition, some explanatory indicators for a possible implementation of L2L in preschool and primary school were selected and presented, namely central indicators such as self-management, self-control, self-assessment, reflection on one's learning processes, and transversal indicators such as collaboration, responsibility, and motivation (Scipione, 2021). From these, it has been proposed to think over the definition, knowledge, skills, and attitudes of competence on the basis of age, and to reflect on what is meant by: L2L in school, priority of the teacher's conception of learning, and relevance of the design of environments and contexts for the meaningfulness of experiences. The choice of no more than two "indicators" and the reiteration of activities were a prerequisite to be met in the design.

Then, the students were asked to draft a Unit of Learning (USR-ER 2018) related to the themes, contents and methodologies proposed by the project team, paying attention to the promotion of innovative and laboratory methodologies. The Unit of Learning did not have to correspond to one or more activities but, inserting itself in the framework of competence-based teaching, required the use of authentic tasks and the promotion of interdisciplinary paths involving authentic assessment. In the intermediate training meeting, the students' initial project proposals were discussed, based on the congruence of the teaching proposal with the theoretical assumptions and the tools and strategies related to L2L. The final step was the presentation of the realized project proposals. This opportunity supported the discussion on L2L and some suitable teaching strategies to foster the development of this competence in preschool and primary school.

Results. Although this is a preliminary study on a small sample, it is worth pointing out some innovative elements because they are promising for trainee teachers in learning-to-learn oriented teaching. First, we explored a methodology from which we can derive useful insights for the future, and this is possible by

means of the documentation of the monitoring phases, the analysis of the students' Units of Learning, the impressions of effectiveness from the supervisors of the experience, but also from the students themselves. The model actually seems promising for overcoming the "narrow view" of learning and L2L and implementing meaningful classroom experiences on distinctive elements of L2L competence. The exercise of the reflective dimension and constant discussion with experts (from Unimore's Department of Education) ensured a positive outcome of the monitoring actions, as reflected in the instructional designs produced. The initial training was an opportunity to make choices about the complex construct of learning to learn, including in itinere, to bring it back to a more procedural and operational level (Marcuccio, 2016), and to deal with some best practices. The researchers then evaluated and discussed, with the group of students and tutors, the consistency of the projects from the point of view of the selected theoretical elements and proposed ideas for re-design in itinere. The last stage was the opportunity of a collective presentation of the projects, in addition to the evidence of the course outcomes, to reflect on the variety of possible actions and activities that can be legitimately ascribed to L2L teaching. In fact, there was a certain variety and originality in the project proposals produced by the students, who made visible the approach of future teachers to the specific content of L2L and to the project's methodologies (producing, for example, a change even in the terminology used) and, during the moments of confrontation and in the final evaluation of the enjoyment of the experience, the awareness of the non-necessity of working jointly on all aspects of competence without cheating its multi-componential nature. Finally, for the entire scientific project team, the training initiative represented a positive opportunity for collaboration between schools and universities, to enhance the innovative practices of schools and, at the same time, to contribute to research on the topic of the professionalism of teachers.

***Keywords:*** learning to learn, pre-service teacher training, preschool, primary school

## "Activities involving students" for Secondary School Teachers in Italy: noncognitive aspects related to L2L

**Hugo Armando Brito Rivera**

Learning to Learn (L2L) is a key competence for lifelong learning. Different local and international organizations have continuously pointed out the relevance of teaching new generations this competency, justifying its importance as an educational goal in the 21st century and a crucial aspect for economic development and social welfare. Although L2L has been studied during the last fifty years, this concept preserves its suitability to face the challenges of current societies, such as volatility, uncertainty and rapidity of changes produced by technological, economic or social development (Brito, 2022). Currently, this concept has influenced educational policies and curricula on a global scale, from the influence of organizations such as the European Commission (2007, 2018), the Organization for Economic Cooperation and Development (OECD, 2009) and the United Nations Educational Scientific and Cultural Organization (UNESCO, 2013). In international perspective, this competence has been considered in the school curriculum from early education to university education. This is the case, for example, in countries such as Argentina or Mexico in Latin America (Patera, 2018; Stringher et al., 2019) or Spain and Finland in Europe (Deakin Crick, Stringher & Kai, 2014).

A predominant research line in the study of L2L competence falls on the metacognitive and self-regulatory aspects associated with students' learning skills and "study techniques" (Marcuccio, 2016). In contrast, its study from a cultural and everyday perspective related to the function and role of teachers has been developed to a lesser extent (Stringher, 2021). The present contribution argues that L2L formation in educational contexts is mediated by representations and meanings linked to non-cognitive aspects (Hautamäki & Kupiainen, 2014) that, together, play a key role in forming part of the reasons why teachers carry out certain teaching-learning activities (Bruner, 1990; Rogoff, 2003; Sfard & Prusak, 2005). Such a premise, of a cultural nature, is necessary to complement the predominant research lines of study on L2L, traditionally focused on students' individual reflective and self-regulated learning capacities (Marcuccio, 2016). From the paradigm of qualitative and sociocultural research (Schwandt, 2000), teacher representations and meanings represent an object of study of notable interest from the hypothesis through

which it is proposed that L2L is present implicitly or explicitly from the function that such constructs have in the didactic choices that teachers make as part of the school activity (Brito, Torti & Carvalho, 2020).
Involving students, fostering interest in the object of teaching and sense making about the activities and epistemic objects of learning are three basic components to protect the motivation to learn and activate the ability to continue learning. Such an approach, represents an essential part of the competence of learning to learn framed from the sociocultural perspective (Well & Claxton, 2002; Claxton, 2014; Ajello & Torti, 2019). The representations and meanings that teachers express as a function of the school activities they carry out can act in favor or against the promotion of the ability to continue learning. From this point of view, the didactic choices that teachers make are based on certain senses and meanings that have the potential to "generate" the competence of AaA as they are inscribed in the daily school dynamics (Ligorio & Pontecorvo, 2010; Vygotsky, 1986). The inquiry about the cultural aspects involved in this process, observable in the reasons why teachers perform certain school activities, represents a developing field of research.
The aim of the present study is, therefore, to identify teachers' representations and meanings associated with school activities that "involve" students and can be considered as being related to L2L competence. Results are presented from the analysis of semi-structured interviews conducted with Italian teachers belonging to the second grade secondary school, as part of the project "L2L in Italy, Europe and Latin America" coordinated by INVALSI during 2018-2020 in Brazil, Ecuador, Spain, Italy, Mexico and Uruguay (Stringher, 2021).
A total of 127 interviews were collected in the six participating countries, of which 40 were in Italy (Stringher, 2021). A descriptive and deductive analysis (Rädiker & Kuckartz, 2020; Richards, 2005) was conducted on 12 interviews from the total corpus of interviews conducted in Italy referring to Secondary School teachers. The overall aim of the analysis was to identify aspects traceable to L2L competence in the interviewees' responses in relation to the second thematic core of the interview guide: "Activities involving students" (questions 3, 4, 5 and 7, vid. Torti, Brito & Patera, 2021). The research question that guided the analysis was: What are the culturally valued characteristics of activities that do or do not involve students according to the interviewees' point of view? To answer such a question, an excerpt coding strategy was conducted based on three categories; 1) Involving activities; 2) Non-involving activities; 3) Cognitive and non-cognitive aspects for or against L2L (Hautamäki & Kupiainen, 2014). The data analyzed in the present contribution correspond mainly to the third category.
The results indicate the prevalence of non-cognitive aspects of a socio-affective nature. Among the non-cognitive aspects that can be attributed to the formation of L2L are, for example, giving prominence to students, demonstrating one's own interest as a teacher in the subject taught, encouraging student autonomy and collaborative work, motivating students, creating meaning around the subject taught, activating the desire to learn and helping to learn. Among the non-cognitive aspects that can be considered against L2L development are, for example, teacher rigidity, frontal class, passive questioning by the teacher, loss of meaning of the subject taught and lack of practical purpose in daily life. It is concluded that, for most of the teachers interviewed, involving students entails an "active" teaching model, while lack of involvement is related to the frontal class model.
The set of non-cognitive aspects identified are eloquent of the cultural connotation of socio-affective emphasis that characterizes the activities through which the teachers interviewed seek to "involve" students at this school level. The data analyzed exemplify how the formation of L2L competence is significantly related to educational aspects of a relational and interactive nature. In this sense, the data point to the importance of educational interaction in engaging students and developing the capacity to continue learning. The results have the potential to inform educational policies and teacher training programs oriented to the development of L2L competence while providing a useful data surface to strengthen its analysis from a sociocultural perspective.

***Keywords:*** learning to learn, socio-cultural analysis, educational research, transversal competencies, teachers

# “Impossible pupils” in teachers’ representations: educational challenges for Learning to Learn in an international qualitative study

**Cristina Stringher – Salvatore Patera**

Introduction. Numerous national and international experts and organizations, from different angles, agree that Learning to Learn (L2L) is an essential hyper-competence in current societies (Ajello, 2018; Binkley et al., 2012; UNESCO Italian National Commission, 2020; Cipollone, 2014; Deakin Crick, Stringher and Ren, 2014; Demetriou, 2015; European Council, 2006; 2018; European Political Strategy Centre, 2017; Hautamäki et al., 2002; Hoskins and Fredriksson, 2008; James et al, 2007; Kupiainen, Hautamäki, & Rantanen, 2008; OECD, 2019; OECD PISA 2010; Smith et al., 1990; Stringher et al, 2021; UNESCO, 2013; UNICEF, 2012; 2018; US National Education Goals Panel, 1995; Wells & Claxton, 2002). Since 2006 L2L has been included among the European key competences, assuming in 2018 the current denomination of Personal, Social and Learning to Learn competence. Over thirty years of research in this field, however, have not provided unequivocal answers on how to support the acquisition of this competence in children, adolescents and adults, also due to definitional and theoretical difficulties, as well as the difficulty of educational systems to draft a L2L curriculum. For this reason, in 2019 fieldwork was conducted for a multi-year and polycentric international qualitative study, with researchers participating from 6 different countries. The study aimed at investigating the characteristics of L2L and the ways in which this competence is connotated in school practices of different cultures. Key informants were teachers from preschool through secondary education, that we interviewed individually.

Object and research hypothesis. This contribution explores a specific aspect of the interview guide developed within the project, i.e. a definition of what teachers consider to be an exemplary "impossible case" in the context of their daily teaching activities. The "impossible cases" are children and teenagers attending schools that teachers encounter in their classrooms, whom they try to involve and instead seem particularly reluctant. They are persons that the school system identifies, evaluates and supports, but with whom it is difficult to work and therefore present the risks of dropout, educational inequality, social and work marginalization. Specifically, in the 40 Italian interviews we try to deepen the characteristics of these cases, through teachers’ representations and their emotions, and we try to enucleate educational and didactic practices to support these pupils. Our hypothesis is that the teachers' reflections may aid the identification of a typology of these cases, with different characteristics. On this basis, it is possible to identify work paths using concepts and components of Learning to Learn as a key, so that no one is left behind.

Data. We analyzed the transcripts of the 40 interviews collected during the project from preschool and primary school teachers (15) and lower secondary school teachers (25), mainly from public schools (36), located in Northern (11), Central (14) and Southern Italy (15) of central (20), peripheral (17) and rural (3) areas and in high (5), medium (24) and low (11) socio-economic contexts. For each school, we also considered INVALSI information relating to positive (21) or negative (19) school added value.

Method. The study was carried out through a focused semi-structured interview. This tool, in line with the qualitative objectives of the project and with its theoretical-methodological assumptions, was chosen to detect teachers' representations from a personal and cultural viewpoint (Fideli and Marradi, 1996; Kvale, 2008; Melucci, 1998; Merton Fiske and Kendall, 1956; Serranò and Fasulo, 2011). Constructed specifically for this study, the interview guide aimed at eliciting authentic responses from respondents while minimizing the influence of social desirability and preconceived responses. For this reason, we employed plain language characterized by a colloquial style, devoid of professional labels. In this specific study, we investigated teachers’ responses to a particular question of the interview guide, asking them to provide an example of an "impossible case", along with a prompt asking interviewees to imagine the answers of their colleagues to the same question. We also tried to make connections among different parts of the interview, linking the answers to this question with those relating to teachers’ conceptions and activities on Learning to Learn. The researchers analyzed the interviews independently, with a combined top-down and grounded theory approach, to then compare and settle any differentiated interpretations, thus arriving at an inter-rater agreement and a richer interpretation of the results.

Results. An initial result was the identification of "impossible cases" in teachers’ answers. Out of 40 interviewees, not all reported examples of these cases: 6 teachers said they have not had any, 2 do not remember any such case, and 32 have instead encountered this type of pupil in their school career or in the classroom in which they currently teach. Although very frequent in our interviews, these examples are not

always present in the teachers' representations and this aspect deserves further scrutiny. According to these teachers' point of view, the characteristics of the "impossible cases" can be traced back to family or social issues (including pupils with non-Italian citizenship), or are attributable, in a simplified way, to the sphere of special educational needs. These are children and adolescents whose families struggle to communicate with and relate to their school and for various reasons do not have the opportunity to follow their children during their school trajectory. These dynamics determine attitudes of refusal and demotivation towards school in the pupils, if not of violence or true and proper social withdrawal. These cases often share a discontinuous school attendance, or an "absence" even when these children are physically present at school, or (in the case of immigrant pupils) have difficulty with the language of instruction. All these aspects lead to reduced learning opportunities and consequently to low school grades, in a vicious circle that is difficult to break, also due to the negative emotions they provoke in teachers: anger, impotence, resignation, frustration. Teachers thus report their inability to find adequate strategies. Some teachers, however, declare themselves resilient in the face of these adversities and affirm that one must not "give up", although in some cases they do not explain which strategies they implement. Finally, we consider the contribution of L2L in taking charge of "impossible cases" from an educational, didactic and training viewpoint, calling into question not only teachers' role of but also that of the wider community, so that this task of supporting "impossible cases" is not only confined within school premises. Trajectories of research and pedagogical reflection emerge regarding the educational and teaching practices adopted by teachers in relation to "impossible cases" and the individual or collective strategies for their solution. Our contribution may be useful for starting a dialogue with schools on "impossible cases" and on ways to support them through Learning to Learn, also thanks to the interventions supporting key competences that can be foreseen within the National Recovery and Resilience Plan.

***Keywords:*** learning to learn, educational inequality, school dropout, teachers' representations, educational and didactic practices

# THEME 5. EDUCATION AND SOCIAL MOBILITY

**ORGANIZER: INVALSI**

**COORDINATOR: DANIELE VIDONI**

**NOVEMBER 23RD: 4.30 P.M. – 6.30 P.M. {ROOM 1 – RESEARCH SESSION 5}**

---

## Born Different: Starting Differences and Academic Career Outcomes of Students through INVALSI Data

**Gian Paolo Barbetta – Luca Stella – Lorenzo Vaiani**

Introduction. Do schools compensate for or exacerbate the differences, determined by socio-economic and cultural background, with which students access the education system? Since the publication of the Coleman Report (Coleman et al., 1966), the literature has been questioning the effect that schools and the education system have on learning inequalities. The results still appear uncertain and are influenced by both data availability and differences in analytical approaches (Dumont, Ready, 2020).

Object and research hypothesis

In this study, we try to answer our research question by analyzing what we have termed "intragenerational school mobility": the evolution of each student's learning levels during their educational journey compared to those of their same-age peers (the cohort). Our hypothesis is that the low level of cognitive skills with which students enter the education system is not easily compensated for during their schooling, particularly for certain disadvantaged populations such as non-Italian students or children from socio-economically or culturally disadvantaged families. For this reason, at the end of compulsory education, the inequalities between students may not have decreased.

Data Used. To observe intragenerational school mobility, we used the results of the INVALSI tests in Italian and mathematics, following the generation of students who, in the school year (SY) 2012/13, were in their second year of primary school. We followed this cohort for six years, encompassing primary schools and lower secondary schools, until SY 2018/19 when they were in the third year of lower secondary school. We started with this cohort, consisting of over 500,000 individuals, and were able to track the educational path of over 300,000 of them until SY 2018/19.

A comparison between the initial population and those whose educational career we were able to reconstruct reveals that it is primarily male students, foreigners, residing in the South, with modest socio-economic and cultural backgrounds, lower academic achievement, and who did not attend nursery school, who are lost over time (their INVALSI test results in the expected year were not recorded). It is reasonable to assume that these individuals experienced academic setbacks, if not dropout, or left the Italian education system. This single piece of evidence already highlights some inequalities that will be further analyzed in our study.

Method. For each student, we calculated their position in the distribution of learning outcomes for each year considered, comparing it with that of their peers. In this way, we constructed "transition matrices" that analyze the temporal evolution of students' relative positions in each of the school years. We also reconstructed the same matrices for significant subpopulations (males and females; Italians and foreigners; children of educated/non-educated parents; in low/high ESCS families; those who attended/did not attend nursery school and preschool) with the aim of verifying whether the transitions from one relative position to another were more or less pronounced in different subpopulations. In addition to these bivariate analyses, we conducted multivariate analyses to test the robustness of the obtained results.

Results. The main results of our analysis are summarized in the following tables and figures . In Table 1, considering the group of students who achieved the lowest academic results in second grade (those in the first row of the matrix), we observe that - six years later, in lower secondary school - 49% of these students (9.8/20.09 = 48.8%) are still in the group that obtains the lowest results (first column); thus, about half of the most disadvantaged students fail to improve their relative position. Furthermore, when looking at the group of students who obtained the lowest results in second grade (i.e., those in the first, second, and third quintiles), we note that over 90% of them remain in the same quintiles in lower secondary school (9.8+5.76+3)/20.09 = 92.4), indicating a strong persistence of the initial distribution of academic results. Finally, the number of students starting from the lowest position who manage to reach the group with the

highest academic results is extremely small: only 0.31% of the total population, which is less than 2% of the worst-performing students in second grade (0.31/20.09 = 1.6).

Table 1 – Transition matrix between Grade 2 and Grade 8 (%)

| | | Quintile at Grade 8 | | | | | |
|---|---|---|---|---|---|---|---|
| | | **1** | **2** | **3** | **4** | **5** | **Total** |
| Quintile at Grade 2 | **1** | 9,80 | 5,76 | 3,00 | 1,21 | 0,31 | 20,09 |
| | **2** | 4,38 | 5,83 | 5,28 | 3,48 | 1,24 | 20,20 |
| | **3** | 2,29 | 3,95 | 5,26 | 5,47 | 3,35 | 20,31 |
| | **4** | 1,61 | 2,34 | 3,80 | 5,57 | 6,18 | 19,50 |
| | **5** | 1,92 | 2,12 | 2,66 | 4,28 | 8,91 | 19,89 |
| | **Total** | 20,00 | 20,00 | 20,00 | 20,00 | 20,00 | 100,00 |
| | | Total popolation | | | | | **312.346** |

Source: Elaboration from INVALSI data

Consistently, when considering the students who rank in the highest quintile of the distribution of academic results in second grade (fifth row of the table), we find that 45% of them (8.91/19.89) are able to maintain their position in middle school (fifth column), and only a small proportion, equivalent to 1.92% of the total population (less than 10% of the students in the highest quintile in second grade), fall into the lowest quintile of the distribution in middle school.
These data suggest that the school results obtained in second grade strongly influence subsequent grades, highlighting the presence of low intragenerational school mobility and a strong persistence of inequalities. Starting one's educational journey from a disadvantaged position means facing a future where it is extremely difficult to improve one's relative position.
The figures from A to H (in appendix 1) provide a concise summary of the results of our analysis for specific subpopulations. It is evident that the probability of starting - and remaining - in the group of students facing the greatest difficulties is particularly high for foreign students, those without any parent holding a degree, and students from modest socio-economic backgrounds.

***Keywords:*** inequality, school outcome

## The social stratification of teachers' grading: How students' social background affect the way students are evaluated in school

**Moris Triventi**

Grading is a ubiquitous educational practice that characterizes the experience of students around the world and it is a central aspect of the relationship with their teachers. Teachers' marks are a crucial tool to provide students feedback about their learning progress and level of knowledge in a particular discipline or subject (OECD 2012). Teachers' marks also represents a visible feedback for students and their families, who cannot easily and directly observe students' learning and achievement. School marks are indeed powerful signals and they may influence students' motivation to study, self-confidence, and effort (McMillan 2001). Moreover, they provide information that can contribute to shape the educational choices made by students and their parents (OECD, 2012) and can be used by employers as a signal about the skills and quality of a potential candidate for a job (Johnes, 2004).
Teachers' grading practices are the result of a multifaceted assessment process, which may not only reflect the objective level of competencies achieved by students, but also the perceived students' effort, motivation, and even their behaviour (OECD, 2012). Indeed, school marks reflect daily interactions between teachers and students and can be used by the former in a "strategic" way (Costrell, 1994). A recent stream of research is showing that not only grading standards may vary across schools and classrooms, but teachers' may systematically grade differently students' with different socio-demographic characteristics, even when they display the same level of competencies in a given subject. Researchers have mainly focused on whether female and migrant students are implicitly discriminated by teachers in their grading practices (Lavy 2008;

Hinnerich et al. 2011; Kiss 2013; Lievore & Triventi 2023). While many works suggest that teachers are positively biased in making school recommendations in favour of upper class children (Batruch et al. 2023), surprisingly, recent scholarship has paid less attention to what extent and at which conditions, students' social origin affect the way in which teachers attribute marks in school.
The aim of this work is to assess whether teachers' grading is socially stratified on the basis of students' social background in Italy and whether high-SES (socioeconomic status) pupils' putative advantage vary across school levels, school subjects and along the ability distribution. I will provide the most comprehensive evidence on this topic, both in Italy and internationally, by relying on more than 20 million student evaluations, covering a period of around 10 years and educational grades spanning from the early grades in elementary school till the end of upper secondary education. Moreover, I will investigate whether the gap between teachers' grades and students' test scores is larger or smaller along the competencies distribution, a rather neglected issue in the literature so far.
The sociological literature suggests that family and personal cultural capital - comprising knowledge, values, and cultural practices - is positively rewarded in schools. Students from privileged backgrounds have familiarity with cultural norms and practices aligned with school expectations, leading to better academic outcomes and evaluations from their teachers (Bourdieu 1987). The socio-psychological literature suggests that teacher subjectivity and personal biases can influence grading decisions. Despite efforts to maintain objectivity, teachers are not immune to their own preferences, beliefs, and experiences (Campbell 2015). These personal biases can inadvertently influence the evaluation of students, including those from lower socioeconomic and cultural backgrounds. Unconscious biases related to students' appearance, behavior, or conformity to societal norms can inadvertently impact grading fairness (Resh 2008). Additionally, the quality of the interaction with the school teachers could matter in the way teachers perceive the students and affect the way they assess their performance. Teachers might be more generous in grading those students who regularly attend classes, are well-behaved in the classroom, actively participate in discussions, demonstrate genuine interest in the topics taught at school and do their homework regularly (Tierney et al. 2011; Calarco 2011). Social and behavioral skills constitute resources that are unevenly distributed across students with different social background (Bowles & Gintis 2002). I thus hypothesize that higher social background students receive on average higher grades by their teachers than equally competent classmates from less advantaged backgrounds and this gap should be larger in the earlier educational stages, when is more difficult to isolate academic performance in the teacher's assessment and the classrooms are more heterogeneous in terms of social backgrounds. Following the compensatory advantage model (Bernardi 2014), I also expect that the grades-test gap is larger among low achieving individuals than among the high achieving ones.
The empirical analysis makes use of data collected by the National Institute for the Evaluation of the Italian School System (INVALSI) on the *whole population* of students in specific grades in primary (2nd, 5th), lower (6th, 8th) and upper secondary education (10th, 13rd). I pooled the data from the 2011-12 to the 2021-22 editions, to provide a comprehensive study of the phenomenon. On the methodological side, I propose a variation of the standard grade-equation model, in particular I introduce the novel application of causal mediation analyses techniques (VanderWeele 2015), which allows me to clarify the assumptions at the basis of the identification of the direct effect of social background on teachers' marks, appropriately controlling for test scores (mediator variable) and possible post-treatment induced variables. In conclusion, the presence of grading biases against students from lower socioeconomic and cultural backgrounds is a complex issue influenced by a variety of factors. It is imperative to raise awareness and foster a greater understanding of these biases among educators to promote equitable grading practices.

***Keywords:*** grading, inequality, social background, teachers, assessment

## How are inequalities generated in schools? An attempt to construct research tools and data

**Giovanni Abbiati – Giuseppina Le Rose – Elisa Manzella – Emmanuele Pavolini**

Introduction. In recent years, there has been growing interest in the quantitative study of the mechanisms underneath inequalities' reproduction in the school system. From a theoretical standpoint, the concept of

"tertiary effects" (Esser 2016) has been conceptualized, referring to the influence of the school system and its actors on inequalities in learning and educational pathways. At the same time, empirical studies focused on various school mechanisms reinforcing the reproduction processes of educational inequalities, both in Europe and in Italy (Argentin and Pavolini 2020). This evolution has been made possible by the creation and dissemination of large administrative datasets in European countries, built for evaluative purposes. Therefore, it is useful to reflect on these issues at the conference promoted by INVALSI.
As often happens, secondary analyses of information coming from standardized assessmets display disavantages too. Indeed, collecting information about the malfunctions of the school system by public evaluation institutions rasises several issues. The purpose of our contribution is to describe an attempt to integrate administrative data with data collected ad hoc by the research group, in order to enable new insights about educational inequalities and the mechanisms generating them.
Data and methods. A pilot online survey was conducted on a random national sample of 100 Italian lower secondary schools, aiming to interview school principals, teachers, and parents of students. The autonomy of this data collection, compared to the institutional one developed for evaluative purposes, provided us with ample freedom in designing new questions, scales and items. We tested them with the different targets involved in the survey.
Less than two-thirds of the schools invited to particpate responded to the survey, leading to selected samples of principals, teachers, and parents. These smaples are not large enough to allow inferences about the related populations, but sufficient to validate the measurement tools we used.
Results. The proposed contribution develops two separate analyses.
Firstly, we investigate the sample self-selection process of schools participating in the survey by relating administrative data on the original sample to administrative data on the subset of respondents. It is thus observed that the self-selection of schools does not appear to be neutral concerning characteristics of the schools/students' populations.
Secondly, we validate the most innovative scales used in the questionnaires for principals, teachers, and parents. We show that the research group's creative effort has yielded mixed results: some scales have proven to be robust and reproducible in future surveys, while others require deep rethinking.
Beyond the presentation of these pilot survey results, we draw implications for future large-scale surveys on inequalities in the school system.

***Keywords:*** inequalities, tertiary effects, survey

# THEME 2. DIFFERENT DATA SOURCES AND THEIR INTEGRATION: POLICY INDICATORS FOR THE EDUCATION AND TRAINING SYSTEM

**ORGANIZER: INVALSI – ISTAT**

**COORDINATOR: BARBARA BALDAZZI**

**NOVEMBER 23RD: 4.30 P.M. – 6.30 P.M. {ROOM 2 – RESEARCH SESSION 6}**

---

## The Italian School System layer by layer

**Lorenzo Maraviglia – Emiliano Campodifiori – Michele Marsili**

Introduction. Many statistical analysis on phenomena related to school, such as the incidence of drop-outs, the distribution of high-performance students, the distribution of choices, transitions to the labour market, are conducted without adequate geographical-spatial characterization.
Even when large and geographically qualified data sets, such as those released by INVALSI or the Ministry of Education, are available the spatial dimension of the analysis is limited to the inclusion of simple ripartitional (NUTS1 level) or regional ( NUTS2 level) variables in regression models. Notwithstanding that, in Italy, in the field of education (as well as in that of the economy) a very high share of variability of observable phenomena can be subsumed under the North/South dichotomy, there are in any case many other dimensions of territorial heterogeneity which are under- explored and deserve to be considered, such as the differences between urban, peri-urban, peripheral areas or the differences that characterize territories with similar morphological characteristics - for example mountain areas - located in different parts of the Country.
In more technical terms, failure to consider the geographical dimension leads to a systematic underestimation of the standard errors of regression parameters and, consequently, a risk of reaching too optimistic conclusions about the existance of significant effects.
Objective. This contribution aims to demonstrate how the spatial dimension represents an information layer that can be used fruitfully for the description, understanding and modelling of phenomena related to school and education.
Starting from the relatively trivial fact that schools are made up of sets of physical entities (buildings) that can be placed with relative precision in space, it is possible to imagine a path of progressive enrichment of the information framework, through which these georeferenced entities are assigned to a space characterized by physical (e.g. orographic features, traffic routes, communication infrastructures), social (the presence of a population with certain ethnic and/or educational characteristics), economic (e.g. concentrations of productive activities belonging to a specific branch of industry ) attributes. In addition to contributing to a more correct treatment and analysis of existing themes, this opens up new research perspectives, which incorporate the geographical/spatial dimension from the outset. What is at stake is the fundamental question of the territorial differences that characterize the outcomes and returns of education in Italy (as elsewhere); but also, the emerging theme of how space is being restructured under the pressure of processes such as migrations, population ageing, the transition to new production models.
Method and Data. The potential of this approach will be explored firstly by composing a framework of spatial information directly linked to the school world; secondly integrating within this framework the georeferenced information that begins to emerge from the new Italian Population Census (Censimento Continuo della Popolazione), for example in terms of education attainments, ethnic-national composition, socio-professional characteristics that qualify certain sub-municipal level portions of the territory. Another aspect that will be taken into consideration is the characterization of the space occupied by the schools in terms of real estate values, using for this purpose the georeferenced quotations provided by the Agenzia del Territorio (Italian Territorial Revenue Agency).
In order to provide an illustration of the perspectives of the proposed descriptive/cognitive setting, an analysis will also be presented that exploits the additional information relating to real estate values and the socio-economic characteristics of the territory to refine the estimate of the socio-economic background indicator (ESCS) released by INVALSI and re-evaluate on a broader basis the information potential of this variable for the classification of the phenomena and dynamics that pertain to the sphere of human capital formation.

Results. At the end of the analysis it will be possible to appreciate the strong links between the socio-economic features of a territory, its projections in terms of demographic trends and development prospects, the characteristics of its school offer, the results achieved by the students who attend its schools.
It will also be possible to evaluate how, in the sphere of education, the traditional North-South dichotomy intersects other dimensions of spatial articulation, producing a highly complex picture that we are just begging to observe with adequate tools and concepts.

***Keywords:*** Spatial Analysis, GIS, Socio-cultural background

## Where have they gone? Following a cohort of students through the intricacies of Italian contemporary society

**Lorenzo Maraviglia – Andrea Bendinelli**

Introduction. The need to integrate data from different sources can arise from heterogeneous needs. In part, it may derive from the need to trace processes that take place in distinct institutional settings or contexts; in part, it may arise from the need to look at the same phenomenon from different points of view, in order to grasp the complexity that characterizes contemporary experience. Education, conceived as a process of accumulation of human capital (in the form of knowledge, skills, attitudes, adherence to values) is an exemplary case in this sense. Education is a phenomenon which, in modern societies, has become centralized in specialized public and private institutions, characterized by a high level of formalization and a certain degree of self-referentiality.
However, starting from the moment in which it is recognized that the school, as it has been designed and built, is not structurally capable of guaranteeing the achievement of a certain level of results for all its students - and, consequently, it becomes necessary to hypothesize alternative channels (through vocational training, the world of work, the family, civil society) for the achievement of these objectives - the problem arises of how to broaden the capacity for observation and self-observation, so as to be able to detect , measure and evaluate things that take place outside the traditional institutional boundaries of the school – but which, in turn, contribute to the achievement of the overall goals of the system.
Moving from this approach, there is room for data, knowledge and analysis that go beyond the self-referential circuit defined by traditional organizational and bureaucratic structures. In other words, we are faced with the need to go beyond the mere school outcomes, in order to reconstruct the true educational, personal and existential outcomes of cohorts of young people living in our country, once the constraints of compulsory schooling have been relaxed and a plurality of paths become possible.
Object of the research, method, data and results. The object of the research that we are going to present is the reconstruction of the educational outcomes of a cohort of students who have recently attended Italian schools. The cohort is picked at the end of the lower secondary education cycle (8th grade); the picture is taken through INVALSI data which, on the one hand provide good guarantees of completeness (participation in the tests is compulsory), on the other hand provide a considerable wealth of information relating to the characteristics of individual students.
The outcomes are reconstructed five years later, when, under the assumption that the higher education course is running smoothly, the students are preparing to take the final exam (maturità). The gap between the number of the outgoing and incoming cohorts - over a quarter of the students are missing - opens up the space for a reconstructive investigation, which allows us to represent the extreme complexity of contemporary educational paths and experiences. Specifically, the survey makes it possible to identify the number and social profile of students who are still in school but have fallen behind, but also that of the small group of students who have already obtained their diploma through the innovative four-year courses made available by the system. From this reconstruction, based on INVALSI data and on the contents of the MIM student registry, however, about 1/6 (17%) of the initial cohort cannot be tracked in any archive related to the school world. These young people "lost" to observation represent the path that introduces us to the complexity of the contemporary situation. Among them, in fact, there are young people who have moved from school to the varied and fragmented system of vocational training in order to obtain a qualification that fulfils the educational goals set by the system; but there are also those who, on the contrary, have abandoned – in many cases definitively – their studies, and who represent the “missing” in the strict sense,

those who have prematurely interrupted the process of accumulation of human capital and who are preparing to enter the labor market and in adult society with few resources and uncertain prospects; finally, there are young people who no longer live in Italy because they moved to other countries, (mostly) following their families.
Estimating the numerical consistency of each of these groups is a primary requirement; this requires an effort to integrate other sources into the cognitive framework, first of all the ISTAT microdata from the Labor Force Survey (for an updated estimate of ELET) and the microdata on transfers of residence; then there is the whole very heterogeneous universe of data on participation in vocational training courses, which are divided into regional data bases that do not always communicate with each other, but which may find some moments of synthesis worthy of exploration, for example below form of OpenCoesione data.
The integration of INVALSI, ministerial, statistical, OpenCoesione data allows (even with some limits) to understand some characteristics of those who leave school and, more specifically, of those who have eventually abandoned the accumulation of formal human capital.
At the system level, the collation of data and their integration into a coherent framework, albeit characterized by a share of uncertainty, allows us to grasp - or, more properly, to make hypotheses about - the differential role that, in different parts of the country, the channels of school, vocational training, emigration play in determining the heterogenoust levels of incidence of the phenomenon of dropping out of studies which, as ISTAT estimates indicate, vary significantly between North and South, and between city and countryside.

***Keywords:*** complexity, outcomes, drop-out

## Projet «VALVAL»

**Simonetta Radi – Arianna Cecchini – Cristina Cosci – Simone Mancini**

Object and research hypothesis. The launch of the National Assessment System (Presidential Decree 80/2013) has marked, in Valdera, the activation of new relations between schools and between schools and local stakeholders in order to detect the most critical issues at the level of both the individual school institution and the local school system and to offer synergic answers to common problems, channelling attention and resources towards the improvement objectives deemed as priorities.
The ValVal Network Projec. 15 school institutions of different order and grade, since 2014/15, have been tackling together, in a systematic way, the self-evaluation process, comparing process and result data provided centrally by SNV, integrated with data available from other sources, and, for
each evaluation cycle, identifying some common priorities to be pursued through improvement initiatives implemented in a shared and synergic way.
Objectives:

- strengthen the tools for reading and analysing data, in order to achieve a complete and accurate self-evaluation process;
- develop comparison (benchmarking) between schools in the area and the dissemination of good practices (benchlearning);
- develop and disseminate a culture of data that accompanies each step of the evaluation pathway and becomes a habitual practice of teaching action;
- develop and disseminate a culture of checking;
- create, through a joint research activity, the conditions for a vertical continuity of teaching action;
- design improvement actions that make use of different experiences and professionalism and that, pursued in a homogeneous and synergic way, have a greater chance of success and impact on the whole school system in the area.

Network priorities identified:

- Decreasing school failure in the transition from junior to senior high school (Evaluative Cycle 2014/2019)
- The fragility of basic skills (Evaluation Cycle 2019/22 and 2022/25)

The Territorial ValVal Project. The positive results of the experience carried out by the schools of the network in the field of self-evaluation have prompted the launch of a new project, in addition to the first one, to represent the Valdera school system. This is the territorial ValVal, which intends to enhance the culture of quantitative analysis, comparison and implementation of common strategies, which represent the main value of the network ValVal experience, in order to create a working method and organisation at educational area level.The basic idea is to develop a synthetic tool that, based on a dashboard of indicators agreed upon, represents the different school systems from the point of view of learning outcomes, internal and external context data, so that comparisons can be made both diachronically and synchronically. All this with the intention, in the face of objectively detected criticalities, of orienting school policies in a possibly coordinated manner and of monitoring the effectiveness of interventions at the level of individual institutes and of the territorial system as a whole.
Dashboard of indicators:

- organisational dimension: both internal to the school (school buildings, learning environments; libraries, laboratories, infrastructures...) and external services (transport, canteens...)
- educational dimension: learning outcomes, school outcomes, integration, relations with the territory...
- territorial dimension: the educational poverty/wealth of the territory and its social cohesion, the opportunities for social and cultural services and the articulation of associative realities.Objectives:
- create a common working methodology between schools and local authorities which, starting from the analysis of data, makes it possible to grasp the strengths and weaknesses of the local education system and points towards the definition of common improvement strategies;
- disseminate a culture of data both for the identification of school and training needs and for monitoring the results of the improvement initiatives implemented;
- strengthen the school-territory alliance so as to make the comparison and search for suitable solutions to support the scholastic and educational success of the Valdera students become systematic.

***Keywords:*** self-assessment, data analysis, evaluation, promoting school success, educational projects, school system governance and strategy, social accounting, educational community

## Digital skills and learning in Italian and mathematics: a crossroads of inclusion, education and innovation

**Valeria F. Tortora – Patrizia Falzetti**

Introduction. By now it is known how much digital skills are increasingly decisive for the world of work and for becoming active citizens in today's society. But how important are they for academic performance?
Since 2013, every five years, the IEA (International Association for the Evaluation of Educational Achievement) promotes the ICILS (International Computer and Information Literacy Study) survey, which aims to evaluate the digital skills of class III students of the school lower secondary school and focuses on Computer and Information Literacy. Participation in the ICILS survey provides countries with reliable and comparable data on young people's development of digital literacy in the 21st century, it also allows investigating the level of digital and IT skills within countries in order to promote the development of programs education suitable for the digital age (Fraillon et al., 2019a).
The national surveys conducted by INVALSI, on the other hand, provide information on the levels of competence in Italian, Mathematics and English (Listening and Reading) achieved by students included in the Italian national school system. In addition to data on skills, INVALSI also collects a series of contextual information useful for creating family background indicators.
This work investigates the effects of digital skills detected with the ICILS international survey on the results in Italian, Mathematics and English (Listening and Reading) detected through the INVALSI tests.
Mayer argues that learning through digital tools allows achieving better results since in the learning process there are three fundamental assumptions underlying the cognitive theory (Mayer, 2014). The first assumption concerns the opportunity for students to organize information using two cognitive structures, namely the visual channel and the auditory one. The second assumption is the limited ability to process information in only one channel; therefore, it is advantageous if the learning context allows to activate both the visual and the auditory channels, in this way there is not an excessive cognitive load on a single channel.

The third assumption is motivation: students need to actively engage and intervene on learning content to acquire skills (Mayer, 2014).
Therefore, from Mayer's study it can be deduced that learning using digital tools obtains positive effects compared to "pencil and paper" teaching methods. Furthermore, the scholar also pays attention to the motivation to learn and the attitude towards learning, topics explored by the scholar Brown, who underlines the determination of the instrumental motivation for the achievement of satisfactory results in the acquisition of mother tongue (Brown, 2007).
Data used. Both surveys refer to students in the last year of lower secondary school in the 2018 survey year. The INVALSI tests, contrary to international surveys, are administered to all students of the school grades subject to the survey (grades 2, 5, 8, 10 and 13). In particular, grade 8, as well as grade 13, as a requirement for access to the state exam, have almost total coverage. The INVALSI tests, national or international, have the additional characteristic of "identifying" the student anonymously but unambiguously; this happens through the SIDI code ; this code makes it possible to link students not only over time but also between the different surveys in which they participate. This code is appropriately pseudonymised and attributed to all the students present in the INVALSI datasets. The merge between the ICILS sample dataset and the grade 8 national survey dataset has a coverage of 98.5%.
Object and research hypothesis. The main research object was to investigate the relationship between the digital competence measured by ICILS and the learning levels of Italian and Mathematics measured through the INVALSI tests. In particular, the study focused attention on the use of digital technologies and the digital activities that students claim to carry out at school and at home by completing the questionnaire at the end of the ICILS tests. The student questionnaire includes an entire section dedicated to digital communication tools useful for making one's own transversal skills useful for learning both Italian and mathematics. Furthermore, the increasing use of digital tools at school may be linked to the improvement of school learning.
Method. We proceeded with a preliminary descriptive analysis in order to identify the existence of any associations between some indicators concerning digital skills and the results of the INVALSI tests in Italian and mathematics.
In order to have a more precise picture of the effect of these indicators on learning, a regression analysis was used which considered the students' entry scores in the disciplines investigated among the descriptors, as well as context variables that often contribute to explaining the dependent variable, i.e. the students' learning outcomes.
Results. The study's analysis strategy relies on a rich set of individual, school, and class control variables that are not commonly available in ICT studies. For this purpose, the results in Italian and mathematics of the students of the INVALSI Survey 2018 and the digital skills detected by ICILS of the same students also involved in the international research were analysed; this was also useful to evaluate whether the obtained results were controlled by factors related to the student, the students' engagement with ICT, their socio-economic background and the technological tools used in the classroom. The results indicate that, overall, digital skills have a positive impact on educational outcomes: this effect is stronger for students with high academic performance and family background.

***Keywords:*** CBT, digital skills, learning Italian, learning mathematics

# THEME 3. INTERNATIONAL LARGE-SCALE ASSESSMENTS (ILSAS)
## METHODS AND RESULTS
### ORGANIZER: INVALSI
### COORDINATOR: MARIA MAGDALENA ISAC
### NOVEMBER 23RD: 4.30 P.M. – 6.30 P.M. {ROOM 3 – RESEARCH SESSION 7}

---

## Exploring factors affecting Italian students' attainment in standardised EFL tests

**Clelia Cascella – Sara Giannone – Simona Incerto – Adriana Timpone**

Introduction. Understanding the factors that affect students' academic achievement is critical for guiding educational practices, and it must therefore be prioritized on both the research and policy agendas. Among them, the role of students' family socio-economic and cultural status has received an increasing attention, especially in predicting students' attainment in mathematics and reading skills. Nonetheless, to our best knowledge, studies aimed at reconnoitring and exploring not only individual factors but also their intersectionality are scant, if any.

Research's aims and research questions. The proposed research aims to fill this gap by analysing large-scale assessment data collected by the Italian national institute for the evaluation of the educational system (INVALSI) specifically for the standardized English as a foreign language (EFL) test administered in Grade 5. In addition to the assessment data, we will gather information about students' personal characteristics and about their families and the classrooms they enrol in. In particular, our proposed research aims to answer the following two RQs:

1. How does students' family socio-economic and cultural status (ESCS) affect their performance on the INVALSI standardized EFL tests?
2. How does linguistic background impact students' performance on standardized English language tests?

Data and methods. To address our research questions, we analysed sample data from INVALSI, which was collected at the end of Grade 5. We employed a multilevel regression model, following the approach outlined by Hox (2010). This model allowed us to examine the relationship between students' academic achievement, measured using the Rasch model (Rasch, 1960/1980), and two key variables: family socioeconomic status (RQ1) and linguistic background (RQ2). In particular, as with previous studies, we may hypothesise that non-native students may face challenges in performing in EFL as it has been shown that L1 (i.e., mother tongue) competence and EFL outcomes are directly related (see, for example, Carroll, 1971). Nonetheless, we also acknowledge and suppose that non-native students (especially, second-generation foreign students) may have studied English before moving to Italy, and/or may come from anglophone countries, and thus be advantaged compared with Italian native students. An *ad hoc* exploration of such an intersectionality between students' attainment in EFL, their proficiency in Italian as measured via INVALSI tests, and their linguistic background is an example of the analysis proposed in our paper. Moreover, all the other information collected by INVALSI about students, their families, and the educational context within which they learn will be used as control variables thus allowing us to carry out that "comprehensive" analysis of factors affecting students' attainment in English as a foreign language. This comprehensive analysis will address the current gap in educational research and provide valuable insights into the factors that contribute to students' achievement in English as a foreign language.

Results. For the present study, 11 a hierarchical regression models were exploratively estimated to study the relationship between students' achievement in EFL test and different combinations (with or without interaction effects) of variables reporting on personal and/or contextual characteristics. Results showed that girls are slightly (less than a quarter of SD on a Rasch scale with mean 200 and SD 40), but significantly, advantaged compared to boys and that such a positive association is not affected by any other variable. As regards students' citizenship status, it is worth nothing that the citizenship coefficient is small and statistically not significant, thus suggesting that the variable 'citizenship' seems to be not associated with students' achievement in EFL test. Yet, results indicate that, when individual students' family socioeconomic and cultural background (ESCS) is accounted for, 'citizenship' become statistically significant, even if just weakly associated with students' achievement in EFL test. Such a result may suggest that 'citizenship' has not to be considered as a "pure" predictor of students' achievement in EFL tests, but as a proxy of other variables, such as (i) students' family ESCS, or (ii) students' linguistic background, as suggested by results

from Model 5, in which 'home language' (i.e., the language most frequently spoken at home) was accounted for. So, for example, and of course not surprisingly, students speaking English at home show a large advantage (more than half SD on a Rasch scale with mean 200 and SD 40) in INVALSI test compared to Italians (used as reference category for the variable 'home language'). In contrast, as with previous studies (e.g., Abbott, 2006), other languages - such as Arabic and Chinese - are strongly (around a quarter of SD), but negatively, associated with achievement in EFL tests. These results indicate that the variable 'citizenship' (with its three categories i.e., 'native, 'first-generation' and 'second-generation'), overtly and covertly used in educational research to predict students' achievement in other subjects (such as Mathematics), cannot be used to predict students' achievement in EFL tests as it provides information about students' migratory background, but not about their cultural (and thus the linguistic) background. Those reported above are just an example of the results from our proposed analysis that, considering also the other variables provided by INVALSI, purports of contributing to knowledge by providing a 'comprehensive' overview of the factors associated with students' achievement in EFL tests.

***Keywords:*** socioeconomic status, English as a foreing language, attaiment, INVALSI data

## Are older adolescents more rational in their trust assessment than young adolescents?

**Linde Stals – Ellen Claes – Maria Magdalena Isac**

Introduction. Political attitudes are said to be subjective to change during the impressionable years (Neundorf & Smets, 2017). The scarce research in the field indicates that attitudes in (early) adolescence differ from the attitudes of adults (Schoon & Cheng, 2011). This study wants to contribute to the field by assessing if young adolescents (14-year-olds) and adolescents (18-year-olds) differ in their perception of political trust. Theoretically, trusting attitudes can be placed 'on a continuum from rational evaluation to personal disposition' (van Elsas, 2015, p. 1161). First, rational trust is a knowledge-based evaluation of whether a specific object operates in line with a subject's criteria (Bauer, 2013). This relational aspect of trust implies that political trust attitudes might not remain stable during the life course: at 18 you are more knowledgeable to assess political institutions' performance. Moreover, this rational evaluation views different kinds of trust as domain-specific (e.g., social or political domains): levels might differ depending on the domain. Second, some scholars regard trust as a stable belief about the goodwill of others or as a general optimistic disposition towards society (Uslaner, 2002). Two consequences follow from this argument: specific trust attitudes will not fluctuate much during the lifespan, and we can speak of a general syndrome of trust which affects all trust decisions equally. The present study will test these conflicting arguments regarding the rationality of political trust among 14- and 18-years-olds.
Research objects: The main aim is to contribute to the literature on the development of political trust attitudes among adolescents across different age groups. We build on the three propositions that define political trust as rational and translate these into three hypotheses. Rational political trust is defined by three elements. First, all expressions of political trust (i.e. trust in different political actors and institutions) should be structured by a single latent dimension as they reflect the same underlying attitude of political trust. Second, political trust should be domain-specific, thus distinct from trust in institutions outside the political domain. Third, political trust is based on knowledge about the political domain itself. Our general proposition reads that older adolescents trust state institutions in a more rational way than younger adolescents. To investigate this, the following three hypotheses are tested: The internal consistency of political trust is stronger for older adolescents than younger adolescents (H1); The positive correlation between trust in state institutions and trust in non-political institutions is weaker for older adolescents than younger adolescents (H2); The positive correlation between political knowledge and political trust is stronger for older adolescents than younger adolescents (H3)
Data: Our study draws on two data sources. The first and main data source is the Flemish data from the 2016 International Civic and Citizenship Education Study (ICCS) (Grade 8 students: 14-years-olds on average: N= 2,880) (Schulz, Ainley, et al., 2018). The second data comes from the Flemish 2016 National Assessment on Civic and Citizenship Education (NACCE 2016). NACCE 2016 was organized in the same year with the aim of complementing ICCS 2016 findings and applied the same measures of institutional trust as

ICCS 2016 (Grade 12 students: 18-years-old, on average: N=4308). Acknowledging that context matters in trust evaluations, in our design, we keep context (time and country) stable (van der Meer & Dekker, 2011). Measures: To measure *trust*, respondents were asked to what degree they trust the following groups, institutions or sources of information: the national parliament, political parties, the national government, the police, the armed forces, the courts of justice, media, social media, and people in general. Items are rated on a four-point Likert scale ranging from "Not at all" (= 1) to "Completely" (= 4). Four distinct types of trust were differentiated: Trust in order institutions, Trust in representative institutions, Trust in media, Generalized trust. Civic knowledge what the only covariate included in the research, which was differently assessed among 14- and 18-year-olds via a timed cognitive test.
Methods: Data was recoded and merged using the IEA IDB analyser (IEA, 2017) and IBM SPSS (IBM Corp., 2021). All analyses were performed in Mplus 8 (Muthén & Muthén, 2017). The estimation considered the complex sample design (students nested within schools). To account for the ordered-categorical nature of the data, model estimation was run with the weighted least squares mean and variance estimator (WLSMV) with theta parametrization. To handle missing data, the full information maximum likelihood (FIML) method was used. Hypothesis 1 and 2 require Multi-Group (i.e., grade) Confirmatory Factor Analysis (MGCFA) and Measurement and Structural invariance Testing. Hypothesis 3 requires a Structural Equation Modelling (SEM) approach.
Preliminary results: To test H1 we use MGCFA to estimate a two-factor model of political trust (i.e., trust in order institutions and trust in representative institutions) for each age group separately. By constraining the factor loadings to invariance over groups and comparing the fit of the constrained and unconstrained model, we can assess whether the factor loadings (and communalities) vary across both grades. A formal test assuming equal factor loadings reaches good model fit (RMSEA=0.044, CFI=0.996 and TLI=0.996). This indicates that factor loadings are invariant across both groups, thus refuting H1. To test H2, we use MGCFA to estimate a one-factor model of trust and a four-factor model of trust (i.e., order institutions, representative institutions, media, people in general) for each age group separately. Consequently, we assess whether all selected indicators load better on one or four separate factors. Next, by constraining the factor covariances to invariance over groups (i.e., structural invariance) and comparing the fit of the constrained and unconstrained model, we can assess whether the relations between the factors vary across both grades. The results show that the four-dimensional model (RMSEA=0.064, CFI=0.984 and TLI=0.983) fits well and better than the one-dimensional model (RMSEA=0.114, CFI=0.960 and TLI=0.962). Furthermore, a formal test assuming equal factor loadings and covariances reaches good model fit (RMSEA=0.049, CFI=0.990 and TLI=0.990). This indicates that factor covariances are invariant across both grades, thus refuting H2. Lastly, to test H3 we use SEM to estimate how political knowledge (as exogenous variable) affects the two-factor model of political trust (i.e., trust in order institutions and trust in representative institutions) for each age group separately. By comparing the standardized regression coefficients, we can assess whether the coefficients vary across both grades. The results show stronger (positive) correlations in grade 12 (18-year-olds) compared to grade 8 (14-year-olds), thus supporting H3. We conclude that political trust is not a fundamentally different construct for old and young adolescents in Flanders. Political trust is a coherent and domain-specific attitude for both 14 and 18 year old Flemish adolescents. However, older adolescents seem to use their knowledge about the political system more when deciding about its trustworthiness than younger adolescents. In that sense, political trust among 18-year-olds is somewhat more rational (i.e., knowledge-based) than political trust among 14-year-olds. The discussion focusses on how these findings may contribute to the field of political socialization research. During this, we will also reflect upon the Italian context and how this study may advise current and future research and measurements among Italian adolescents.

***Keywords:*** political trust, rationality, early adolescence, structural equation model, international civic and citizenship education study 2016

# Comparative Analysis of Environmental Knowledge: A Focus on Italy

**Yuan-Ling Liaw – Sabine Meinck – Andres Strello – Rolf Strietholt**

The acquisition of environmental knowledge is a fundamental precondition for shaping attitudes and behaviors towards environmental sustainability among future citizens. Educational institutions, particularly schools, have long been recognized as central platforms for knowledge acquisition, and many countries' educational systems have acknowledged the need to develop frameworks for environmental sustainability competence (European Commission, Joint Research Centre, 2022). With the rising prevalence of misinformation through social media and the transformative potential of education, there is a critical need to examine the current state of environmental knowledge among children. This paper aims to address this research gap by utilizing international data to shed light on the environmental knowledge of students. The conceptualization and measurement of Environmental Knowledge (EK) as a cognitive outcome of education for sustainable development will be presented in this study. Large-scale assessments typically follow a standard procedure of developing assessment items based on a theoretical framework. However, in the case of the TIMSS 2019 data collection, the "environmental awareness scale" was established post hoc. This departure from the standard approach was prompted by the growing demand for data on environmental literacy among school students. Consequently, the lack of a structured item development process may have implications for the construct validity of the environmental knowledge assessment. Thus, this paper will thoroughly examine how EK is operationalized within the TIMSS 2019 eighth-grade assessment. Issues such as construct-underrepresentation and construct-irrelevant variance will be scrutinized, alongside detailed descriptions of individual items, including their difficulty levels based on percent correct. Moreover, the dimensionality of the assessment will be explored using bi-factor models. In order to promote substantial and sustainable changes in behaviors and attitudes towards the environment, it is crucial to ensure equal opportunities for all students to acquire environmental knowledge. This study will investigate disparities in environmental knowledge among eighth graders, utilizing TIMSS 2019 data. The analysis will specifically focus on various factors, including gender, socioeconomic background, immigration status (where applicable), and urban-rural divides within countries.This presentation aims to provide a comprehensive comparative analysis of environmental knowledge, with a specific focus on Italy. By juxtaposing Italy's environmental knowledge against that of other countries, a contextual framework will be established to interpret the findings within an international context. This comparative approach will facilitate a deeper understanding of Italy's strengths, weaknesses, and potential factors contributing to environmental knowledge among its eighth-grade students. Furthermore, this study will explore the linkages between environmental knowledge and measures of economic wealth. By comparing Italy with countries from both the global south and north, we will examine whether economic disparities have an impact on environmental knowledge levels. Additionally, environmental-related indicators, such as carbon oxide output and pollution levels, will be investigated to provide insights into the environmental challenges faced by countries and their potential influence on environmental knowledge. The research presented in this study aims to contribute to the development of effective strategies and interventions for promoting environmental literacy and sustainable practices among Italian students. The findings will have significant implications for policymakers, educators, and stakeholders in fostering environmentally responsible behavior among future generations, ultimately contributing to a more sustainable and resilient global ecosystem. By conducting a rigorous analysis of environmental knowledge and its comparative context, this study aims to inform evidence-based decision-making and policy development in the field of environmental education in Italy and beyond. Overall, this research study seeks to provide valuable insights into the state of environmental knowledge in Italy, compare it with other countries, and identify areas for improvement. By understanding the strengths, weaknesses, and factors influencing environmental knowledge among Italian eighth graders, stakeholders can work towards designing targeted interventions to enhance environmental literacy and foster sustainable practices, thereby contributing to a more environmentally conscious and sustainable future.

***Keywords:*** environmental knowledge, comparative research, TIMSS

# Digital Citizen Participation in Chile and Italy: description of gaps and their relationship with other forms of participation

**Daniel Miranda – Juan Carlos Castillo**

General Background**.** Digital technologies have reshaped power structures and public life. However, the promise of inclusion with the massification of digital tools has collided with inequalities and prejudices in the production, transmission and uses of digital media in terms of content and digital devices. The pandemic made visible the gaps that affect vulnerable groups at the population level and particularly in the school system. The digital divides observed for women, low-income people, rural communities, and older people became apparent (Beaunoyer, et al., 2020; Hargittai, 2021).

The rapid massification of information and communication technologies together with an accelerated development of Internet access has the capacity to influence the daily life of citizens. This accelerated development is challenging in at least two ways. First, some emerging phenomena that require theoretical and empirical attention are observed, such as the development of false content or the formation of Digital Citizenship (Chen, et al., 2021). On the other hand, not everyone has managed to access and use digital tools equally. Digital inequalities threaten the possibility that all social groups benefit equally from the opportunities offered by digitization (Halsper, 2021).

While the pandemic highlighted the urgency for better coverage, it also highlighted that digital inclusion is not complete with physical access alone. Digital inclusion is a major concern of the public and private sectors. A more nuanced understanding of how digital content is produced and consumed from the school age, and the factors that shape user experiences with digital tools, is key to designing policies and programs that affect digital inclusion. In this sense, this project intends to address the existing barriers and challenges in the effective implementation of digital inclusion in schools, in light of the concept of Digital Citizenship (hereinafter CD), in order to develop educational strategies and policies that promote equitable access and meaningful participation of all students in the digital society.

This investigation. The accelerated expansion of the use of digital tools in different spheres of life has modified the ways in which citizens manifest themselves in the public space, raising the need to understand the ways in which Citizenship is exercised using digital media (Kahne et al. , 2016). Particularly, given the extensive use of digital tools among the youth population, the question arises to what extent digitization deepens democratic participation, if it widens already existing gaps, or if it generates new inequality gaps.

The literature reveals a broad concern about understanding these new ways of exercising citizenship. However, there is a significant lack of empirical research in the school population and the absence of common instruments to address DC. In these two aspects, this proposal intends to move forward.

There is a wide discussion about how to define these new ways of exercising citizenship. An exhaustive review of the concept of DC in multiple disciplines suggests that the definitions can consider aspects such as: attitudes towards DC, civic participation using digital media, literacy, skills or competencies for the use of digital communication tools; use of technological and digital tools, or a combination of two or more of the dimensions mentioned (Chen et al., 2021). A common aspect in the definitions is that participation through digital media is considered as a relevant aspect. Following this emphasis, this paper focuses on DC as the use of digital media to learn about politics, comment or share content on social and political issues, following the work developed by Christensen and collaborators (2021).

This study proposes to address the following research questions:

- What are the levels of Digital Citizenship observed in the Chilean and Italian school population?
- How are individual factors (academic achievement, gender, educational level of the family and type of school) related to inequalities in the development of Digital Citizenship?
- What is the relationship between the development of Digital Citizenship with other forms of citizen participation in the school?
- In the context of this research, the following specific objectives are proposed:
- Explore and map the levels of Digital Citizenship in students from Chile and Italy.
- Examine the gaps of Digital Citizenship in terms of academic achievement, gender, educational level of the family, type of school and rurality.
- Examine the relationship between the levels of Digital Citizenship and other forms of Citizen Participation.

Using data from the 2016 ICCS International Study of Citizenship Education, developed with students from multiple countries around the world in grade 8, we analyzed the digital participation of students from Chile

and Italy in areas such as seeking information about politics, commenting on posts with political content, and sharing political content online. Additionally, DC gaps are described based on a series of relevant aspects, such as academic achievement, gender, educational level of the family and type of school. Finally, DC levels are expected to be associated with other Citizen Participation indicators.
Preliminary results, based in multilevel estimation, show that searching for information is a more frequent activity than commenting or sharing information on social networks. Along with this, relevant gaps are observed in digital participation according to socioeconomic level, but not according to gender. Finally, the results show that digital participation is positively associated with other forms of participation, such as participating in marches or the intention to vote in elections. The presentation will discuss the specific implications in each national context.

***Keywords:*** digital citizenship, digital inequalities, ILSAS, international study of citizenship education – ICCS

# THEME 7. FOSTERING THE EVALUATION CAPACITY BUILDING FOR SCHOOL IMPROVEMENT

**ORGANIZER: INVALSI**

**COORDINATOR: DONATELLA POLIANDRI**

**24TH NOVEMBER: 9.30 A.M. – 11.30 A.M. {ROOM 1 – RESEARCH SESSION 8}**

## Supporting teachers' self-evaluation and competence development for the use of artificial intelligence in the classroom: the Ersamus AI4T project

**Sara Mori – Jessica Niewint – Francesca Storai**

Introduction. Since the pandemic period, there has been a growing interest in the integration of technology into everyday school practice (Molinaar, 2021), making it even more necessary to ask how technology can improve teaching practice (Hattie, 2023, p.408). More recently, the topic of artificial intelligence and its use in education has entered the scholarly debate in a big way. Education must evolve in line with technological advances and the continuous influx of new knowledge (Pedro et al., 2019). Improving 21st century skills is crucial to bridge the growing gap between students' skills and the needs of the labour market (Brun-Schammé and Rey, 2021). However, adapting teaching methods to modern complexities is often a challenge (Zawacki-Richter et al, 2019; Korinek et al, 2021). The impact of artificial intelligence (AI) on education is a focal point of discussion (Selwyn, 2022;) and its integration into educational settings promises great potential to improve learning and to provide more support for teachers (Niewint-Gori, 2023). Yet many still struggle to understand AI, its algorithmic decisions, and the role of humans in interacting with AI (Fjelland, 2020). According to UNESCO (2022), changes should include the promotion of equity, shared prosperity and inclusiveness in education. Curricula should promote interdisciplinary, intercultural and ecological approaches, cooperation and solidarity, using digital technology for networking (ibidem). The use of AI in education has both potential and risks. Ethical implications need to be considered, including the use of AI to enhance and improve teaching, learning and educational management (Miao et al., 2021). The impact of AI on pedagogy highlights the importance of transversal human skills such as creativity, complex problem solving and collaboration to interact effectively with AI tools in life, learning and work (Carvalho, 2022). Thus, there is a need to support teachers in reflecting on these issues and on their own attitudes and competences in relation to the use of AI systems in the classroom. Teachers' perceptions are essential for the implementation of new tools and methods in schools (Owston, 2007). The opportunity to be supported in the implementation of new teaching models is an important factor in gaining confidence in the use of innovative practices (Capperucci et al., 2021). This applies in particular to the possibility of accompanying teachers in rethinking their role in the use of technological tools such as artificial intelligence in the classroom (Gillani et al., 2023). In this context, the Erasmus Plus AI4T - Artificial Intelligence for and by Teachers - KA3 Policy Experimentation Programme was born, a three-year experimental initiative that aims to explore and support the use of AI in education. The project is the result of a collaboration on two levels: on an international level, between a network of organisations (17 European research institutes, universities and ministries) dealing with innovation and evaluation in education from France, Ireland, Italy, Luxembourg and Slovenia. The hypothesis underlying AI4T is that tailor-made training can help teachers to deconstruct their preconceptions, develop their knowledge and encourage a conscious use of AI in an educational context. INDIRE, together with its partners, has contributed to the development of an evaluation protocol to measure teachers' attitudes and behaviours.

Object and research hypothesis. The aim of this research is to collaboratively explore and delineate tools and constructs that can enhance educators' ability to self-assess their proficiency in using digital tools and artificial intelligence (AI) in the classroom. To this end, data from interviews with teachers and school administrators conducted after the pilot of the AI4T project will be used. The hypothesis guiding this research is that insights gained from the experiences of educators who have already been trained in the use of AI in the classroom can enable the identification of specific requirements for training activities and tools. These insights could potentially be used to strengthen their capabilities in this area.

Data used. This study incorporates the insights of a headteacher and two teachers who were interviewed at the end of the pilot phase of the project. The input from these participants played an important role in

refining the tools used and shaping the subsequent pilot tests. The discussions are situated in the context of the proposed protocol for impact evaluation associated with the Erasmus project. This approach makes it possible to explore the views of trainers from the point of view of self-evaluation and improvement.
Method(s) employed. This study will use a counterfactual research design to explore the AI4T professional learning experience (Guskey, 2000), incorporating constructs explored in the Technology Acceptance Model (Davis et al., 1989; Kemp et al., 2019). The interviews are analysed using a primary content analysis method. The following areas were explored in conversations with teachers: AI-specific training requirements, the role of training in self-assessment of individual competencies, perceptions of AI application in the classroom and potential areas for growth, perceived key ethical risks, and suggestions for AI implementation in the classroom. The headteacher interview also explored systemic requirements for the subject and teachers' perceived motivation to engage with the subject.
Results achieved. Preliminary findings highlight the need for teachers to receive hands-on training, as AI is often perceived as a distant and potentially harmful concept. Through the training, teachers were given the opportunity to expand their knowledge of AI, which led to the dismantling of certain prejudices and facilitated a more nuanced understanding of the potential of certain tools. There is an ongoing need for collaborative discourse among teachers on these facets and exploration of the potential impact on students, particularly in relation to the learning process.

***Keywords:*** attitudes, training, artificial intelligence, evaluation model, self-evaluation

## School principals' viewpoint on a training program for supporting school self-evaluation

**Giuseppe C. Pillera – Letizia Giampietro – Donatella Poliandri**

Introduction. Since 2014, Italian schools conducts a process of self-evaluation, as required by DPR 180/2013. The introduction of a semi-structured self-evaluation report required to carefully consider school staff's evaluative skills and the training strategies to support evaluation capacity building (Poliandri, Freddano, Molinari, 2019).
According to the OECD (2013), key aspects for fostering a culture of evaluation in schools include: supporting during evaluative processes; networking with stakeholders in the local community; promoting spaces – including online ones – for discussion, exploration and collaboration within and among schools. Literature highlights further elements, such as the need for a decentralized perspective in supporting schools (Ehren et al., 2017) and the development of adequate skills among personnel (Schildkamp, Poortman, Handelzalts, 2016). Furthermore, research evidence in Italian context has shown: the urgency of moving from a descriptive to an interpretative/evaluative approach, aiming at improvement (Robasto, 2016); the necessity of proximity support, tailored to specific school needs (Fortini et al., 2016); the demand for training activities not only addressed to school principals (Fiore, Torelli, 2019).
On this framework, the *Valu.E for Schools* (VfS) project[1] aimed to strengthen school teachers' (STs) and school principals' (SPs) professional skills in self-evaluation. In this intervention-research, a sample of 45 schools (primary, low-secondary and comprehensive institutes, including both) was selected[2], 15 for each macro-region (North, Central, South), in which three training courses on self-evaluation was provided by three partners. These courses, one for each macro-region, were delivered between 2020-2021 using a peer learning or a situated training approach (Gomez Paloma, Poliandri, Giampietro, 2020) and were attended by 400 STs and 42 SPs. As part of the project's evaluation design, 13 online focus groups were conducted (Poliandri et al., 2023), four with SPs and nine with STs (Poliandri et al., 2022; Perazzolo et al., 2023).
Research object. In this contribution we compare the main results of the textual analysis conducted on the focus groups with SPs and with STs, exploring redundancies and peculiarities of viewpoints emerging from similarities and differences in the articulation of the discussion topics: 1) alignment of the provided training with schools' expectations and needs; 2) developed evaluative knowledge and competences; 3) promoted collaboration and peer learning activities; 4) dissemination of the learning outcomes within school.
Used data. Analyses were conducted on the corpus resulting from the transcription of the 13 focus groups[3]: one for each region involved in the case of STs, one for each macro-region plus an additional

catch-up one in the case of SPs. The overall sample of focus group participants consist of 37 SPs and 101 STs (on average, 9.3 and 11.2 participants per focus group, respectively). The sub-corpora contain 357 interventions by SPs and 782 by STs (excluding the moderator and closing phase). Each intervention, considered as a case, is associated with a set of background variables, first and foremost the professional role.

Methods. A computer-assisted textual analysis was conducted using a coding methodology (Adu, 2019; Pagani, 2020) inspired by an abductive approach (Fereday, Muir-Cochrane, 2006), which combines the framework derived from the stimulus questions of the focus groups (top-level categories: contextual information and expectations, characteristics of the training course, collaboration/debate, outcomes/impacts) with the emerging evidence from direct exploration of the empirical base (second-level codes, providing greater detail). Text data mining techniques were used to examine the coded corpus[4].

Results. The comparison of coded SPs and STs sub-corpora at a general level (i.e., the top-level categories) does not reveal many major differences about the general thematic articulation of group discussions.

However, significant differences emerge in topics structure at a more granular level (i.e., the second-level codes). Indeed, SPs, in comparison to the STs:

- offer a more extensive presentation of their own school context and self-evaluation work being carried out, put in relation to a more in-depth examination of expectations on the learning;
- comment to a lesser extent on the training course (especially on content and materials, methodologies and timing, issues that has been variously debated by STs, even with critical positions);
- have a greater awareness of the emergence of new training needs in the field of evaluation, which they consider both in the dissemination of the VfS project results to the entire school community (a process that has already started in some cases) and in the planning of future professional development activities.

To summarise, while SPs participate less than teachers in training (as is also shown by the Learning Analytics in the Moodle platforms on which the courses were delivered), we observe their (more or less) explicit adherence with the purposes, contents, working methods and organisation of the training. Moreover, in relation to the role played, SPs' responsibility for the preparation of training activities on evaluation aimed at all staff emerge, as well as an interpretation of the VfS as consistent with the training needs and as an evolution in continuity with school's evaluation practices.

Grasping these differences based on the role makes it possible to set up targeted courses according to SPs' and STs' specific responsibilities within the evaluation and governance processes of the school: on the one hand, it is a question of strengthening SPs' competences, so that they can conduct self-evaluation processes at the level of the entire school, and on the other hand, of supporting the teachers' evaluation and data reading for the improvement, as theorized in the preparation of the VfS intervention-research.

[1] "Valu.E Valutazione/Autovalutazione Esperta" Azione 2 - Progetto Valu.E for schools, Valu.E 10.9.3.A – FSE PON 2015-1, approvato con comunicazione MIUR prot. AOODGEFID/23772 del 15/12/2015.

[2] A two-stage stratified sampling was used: provinces were considered as stage I units (a random sampling selected one province for every intervention region: Lombardy, Piedmont, Emilia-Romagna, Latium, Tuscany, Umbria, Apulia, Campania, Sardinia); schools as stage II units (selection of five schools in each province through a random sampling).

[3] The automatic transcriptions produced through Cabolo (www.cabolo.com) were reviewed by VfS research team.

[4] Coding and analysis processes were managed through QDA-Miner (https://provalisresearch.com/products/qualitative-data-analysis-software).

***Keywords:*** self-evaluation, evaluative skills, professional development, computer-aided textual analysis, school staff

# Study of teachers' beliefs and practices about formative assessment in Emergency Remote Teaching. The role of INVALSI data

**Marco Giganti – Emanuele Marcora**

Introduction. The centrality of assessment in the learning paradigm is a theme widely developed in the literature and finds particular expression in formative assessment (Scriven, 1967; Bloom 1971; Black & William 1998; Weeden, Winter, & Broadfoot, 2009; Brown, Campione, Webb, & McGilly, 1992).
The pedagogical potential and social effects of such an assessment have also long been considered by UNESCO and the Council of Europe (CoE), which, during the first phase of Covid-19 spread, saw fit to incentivize its use to support the involvement of students then their learning process (Engzell, Frey, & Verhagen, 2020; UNESCO 2020c, 2020d, 2020e; Hughes, 2020).
In spring 2020, national governments decided to close schools and abruptly switch to distance learning as measures to counteract infection and not disrupt educational activities. In most cases, school systems have adopted DL through e-learning or videoconferencing platforms (UNESCO, 2020a, 2020b). In this regard, it is appropriate to take up the distinction postulated by Hodges et al. (2020) between online distance learning and emergency remote teaching (ERT). The latter is a temporary and alternative mode of education delivery in crisis circumstances. The main objective is not to recreate an educational ecosystem but to provide temporary access to education and teaching aids.
Central to the management of such education has been teacher professionalism, led from conviction and turned into practice. The most problematic aspects emerging from empirical research (Richardson, 2002) concern the link between conceptions and effectiveness of teacher education (Balduzzi & Vannini, 2008). The literature indicates two approaches for the analysis of beliefs and convictions referring to the influence one of individual factors, the other of organizational factors. According to V. Richardson and P. Placier (2002) it is appropriate to integrate them as they support the researcher in studying the ways in which personal experiences act on teacher knowledge acquisition during initial and in-service training. The links between teachers' conceptions and the effectiveness of teacher education are complex issues and postulate training related to teaching practices (Balduzzi & Vannini, 2010). To this end, the analysis is expanded to studies inherent in Teacher Change (Floden, 2002; Richardson, & Placier, 2002) to understand whether changes in beliefs precede or follow those in practices; or whether beliefs and practices are mutually interacting and synergistic (Goffman, 1973; Peterman, 1993). As R. Chin and K.D. Benne (1969) state, on the one hand there are empirical-rational approaches aimed at showing teachers the existence of effective practices, promoting new awareness therefore new beliefs; on the other hand, there are normative-reductive approaches that explicate sociocultural and pedagogical norms and values in order to solicit teachers' collegial reflection and orient their beliefs and choices of action.
Research object and hypothesis. The research took place in the Lombardy context - specifically in the metropolitan city of Milan - and involved, through non-probabilistic sampling of volunteers, three all-inclusive institutions (from primary to secondary school). The main objectives referred to three different areas: theoretical-conceptual to identify whether and what relationship exists between teachers' beliefs and practices on formative assessment in ERT; methodological to analyze and describe the school context, investigate teachers' beliefs and practices on formative assessment, student engagement and learning, and ERT; and transformative to engage school leaders and teachers as active participants in the research and, at the end of the data analysis, design possible training paths in relation to the issues that emerged from the research. Three general questions were identified to guide the study: what is the school context in which teachers might have used formative assessment during ERT? (RQ1); What are teachers' beliefs and practice statements about formative assessment, engagement, and learning in the ERT context? Is there a relationship? What kind? (RQ2); how does the school context relate to teachers' beliefs and practice statements? (RQ3). After defining the problem and research questions, two general research hypotheses were formulated: there is a correlation between teachers' beliefs and practice statements about VF, learning, and student engagement in the ERT context (GHp1); school context is related to teachers' beliefs and practice statements (GHp2). The formulation of the following specific research hypotheses is intended to assume precision and concreteness with respect to the general ones: teachers' beliefs about learning are related to the application of VF in ERT context (SHp1); teachers' beliefs about student engagement are related to the application of VF in ERT context (SHp2).
Data used. Data were used from the semi-structured interviews with school principals, analysis of collected documents (Three-Year Educational Offer Plan, Self-Assessment Report, circulars, policy acts, etc.), a

questionnaire administered to teachers of all-inclusive institutions, and focus groups conducted with some privileged witnesses. In the research in general and in the introduction to the last phase in particular, INVALSI data from the 2019 and 2021 surveys were used in order to propose a comparison and initiate reflections on the performance of the research subject schools. Unlike the others collected, these data are characterized by their standardized nature and exogenous, non-self-reporting origin, lending further validity to the entire study framework.
Method. A multiple case study was conducted on the three identified institutions aimed at observing and describing their context, teachers' beliefs and practice statements about formative assessment, learning and student engagement during the ERT activated during the critical period of the Covid-19 pandemic, the relationship between them and the context. Specifically, qualitative-quantitative analyses were carried out on the results emerging from the semi-structured interviews conducted with managers, document analysis (Three-Year Educational Offer Plan, Self-Assessment Report, circulars, policy acts, etc.), the questionnaire administered to teachers, and the focus groups conducted with the School Headmaster, the Internal Evaluation Core and selected teachers by school order.
Results. At the time of the submission of this abstract, the semi-structured interviews with principals were conducted, questionnaires were administered, and focus groups were conducted with the indicated individuals, and qualitative-quantitative analyses were carried out on all the data collected.
In general terms, it is possible to say that in the three cases considered, there is a correlation and, at times, a causal relationship between teachers' beliefs and practice statements about formative assessment and the other constructs in question in the context of ERT and certain socio-demographic and contextual variables. Specifically, beliefs and practice statements about formative assessment are correlated with those of constructivist learning, student engagement, and are found to be consistent and applicable in the ERT context. Interestingly, the age and experience of the teachers are a conditioning factor as well as the teaching order and emergent situation. School institution also had an effect on formative assessment and learning beliefs: the numerically smaller, identifiably stronger institution with accompanying practices and teacher training was decisive on agreement averages. Primary school also proved to be favorable contexts for formative assessment and for what concerns student engagement and constructivist learning.
In all three schools the management of DDI was complex but was seen as an opportunity to innovate. From multiple angles of view, it emerges that beliefs are difficult to change especially when conditioned by lived experience. Training is considered useful but too distant from practice, weak in its applicability and too narrow in its timeframe. The socio-economic context of the area and the characteristics of the school have been decisive because where there is more availability of human, instrumental and economic resources, smaller size, stronger relationships it has been possible to activate greater processes of change and innovation than in other areas that are more fragmented, less cohesive and poorer in resources, despite the effort of the teachers and even more so of the School Headmaster.
The general and specific hypotheses formulated are thus tested in the contexts examined. The data collected and the analyses carried out confirm that the design identified to conduct this research proved to be appropriate since it allowed for in-depth investigation of specific and delimited contexts, and the triangulation of tools and visual angles attest to its validity.

***Keywords:*** formative assessment, teacher education, emergency remote teaching, beliefs, practices

## Impact analysis of a professional development on self-evaluation: first results

**Donatella Poliandri – Letizia Giampietro – Sara Manganelli**

The international literature points out that schools have different capacities in carry out of self-evaluation paths consistent with the goals set, in the ability to read and interpret data and in the ability to use them for the purpose of preparing improvement paths (Blok et. al., 2008). For this reason, the development of an evaluative culture in the places where learning takes place is assumed to be a strategic objective for the evolution of education systems at the international level (OECD, 2013; Giampietro et. al., 2016). In studies of school improvement, collaborative processes within and between schools are considered catalysts for educational change (Hubers and Poortman, 2017). In this scenario is the Value for schools Action Research (Giampietro and Romiti, 2019), as part of the Valu.E Self-Evaluation/Expert Evaluation Project, which aimed

to test the effectiveness of different training models to support school self-evaluation activities as outlined by the National Evaluation System (SNV).After an preliminary phase of preparing the framework and administrative procedures, the training activities with schools were developed in the midst of the pandemic emergency from spring 2020 to the end of 2021. The training and support activities were delivered by 3 economic operators identified through a public negotiation procedure.Participating schools were selected through a two-stage stratified sampling procedure; 15 schools were sampled in each Macro Area (North, Central and South), for a total of 45. An additional sample of 45 "control" schools was also identified using the same criteria. For each school, the Principal and Data Teams participated in the training. In order to value the impact of the training intervention on participants (teachers), it was decided to use a quasi-experimental evaluation design (cluster randomized trials, Raudenbush, 1997), with a group of schools that received the intervention and a group of control schools. Identified subjects from both groups, intervention and control, were administered two questionnaires, i.e., in the fall of 2020, at the start of the training by the Economic Operators, and in the fall of 2021, at the end of the activities.
In this paper we discuss the impact of the training intervention on questionnaire respondents regarding perceptions about the function of self-evaluation, attitudes toward evaluation, and collaborative behaviours among colleagues. Related to these scales are the intraclass correlation, which indicates the variance of the construct at the school level; the effect, at the school level, of the starting level of schools on the construct (pre-test) on the final level (post-test); and the effect, at the school level, of the training intervention, net of the starting level of schools. Data were analyzed using the Multilevel Structural Equation Modeling (Multilevel SEM; e.g. Cho et al., 2015) approach. A model was tested for each detected construct considering two levels: an within-school level and a between-schools level. The effect of the intervention was examined at the within-school level (between-schools level), controlling for the effect of the level at the pre-test on the level at the post-test, both at the individual level (within-school level) and at the school level (between-schools level). The sample referred to in the analysis consists of 544 subjects. The sample consists of all training participants who responded to at least one of the two surveys provided and belonged to schools in which at least 6 subjects participated in the surveys (57). The measure related to Participants' perceptions of the functions of self-evaluation were developed from the international literature on the topic (Poliandri et. al., 2019; Gaertner, 2004; Scheerens et. al., 2003). In total, the scale consists of 14 items on a likert scale from 1 to 5 steps (1 = not at all agree, 5 = fully agree). The 14 items aim to detect two functions of self-evaluation: the first is the knowledge and development function, consisting of 7 items, and the second is the reporting function, consisting of 7 items. The scale of attitudes toward self-evaluation was developed by Vanhoof and colleagues (2009) and adapted to the Italian context (Poliandri et. al., 2019), aimed at understanding the perceptions of Principals and Data Teams about the self-evaluation process (Block et al., 2005; Meuret & Morlaix, 2003). The scale consists of 12 items placed on a continuum from 1 to 6 and is divided into a dimension concerning purpose and one concerning sustainability. Collaborative behaviours are considered as those that are extra-role and for which there is no reward at the contractual level (De Longis et. al.,2018). The scale consists of 7 items on a Likert scale from 1 (never) to 5 (always) and has been adopted in several studies where it has shown good psychometric characteristics and reliability (for the Italian context see, for example, Alessandri, Vecchione, Tisak, Deiana, Caria, & Caprara,2012).The results of the analyses are differentiated in relation to the scales considered. Regarding the functions of self-evaluation, following the training intervention, there is an increase in both the perception of the knowledge and development function of self-evaluation and the reporting function in the group of schools that received the intervention alone. Participants in the training would appear to have gained a greater awareness of the functions of self-evaluation as both a knowledge and learning tool; at the same time, in the same intervention group, awareness of the reporting function of self-evaluation as a means of disseminating the results achieved to its target community also increases. Significant effect of the training intervention in the group of intervention schools was also found in the construct related to collaborative behaviours among colleagues. Indeed, the training methodologies favoured collaborative modes both within the group being trained and with colleagues in other schools. Following the training intervention, these behaviours were found to increase in frequency. This finding is also confirmed by other studies conducted on the same sample of subjects (Poliandri, 2022; Perazzolo, 2023). Attitudes toward evaluation, on the other hand, are not changed as a result of the training in terms of both purpose and sustainability.

***Keywords:*** self-evaluation, evaluation capacity building, impact evaluation, teachers

# Small schools for self-evaluation

**Graziella Arazzi**

Introduction. The contribution intends to highlight to what extent small schools (reduced complexes far from the center, multi-classes, mixed classes) - in their morphology and in the dynamics that unfold within them - contribute to strengthening attitudes and styles of self-evaluation within of the comprehensive institutions in which they are inserted. INDIRE, with the "Movimento delle Piccole Scuole", has given rise to research plots on the nature of an organizational and managerial form (not just didactic), which highlights the possibility of developing cultures and practices of widespread self-assessment, made evident by the following factors : planning for groups; possession of active research methodologies on the territories; use of discontinuous tenses in teaching; building educational communities, nurtured by informal networks with a focus on digital strategies; teacher training understood as peer research and experimentation of new professional contexts (cf. G.R.J. Mangione, G. Cannella, L. Parigi, R. Bartolini, Community of memory, community of the future. The value of the small school, Rome, Carocci editore, 2020). In this context, however, research on the evaluation aspects and on the contribution that school micro-realities (fluid structures, civic centers or training hubs) can provide to the institutes to which they belong has been almost absent to date. However, the fact that INDIRE formulated in 2019 - within the National Repertoire of professionalism for innovation - the figure of tutor of Small Schools can signal the debut of a new research sector, in which the realities investigated create processes improvement, not so much by overcoming critical factors as by recognizing the specificity of the values and points of view with which self-evaluation is played out. The tutor, trainer and facilitator, by promoting the analysis of the impact of the experimented innovations (multi-age, multi-level, combined classes), triggering relationships between isolated complexes and other locations, "participates in the school's self-evaluation processes and undertakes to redefine his intervention in relation to the improvement of the organization of the small school" (Programmatic document INDIRE, 2019). AA.VV., Towards didactic sustainability in small schools: interventions for training, professional development and scaling up of innovation, edited by G. Cannella, M. Garcia, G.R.J. Mangione, M. Repetto, <Training&Teaching. International review of education and training sciences>, XVIII, n. 1, 2020, p. 71-84, ISSN 1973-4778

Research object and hypothesis. The research takes place in the Ligurian territory and arises from the collaboration between Polo Schools for training and the Regional School Office. The research hypothesis is that where the NIV involves the teachers of small schools or includes such teachers within it, the following objectives are achieved: climate of sharing and comparison; mutual exchange on self-evaluation issues; greater effectiveness in governing self-evaluation processes; construction of reflective practices common to the various school levels; expansion of critical thinking on the dynamics of the entire school (https://www.INVALSI.it/value/index.php); enhanced ability to read data; identification of new indicators; creation of tools to define priorities and goals in a more understandable and coherent way; perception of training needs on the communication of the self-assessment and on the negotiation of its various phases.

Some research questions:

- To what extent does the self-assessment of the institution take account of small schools?
- To what extent do small schools promote assessment skills for school staff?
- In small schools, what approaches, methods, tools exist for strengthening self-evaluation skills?
- Constraints and opportunities, inside and outside the school;
- Role of networks;
- New professional skills that are necessary in small schools;
- New training models for teachers and operators in small schools;
- Evaluation and improvement: spaces for collaborative study with the aim of supporting reflective processes (for example, through peer to peer, virtual learning environments, the creation of networks with subjects and local agencies).

Data used. The data used refer to two activities:

1. a focus group - in online mode - dedicated to comprehensive institutes including micro-realities (multi-classes, complexes far from the center or in internal areas). Managers were asked to involve a member of the NIV and a teacher working in a small class in the discussion table. There were 27 participating institutes. In 8 cases the same school leaders also took part, with indications of distributed forms of leadership;
2. analysis of the answers to the in-depth questionnaire administered in the following phase.

Method. The methodologies followed are listed: discussion table among peers; construction of an in-depth questionnaire, following some indicators of training courses of the ValuE for Schools project.
Results. Some outcomes from the focus group. Many teachers of small schools, also belonging to the NIV, point out that the reflection shared between the central plexuses and peripheral plexuses is undoubtedly an opportunity for improvement. They require specific training on the following dimensions:

- Data literacy;
- Ability to better reflect and argue judgments;
- Ability to identify priorities and improvement paths;
- Ability to analyze data and trace connections between school weaknesses and action priorities.

In general, in the new 2022-2025 cycle, some multi-class or isolated school teachers were involved in drafting the RAV and also the PTOF. Possessing a good knowledge of local constraints and opportunities, they were delegated to construct perception questionnaires, aimed at families, local authorities or associations. The presence of many of them in preparatory commissions for the self-assessment has made it possible to refine the interpretation of the data and to focus the plots of improvement. According to some members of the NIV - non-teachers in small schools - the involvement of the latter in the reflection paths for change has made it possible to intercept unprecedented evaluation needs, leading to reflection on the experimentation of strategies to achieve improvement goals. The triggering of face-to-face or remote discussion groups led to the sharing of critical points and development milestones. Improved skills in active listening, collaboration, empathy, with a significant decrease in competitive dynamics. To consolidate the peer review process in the self-assessment. The moment of drafting the RAV was understood as the generation of a powerful link of cohesion between the various complexes. Only in some cases was the presence of a network for self-assessment found, which includes local schools and companies, libraries. Self-assessment is perceived as "getting involved, asking those who have more vision and experience" but at the same time it outlines a space in which multi-classes find recognition of a constant work of mediation, in overcoming unexpected events, risks, obstacles. Among the methods of training on self-evaluation, small school teachers report: collaboration between different grades; peer learning and visiting; mobile and variable groups with intersecting spaces; moments of co-planning with stakeholders (beyond the logic of accountability.).

***Keywords:*** small schools, self evaluation, educational networks, negotiation, improvement

# THEME 4. GENDER STEREOTYPES

**ORGANIZER: INVALSI**

**COORDINATOR: ELLEN CLAES**

**24TH NOVEMBER: 9.30 A.M. – 11.30 A.M. {ROOM 2 – RESEARCH SESSION 9}**

---

## Grade gender gap: evidence from RiminiInRete

**Beatrice Franzolini – Filippo Albertini – Lorella Camporesi**

Introduction. The *grade gender gap* refers to the disparity in academic performance between male and female students in different subject areas and educational levels. In recent decades, this phenomenon has drawn the attention of researchers, educators, and policymakers. An adequate quantification of the *grade gender gap* may highlight imbalances within our educational systems and potential consequences for individuals and society as a whole. This study aims to quantify and analyze the *grade gender gap* in primary and secondary schools (first and second level) in the province of Rimini. Scientific literature focused on measuring and understanding gender differences in school grades and standardized tests continually expands. We report here some key findings from the most recent literature to frame the results obtained in this research within broader trends of the *grade gender gap* in industrialized countries like Italy. One important result regarding the *grade gender gap* emerging from the literature is its variability across different historical periods and countries (see, for example, the PISA 2022 report). This suggests how the gap can be greatly influenced by the cultural, social, and/or economic context, thereby identifying interventions and policies aimed at reducing the grade gender gap. Some recent studies report how girls tend to have higher and less variable grades on average than boys, with a narrower advantage in STEM subjects (see, for example, O'Dea et al., 2018; Voyer & Voyer, 2014). However, gender gaps in favour of female students are not consistently observed in STEM subjects. For instance, the 2022 report of the Programme for International Student Assessment (PISA) by the OECD states that "*girls obtain slightly - albeit significantly - better results than boys in reading, while boys obtain slightly - albeit significantly - better results than girls in mathematics and science.*"

Research objectives. The objectives of this investigation are: (1) identify the *grade gender gap* in end-of-year average grades and across different subjects, (2) study the evolution of the *grade gender gap* throughout students' educational journey from primary school through first-level and second-level secondary school, (3) investigate the temporal trends of the *grade gender gap* by comparing results from different cohorts of students, and (4) quantify the effect of students' nationality on the *grade gender gap*.

Data used. The data used for this investigation are from the RiminiInRete (RIR) platform. RIR is a project that was born in 2015 from an idea by Lorella Camporesi and Roberta Badioli, dirigente and docente vicaria of the Istituto Comprensivo Centro Storico of Rimini, Italy. The project aims to create a database for collecting periodic and final evaluations of all schools in the province of Rimini. The primary goal of the project is to provide schools with the necessary data to develop longitudinal studies on student performance, allowing for the verification of the effectiveness of the evaluation system, especially during the transition between different levels of education, and the impact of evaluation methods on students' academic progress and school activities planning. To create the database, a network was established among all educational institutions, and a project group was formed, consisting of Filippo Albertini, Arianna Morri, Alessandra Fabbri, Elisabetta Mazzoli, and Emanuele Lucarini, coordinated by Lorella Camporesi. This group was responsible for creating the database, organizing data collection (including awareness-raising and technical training for teachers and administrative staff involved in the project), and disseminating the research results through collaborations with institutions and universities (University of Bologna, Rimini Campus, Cesena Campus) and researchers such as Camilla Giustiniani and Beatrice Franzolini, as well as individual schools within their own self-assessment activities. The RIR database includes data on student evaluations and attendance from primary to secondary schools in the province, starting from the 2012-13 school year, and it is currently also collecting guidance counselling data. The longitudinal nature of the RIR database allows for the analysis of the *grade gender gap* and its comparison among different student cohorts and stages of the educational journey. The sample for the current study is extracted from the population of students who received at least one end-of-year evaluation at a primary or lower secondary school in the Rimini province during the academic years from 2013-2014 to 2019-2020. The analysis

includes 2,927 students, of which 1,540 (52.6%) are male and 1,387 (47.4%) are female. Information is available for each student across multiple school years, up to a maximum of 7 years. Each student enters the study when they receive their first evaluation in a primary or lower secondary school in the Rimini province. The end of the follow-up period can be determined by one of the following events: the completion of the second cycle of studies, discontinuation of studies, enrollment in a school outside the Rimini province (censored data), or not completed second cycle of studies by the 2019-2020 school year (censored data). The average follow-up period per student is 3.94 years. Censored data, such as enrollment outside the province or non-completion of studies, are not expected to introduce biases in the results as they are not likely to be associated with a particular trend in student evaluations. The variables considered are (i) grades for Italian, English, Mathematics, Mathematics+Science, and Behavior, (ii) end-of-year average grade, (iii) gender as reported in official documents, (iv) nationality of students (foreign if born in a country other than Italy or if they had non-Italian citizenship at birth), (v) grade level, and (vi) cohort.
Method. The results were obtained through mixed-effects linear models, which allow for the simultaneous evaluation of the effect of the gender gap, grade level, and nationality of students, potential interactions between these variables, and a latent effect specific to each student.
Results. The average grade differences highlight an advantage for girls in all areas, with a more significant advantage in non-STEM areas (Italian and English). Specifically, the average grades are on average 0.34 higher for girls compared to boys, and the grade in Italian is on average 0.43 higher for girls compared to boys. The average differences in Mathematics and Mathematics+Science are 0.21 and 0.25, respectively, both in favour of girls. The value for English evaluations is 0.41. For Behavior, the differential between females and males is the highest at 0.62. All six differences are statistically significant (p-value < 0.05). Restricting the analysis to the top 5% of students (identified through the latent parameter of mixed-effects regression), boys achieve higher results in Mathematics and Science, while the advantage of girls becomes non-significant. Considering the bottom 5% of students, no statistically significant gender differences are observed. For the average grades and the areas of Mathematics and Mathematics+Science, there is no statistical evidence of a trend in the average gender differences during the school journey. However, a weak increasing trend is observed for Italian and English during lower secondary school. The study does not show differences in the gender gap among the different considered cohorts or historical trends of the phenomenon that cannot be attributed to natural variability among students. Finally, foreign students obtain lower grades compared to Italian students, with average differences of almost one grade in Italian. However, these differences do not vary based on the student's gender.

***Keywords:*** evaluations, gender gap, grades, inequalities

## A new scale of Perceived Self-Efficacy in Management and Use of Digital Devices: psychometric structure and gender invariance

**Rosalba Ceravolo – Marta Desimoni – Michela Milioni**

Introduction. In the current era, Governments and international institutions recognize as a strategic developmental priority in building active and inclusive citizenship to support and promote people's ability to adequately access digital devices for work, learning, and social life (e.g., EU's Digital Education Action Plan 2021-2027). The competent use of digital devices and internet tools, known as Information and Communication Technologies (ICT), represents a fundamental cluster of skills in contemporary society from the earlier phases of development till old age (OECD, 2016; 2020). In the EU context, the Digital Competence Framework for Citizens (DigComp) at European Commission states as digital competence"…the confident, critical and responsible use of, and engagement with, digital technologies for learning, at work, and for participation in society. It is defined as a combination of knowledge, skills, and attitudes." **(**Council Recommendation on Key Competences for Life-long Learning, 2018). Experts (Ayllón et al., 2020; Fraillon et al., 2019; Lorenceau et al., 2019; Rohatgi et al., 2016; van der Vlies, 2020; Xueliang & Jie, 2020) in the field started to investigate which aspects are more powerful in impacting the competent and goal-oriented use of ICT, which can be identified as a set of multidimensional skills that includes a heterogenous series of technical, non-technical, and knowledge issues and coping strategies. On the other hand, in the frame of the social-cognitive theory (Bandura, 1977), self-efficacy beliefs are recognized as the

more pervasive and influential mechanisms of human agency. Within these frameworks, an increasing number of studies have focused on gender differences in ICT self-efficacy believes (Hatlevik et al., 2018). Recently published meta-analysis on gender differences in ICT-related attitudes and motivation reported a small but significant positive effect towards boys ($g$ = +0.18), suggesting that boys have higher ICT self-efficacy than girls (Cai, Fan, & Du, 2017). Some studies suggested that gender differences in ICT self-efficacy depend on specific ICT activities. For example, some studies reported higher levels of self-efficacy regarding communicative Internet activities than for explorative Internet activities for girls (Tsai & Tsai, 2010; Vekiri, 2010). Additionally, in a German study, Lorenz et al. (2014) found higher self-efficacy regarding advanced ICT literacy for boys despite the lack of gender differences in self-efficacy regarding basal ICT literacy. Guided by the Social Cognitive Theory (SCT) framework (Bandura, 1997; 2006) and considering the relevance of high-level ICT competence, the heterogeneity of ICT skills, and the power of self-efficacy beliefs in influencing human agency and behavior and outcomes, we aimed at developing a reliable and valid scale to measure ICT self-efficacy beliefs in student attending the last year of upper secondary school. Further, based on literature on gender differences in ICT-related constructs, we aimed at testing gender invariance of our ICT self-efficacy scale.

Object and hypothesis. The present contribution presents a 19-item measure of ICT Self-Efficacy beliefs and assesses its psychometric properties. The *ICT Self-Efficacy scale* focuses on three areas: the students' judgments of their personal efficacy in managing the Internet content (e.g.: "How much can you do to understand if online contents are true, updated and complete?"); the students' judgments of their personal efficacy in handling basic competencies when using digital devices (e.g.: "How much can you do to create or modify a text doc?"); the students' judgments of their personal efficacy in handling advanced competencies when using digital devices (e.g.: "How much can you do to create or modify a webpage?"). The sample's characteristics, such as grade, gender, and school type, are considered in the study. We hypothesize that some ICT skills, even if universally relevant for students, could reveal a different functioning for males and females as well as for students attending different types of schools. To test our hypotheses, (a) the scale dimensionality (through exploratory and confirmatory factor analyses) and (*b*) the scale invariance across genders and school types are tested.

Data and method. A sample of 3511 students graded 13 has been recruited for the Computer Based (CB) administration of the Maths' grade 13 INVALSI pretest in the school year 2021-2022. The final sample of 3263 students, of which 1585 (48,6%) were girls, 1610 (49,3%) were boys, plus 68 (2,1%) students not reporting this information (missing). The youth attended different types of high schools. Specifically, 1470 (45,1%) students were from scientific lyceum and technological institutes; 500 (15,3%) were from vocational schools, and 1293 (39,6%) were from other types of lyceums. A questionnaire on ICT-related constructs, also including our scale on ICT *Self-Efficacy,* was administered to this sample. Students were asked to rate (from 1 =not at all to 5 =completely) their perceived ICT self-efficacy on the 19-item scale encompassing a wide range of activities they have to manage in using digital devices and accessing to the Internet. Scale's items asked in terms of "can do" and not "will do" or "be able to" (Bandura, 2006).

Results. Preliminary descriptives statistics reported an appropriate distribution for each of the 19 items for both calibration and validation subsamples. We first performed an Exploratory Factor Analysis (EFA) using the principal axis factoring (PAF) method and Promax rotation to test the factorial structure of our scale. Parallel analysis revealed a 3-factor model (explaining together 52.5% of the common variance) as the one to be preferred and retained for the subsequent Confirmatory Factor Analysis (CFA). The three dimensions resulting from EFA show high internal consistency (Cronbach's alphas >.80), and the correlations between the three subscales are all positive and statistically significant. Further, we tested and compared three alternative models through Confirmatory Factor Analysis (CFA): a one-factor model, hypothesizing all items loading on a single latent factor (M.1), a correlated three-factor model resulting from EFA (M.2); a second-order latent factor model (M.3). The best-fitting model included a second-order factor of general self-efficacy in the management and use of digital devices which is defined by three interrelated sub-dimensions: the first factor includes items that analyze the student's ability to manage and evaluate the online contents, called "Management"; the second factor includes items that evaluate the student's ability to have basic competencies in using digital devices, both for individual and group activities, called "Basic"; finally, the third factor includes items that evaluate the student's ability to have advanced competencies in using digital devices, called "Advanced". After establishing a well-fitting structure, we tested the measurement invariance across gender and high school type, following the guidelines suggested by Rudnev, Lytkina, Davidov, Schmidt, Zick (2018) to test measurement invariance for a second-order factor. For this analysis, the fit of the nested models was compared using Delta CFI (Cheung & Rensvold, 2002), which is referred as

more powerful in exploring the lack of invariance in large sample sizes (Meade, Johnson, & Braddy, 2008). Full metric invariance was empirically confirmed, thus supporting the invariance of factor loadings across gender and school type, for both the first- and second-order models. However, consistently with previous studies, results on scalar invariance suggested that some of the items were not invariant across gender.

***Keywords:*** information and communication technologies, self-efficacy beliefs, gender, EFA, CFA, measurement invariance

## Gender differences in computer user-profiles and performance on mathematics computer-based tests: an application of latent class analysis

**Marta Desimoni – Donatella Papa – Cristina Lasorsa**

Over the last decades, empirical research has been challenging the "myth of digital natives" (for a review, see Eynon, 2020), suggesting the need for a more complex and fine-grained picture of technology access, use, and skills in young people. Further, the extensive use of digital devices during the COVID-19 pandemic leads to a renewed interest in the pre-existing digital inequalities among students, as well as on the associations between differences in ICT use, and educational achievements. However, the gender digital divide community of knowledge is overwhelmingly focused on adulthood, whereas less is known about the digital realities for adolescents. Results from adults suggest that, compared to men, women have a lower frequency of use of ICT and a narrower range of online activities (for a review, see Robinson et al., 2015). Some studies suggested a similar pattern of lower access and use for girls, as for women (for a review, see Tyers-Chowdhury & Binder, 2020), and that girls have fewer positive attitudes toward technology use compared with boys, as boys are more likely to believe in the benefits of technology use in society (Cai, Fan & Du, 2017). However, other pieces of evidence suggest that the difference between female and male students in terms of frequency of use decreased over time (for a review, see Gebhardt et al., 2019). Recently, the ICILS 2018 international report (European Commission, Directorate-General for Education, Youth, Sport, and Culture, 2020) indicated that female students are stronger users of ICT for general school-related tasks, although this difference varies across countries. In response to the need for a more nuanced understanding of gender inequalities in early adolescents, the current study adopts a person-centered approach to explore the eventual differences in their patterns of computer usage outside of school between boys and girls, just a few months before the spread of the COVID-19 pandemic. Further, the study aims at investigating the associations between students' gender, the eventual "typologies" of computer users, and digitally assessed mathematics by integrating latent classes into auxiliary models. According to international large-scale assessment data, Italy is one of the countries in the OECD still displaying a relatively large gender gap in mathematics (Contini et al., 2017; OECD, 2021). Less is known about the interplay between gender, the use of digital technologies, and students' achievements.

Aims and hypotheses. This study aims at exploring heterogeneity in computer usage outside of school among early adolescents in Italy, about one year before the spread of the COVID-19 pandemic. In particular, the study pursues the following Research Questions (RQs):

RQ1. Do different computer-user profiles exist among eighth-grade students based on the time they spent on computers (desktop/laptop/tablet pc) and their activities with computers outside of school?

RQ2. Does students' gender predict the likelihood to belong to a given computer user subgroup?

RQ3. Does computer-user subgroup membership predict mathematics ability, beyond the (direct) paths from students' gender to mathematics?

Method. Participants were 1155 students (49.5% female and 50.3% male; gender missing data= 2; 12.6 % of students with an immigrant background) attending Grade 8 in Italy (North-West = 19.3%; North-East = 18.5%; Center = 30.2%; South = 31.9%; n of classes = 72). All students participated in the 2019 Field Trial (FT) of the INVALSI National Assessment Study, carried out at the end of the school year. Mathematics tests were administered to all students of a representative sample of classes (two randomly selected classes within each randomly selected school). Students were assessed at school with an external observer. Five mathematics computer-based test forms were randomly assigned to the students, along with a short questionnaire on computer (desktop/laptop/tablet pc) usage. All mathematics tests were administered online, as well as the additional questionnaire. All data were collected and kept anonymously.

Results. Latent class analysis (LCA) with mixed-metric indicators was applied to explore the presence, and nature of, subpopulations of computer users. The indicators of the latent class analysis were items from the short questionnaire on computer usage outside of school. Relative fit indices (Information Criterion, IC), entropy, and the interpretability of the proposed optimal solution were evaluated to determine the best solution. Vuong-Lo-Mendell-Rubin and Lo-Mendell-Rubin adjusted LRTs, which quantify specific comparisons between a given model and a model with one fewer class, were also performed. A three-step VAM approach (Asparow & Muthén, 2014; McLarnon & O'Neall, 2018) was then used to test auxiliary models on the associations between the categorical latent class variable, students' gender (and other socio-demographic characteristics), and performance on mathematics computer-based tests. Results from latent class analyses (LCA) on data collected in Italy, just one year before the outbreak of the COVID-19 pandemic, are consistent with the hypothesis that grade 8 students were not homogeneous in their computer use outside of school. Students' gender predicts students' class membership which, in turn, predicts digitally assessed mathematics, even beyond the direct paths from gender and other sociodemographic variables to the outcome.

***Keywords:*** gender, latent class analysis, computer-based assessment

## Get off with a good start: predicting power of INVALSI mathematical test on STEM career at university according to gender

**Patrizia Falzetti – Patrizia Giannantoni**

A major interest in the most recent years has been devoted to STEM disciplines, for the increasing job demand and higher wages associated to this field of studies. In this perspective particular attention has been directed to gender differences into STEM studies, which if often cause of the widening gender gap in work career and life opportunities.
The data about Italian situation show a different composition by gender at University: among STEM graduates the male component is higher, reaching 59%, of students while among non-STEM graduates women prevail (they are almost two out of three) (source: Almalaurea).
In a previous work presented at INVALSI Seminar 2022 the authors investigated the impact of performance at INVALSI test in Math for students in the last year of upper secondary school (so called grade 13) on the propensity to enrol in a STEM course at University the following year. Findings showed that there was an important influence of cultural status of the family, measured by educational attainment of the mothers, and an influence both of teachers marks and performance at INVALSI test in math, with a slightly higher impact of the former, on the probability to choose a highly scientific course of study at university.
However, the academic career and eventually the obtaining of the final degree are a much relevant outcome to consider in terms of success into a STEM path for students. Therefore, with this contribution we decided to make a step further and to analyse not just the enrolment, but the university career as predicted by INVALSI test, always in a gendered perspective.
OBJECTIVES. The aim of this contribution is to deepen the study of factors predicting a STEM career, by looking more to the university success in terms of exams and credits gained, rather than merely to enrolment in different university courses. Some questions of interest are: do the mathematical skills achieved in high school count in a different way for a highly scientific career of boys and girls? With the same mathematical skills, how important are context factors, such as the socio-economic background of the family of origin, and / or the evaluation by the teachers, the country of origin in proceeding into a STEM path at university? Once inserted in the STEM University courses, do the girls have an equal performance with their male colleagues or even higher?
DATA. We used as a base the original dataset already created ad hoc for the previous step of this research project, i.e. a dataset built on the combination of different data sources: MIM (Ministry of Education and Merit), University Register of students and INVALSI (National Institute for the Evaluation of the Education and Training System) data. This combined data-source will allow us to follow different cohorts of students in a longitudinal perspective that covers different school grades (INVALSI data), with an enrichment of information from Ministry data, and is eventually linked to university enrolments and to the number of educational credits obtained for each year of university course.

It will therefore be possible to have data about students' performance, in all subjects (Italian, mathematics and English) and throughout the whole school career, as well as to have information about the student's family background (e.g. parents' educational level, socio-economic status of the family, country of origin, language spoken at home).
Furthermore, data will be available at the University level and with a very specific course classification, which can be recode into a binary variable (STEM / non-STEM).
METHODS. Analyses about the development of career of students into scientific tracks will be carried on by looking at similarities and differences between boys and girls in the speed and the average mark they obtain in scientific courses.
More in depth studies would be related to regression analyses and survival analysis.
On the one hand, regression models allow the estimation of the weight of the various factors in determining the progression into the university career, as measured in terms of credits gained.
Having a wide set of socio-demographic and contextual characteristics it would be possible to estimate the impact of the different characteristics on the speed and success in progression through the university career for boys and girls, keeping as a key variable that of "mathematics skills" measured as a continuous variable (WLE score) or as a categorical variable based on the WLE score levels (ordinal scale from 1 to 5) during the INVALSI test of level 13. On the other hand, survival analysis, could improve the understanding of STEM promoting factors, highlighting characteristics of those students who survive compared to those dropping out. Of a major interest would e at this point, to observe whether the scientific career for boys and girls has significant differences
EXPECTED RESULTS. After giving a first descriptive analysis about gender differences in academic performance and university choices, we aim with this contribution to analyse more in depth the characteristics of students having more success into a STEM discipline at University, keeping constant the interest on gender in the analyses. We intend to use regression models, to study the impact of different factors, both individual and contextual, on the university success into STEM discipline. Particularly, we intend to estimate the weight of the mathematical skills, which we assume have an expected and legitimate influence on the choice of academic orientation, compared to the weight of the "social" factors that instead lower the propensity of girls towards the more scientific-technological university paths.
Previous results showed that math skills are a strong predictor of choosing STEM degree (students with very high mathematics performance, i.e., Level 5, during the last year of high school are 5 times more likely to enrol in a STEM-university degree compared to students who have only reached the lowest level at INVELSI test, Level 1.
It is therefore essential to deepen how, on a personal level, all these contextual factors (family and school as the main agents) and personal factors (both academic and non-academic skills) shape the probability of success in university career for students who have chosen a STEM degree, and particularly for girls.

***Keywords:*** STEM, gender gap, math, predictive factors

## Exploring the Influence of Maternal and Teacher Role Models on Gendered Differences in Mathematics

**Alice Bertoletti – Marta Cannistrà – Mara Soncin**

Gender disparities favouring boys in mathematical performance continue to be significant in Italy, with the country exhibiting one of the largest gender gaps in Europe (Contini et al., 2017). The underachievement of girls in mathematics has raised concerns as it appears to influence their career choices, with female students being less likely to opt for STEM (Science, Technology, Engineering, and Mathematics) field careers (Cheryan et al., 2017; Stoet & Geary, 2018). Therefore, gender inequality in mathematics remains an unresolved issue encompassing limited female representation in certain fields of labour market, salary disparities, and the persistence of gender stereotypes in career choices (Weiner et al., 2010).
Evidence from INVALSI (National Evaluation Committee for Education) data indicates that these gender differences in mathematics achievement become more pronounced as Italian students progress through the educational system. Notably, a gender gap in math becomes evident from grade 5, where boys achieve an

average score of 199, compared to girls who score an average value of 186 (see INVALSI, 2021). This gender gap further amplifies to over 6 points by grade 8 (see INVALSI, 2021).
The present paper addresses this issue by focusing specifically on the influence of role models in explaining and mitigating the gender gap in mathematical achievement. Indeed, existing literature highlights the relevance of role models in inspiring students to overcome gender stereotypes and pursue careers in STEM fields. Specifically, two primary types of role models are commonly discussed: parental role models and teacher role models (De Gendre et al., 2023). For instance, extant research suggests that a mother's profession positively impacts the academic performance of university students, particularly females (Hussain, 2022). Also, children of educated mothers tend to exhibit better academic performance, compared with those whose mothers have lower levels of education (Awan, 2015). Another crucial role model explored in the literature is the effect of same-sex teachers. A recent meta-analysis by De Gendre et al. (2023), covering 90 countries worldwide, demonstrates a small but statistically significant and positive effect of same-sex teachers on student performance (0.03 SD). Interestingly, the effects are more substantial in countries with wider gender gaps.
Thus, we aim at investigating the influence of both mother and teacher role models on student performance, with a particular focus on the standardised math test scores of female students in Italy. The paper seeks to shed light on a possible mechanism driving the gender gap in math performance throughout the educational careers of students.
More specifically, the study addressed the following research questions:

- How do mother and teacher models influence the standardised math test scores of female students in Italy?
- How do these effects vary across different grades and regions?

The study employs a longitudinal analysis of INVALSI data, covering the period from the school year 2012/13 until 2018/19. We exclude data from 2020/21 onward to avoid of confounding the role-model effect with the one due to COVID-19 pandemic.
We examine the data of students in grades 2, 5, 8 and 10. This allows us to track the same cohort of pupils throughout their educational pathway, from an age around 7 years old until around 13 years old. The extended time frame allows us to gain valuable insights into the changing impact of role models on academic performance across different educational stages. The existent evidence suggests that the influence of mothers as role models should diminish as students progress in their educational journey. Conversely, the relevance of teachers as role models may exhibit an opposite trend, becoming increasingly relevant as students attain higher levels of education. The longitudinal dimension of the data represents an important novelty of this study. This diachronic perspective is normally not feasible in studies available in the literature that often rely on international large-scale assessments focused on specific-age cohorts of students (such as PISA, TIMSS and PIRLS)
As we aim at exploring the effect of female role model on mathematical performance of girls, we compare the educational achievements of female students (the group of interest) with the ones obtained by their male peers (the control group). Then, to enhance the comparability between these two groups, we apply propensity score weights based on key student characteristics.
The analysis adopts a multi-level framework that takes into account the hierarchical structure of the data, with students nested within schools and schools nested within regions. This analytical strategy allows for the examination of regional heterogeneity in the effects under investigation, which is particularly important given the significant differences that exist between Italian geographical areas. Several studies have, indeed, shown that students in the north perform significantly better than those enrolled in southern schools (see, for instance, INVALSI, 2021; and Bratti et al., 2007).
Drawing on the INVALSI data, the influence of maternal role models will be assessed by considering both the mother's occupation and education level. Additionally, the role of teacher models will be captured by the gender of the teachers and their educational qualifications. Examining the joint impact of teacher and mother role models represents an original contribution to the existing body of research.
The findings offer valuable insights for addressing the gender gap in mathematics and promoting equality in educational and occupational opportunities. The research highlights the potential of role models as a potent policy tool for reducing educational inequality, underscoring the need for targeted interventions to encourage greater engagement of female students in STEM fields.

***Keywords:*** role model, gender gap, mathematics, standardized test scores

# THEME 3. INTERNATIONAL LARGE-SCALE ASSESSMENTS (ILSAS)
## METHODS AND RESULTS
### ORGANIZER: INVALSI
### COORDINATOR: ERNESTO TREVIÑO
### 24TH NOVEMBER: 9.30 A.M. – 11.30 A.M. {ROOM 3 – RESEARCH SESSION 10}

---

## When almost all items are endorsed: Extreme responses or substantive classes?

**David Torres Irribarra – Diego Carrasco**

Introduction: Citizenship norms present in the International Civic and Citizenship Education Study (ICCS) are injunctive solcial norms (Torres Irribarra & Carrasco, 2021; Cialdini, Kallgren, & Reno, 1991). This includes norms such as voting, obeying the law, working hard, engage in political discussion among others. Injunctive norms describe what people ought to do in contrast to descriptive norms (what people tend to do) (Cialdini & Goldstein, 2004). Citizenship norms endorsement are relevant because these predict the likelihood of students to vote, participate in protests and obey the law (Gerber & Rogers, 2009; Köbis, Van Prooijen, Righetti, & Van Lange, 2015; Rees & Bamberg, 2014; Wenzel, 2005).

Previous studies in citizenship norms endorsement (Hooghe & Oser, 2015; Hooghe et al., 2016) proposed the use of latent class models to represent adherence to citizenship norms. Citizenship norms endorsement in these response models are represented with different response patterns. Previous findings have found diverse configurations of endorsement of citizenship norms (Hooghe, Oser, & Marien, 2016; Reichert, 2017), and latent groups that resemble Dalton's distinction (Dalton, 2008) between duty-based and engaged citizenship.

Based on the work of Hooghe and colleagues (Hooghe & Oser, 2015; Oser & Hooghe, 2013), and the literature of latent class models for multiple groups (Kankaraš & Vermunt, 2015; Kankaraš, Vermunt, & Moors, 2011; Masyn, 2017) we discuss what latent class model specification is better suited for a comparative approach. We identified two limitations of the previous research. The first, consists of the use of heterogenous latent class models, which we assert do not provide comparable latent groups between countries (see Torres-Irribarra & Carrasco, 2021). The second limitation of previous approaches are its limited capacity to distinguish extreme norms endorsers from extreme response styles (e.g., Hooghe, Oser, Marien, p123). This second limitation is of importance, as it constraints the substantive interpretation of the unobserved groups that present the higher variation between 1999, 2009 and 2016 of ICCS rounds (Oser et al., 2022).

The present study offers a new approach, consistent with the use of latent class models, to differentiate substantive extreme endorsers, from extreme response styles.

Object and research hypothesis: In the present study our aim is to compare different responde models fitted onto citizenship norms endorsement items, to discuss their trade offs. In particular we discuss the interpretation of two latent class model specifications, the structurally homogeneous model and the partially-homogeneous model, as tools for comparative research.

Data: Data. We used students' responses from the International Civic and Citizenship Education Study (ICCS) 2016. This study obtained responses from a representative samples of grade 8 students (Schulz, Carstens, Losito, & Fraillon, 2018). In 2016, 24 countries participated in the study from Europe (Belgium (Flemish), Bulgaria, Croatia, Denmark, Estonia, Finland, Italy, Latvia, Lithuania, Malta, Netherlands, Norway, Russian Federation, Slovenia, Sweden, North Rhine-Westphalia (Germany)), Latin America (Chile, Colombia, the Dominican Republic, Mexico and Peru), and Asia (Chinese Taipei, Hong Kong SAR, and the Republic of Korea).

Variables. Students responded there adherence to different citizenship norms, using a four-point Likert-type scale, with the response options of "Very important", "Quite important", "Not very important", and "Not important at all". Examples of these indicators, are "Always obeying the law", "Voting in every national election", "Working hard", "Joining a political party" and "Engaging in political discussions". In the present study we used 12 of these indicators, to ensure comparability with previous studies (Hooghe & Oser, 2015; Hooghe et al., 2016)

Methods: In the present study we fit a series of structurally homogeneous model to analyze profiles of good citizenship across countries in order to prioritize the interpretability of international comparisons

(Kankaraš & Vermunt, 2015). This models fixed the same response pattern across countries — thus ensuring comparable interpretations across them —, while the prevalence of each class is free to vary between countries.
We used a model based approach, to try to separate respondentsthat would be more likely to be suspected as extreme responders from respondents who simply endorse all or most of all citizenship norms. To do so, we specify two extreme latent classes to capture the extreme responders, while competing to our previosly fitted model with five latent classes. The fitted mdoels are not exploratory latent clases, but comfirmatory latent class models.
All estimates were produced using Latent Gold 4.5 software (Vermunt & Magidson, 2013), accounting for the study survey design, and use Taylor Series Linearization for standard error estimation (Asparouhov & Muthén, 2010; Stapleton, 2013).
Results: Our results allows to show that our originally fitted model with five latent classes (Torres-Irribarra & Carrasco, 2021) miss classified students between 0.06% to 3.01% under the presented method, well bellow a nominal measurement error (5%).
We compare the present approach to other available indexes in the literature develop to identify extreme response styles and careless responses. We believe that the model based approach using confirmatory latent class models is a viable tool to gather evidence regarding the presence of extreme response patterns, while allowing the substantive interpretation of extreme endorsers in citizenship norms endorsement. We hope the present study offers an additional useful tool for the study of highly skewed response patterns, present in non-cognitive measures.

***Keywords:*** extreme response styles, social desirability, mixture models

## Implications of ILSA Results for Citizenship Curriculum Reforms

**Citlalli Sanchez-Alvarez**

The findings of the 2016 International Civic and Citizenship Education Study (ICCS) have played a pivotal role in informing and strengthening the latest curriculum reform proposal in Mexico. These results have presented a valuable opportunity for policymakers to incorporate key insights related to equity, quality, inclusion, and a humanistic approach, which are foundational aspects of the current educational model. This proposal aims to contribute to the discourse on citizenship curriculum reforms by examining the impact of the 2016 ICCS findings on the current educational model and the most recent curriculum reform in Mexico. Furthermore, the implications of this work seek to stimulate a discussion on how policymakers in other countries, such as Italy, can leverage ILSA-derived data to further develop and enhance their own curriculum reforms.
This proposal seeks to analyze the implications of ILSA results, specifically focusing on the 2016 ICCS cycle, for citizenship curriculum reforms. By examining the influence of these findings on the current educational model and curriculum reform in Mexico, the objective is to provide insights into the benefits and challenges of utilizing ILSA-derived data to inform policy and curriculum development.
The analysis centers on the results obtained from the 2016 ICCS cycle, which assessed the civic knowledge, attitudes, and engagement of second-grade students in secondary education in Mexico. The national sample comprised 5,526 students and 1,918 teachers from 213 schools, representing a diverse range of educational contexts. The study collected data on various aspects of civic education, including students' perceptions of the state and its institutions, peaceful coexistence, social cohesion, diversity, and attitudes toward civic values.
Utilizing results derived from large-scale assessments poses a challenge in educational assessment. However, these assessments provide valid and reliable information that enables evidence-based decision-making. This presentation will reflect on how the results from the 2016 ICCS cycle played a significant role in shaping the current educational plan and its study programs, taking into account the educational goals and focus on equity, inclusion, quality, and a humanistic approach set forth by the current government.
Among multiple findings, one example of what the study revealed is that 47% of students considered it neither good nor bad for one single company or government to own all the newspapers in a country. This finding suggests a limited perception of the importance of freedom of expression in the media, pluralism,

and the right to access information, which are integral components of democratic societies (INEE, 2018). Contextualizing this information within the framework of curriculum reform calls for a comprehensive review, encompassing aspects such as curriculum content, school context, and teaching characteristics. Strengthening issues related to the importance of freedom of expression within democracies become crucial in such scenarios.

During the evaluation process of the curriculum design for citizenship education, a notable finding emerged: most issues related to citizenship education were limited to the subject of Civic and Ethics Education, despite the educational model stating that civic and ethics core principles should permeate all subjects (INEE, 2019). These findings, in conjunction with the ICCS results, shed light on a significant curricular inconsistency and provided an opportunity to reinforce specific points that impact civic and citizenship education.

Citizenship education in Mexico is a dynamic and evolving phenomenon, which can be approached from two perspectives: education received in schools and knowledge acquired from the surrounding context. The most recent change occurred within the framework of the 2017 Educational Model for Compulsory Education, which places citizens' education as a central goal of compulsory education. However, as mentioned earlier, a new revision and its implementation are currently underway under the principles of the current administration (2019-2024).

Mexico's participation in ICCS began in 2009 when the Secretariat of Public Education (SEP) embarked on efforts to join an international study investigating how countries prepare students to become informed and engaged citizens. In the 2016 cycle, Mexico participated for the second time, with a national sample comprising 5,526 students and 1,918 teachers from 213 schools. ICCS aims to assess the preparedness of young people to assume their roles as citizens in the 21st century, addressing the challenges of democracy and civic participation within evolving contexts.

The reference frameworks of the ICCS 2009 and 2016 studies have been widely utilized, surpassing the use of data for research purposes. However, there has been a lack of systematic efforts in this regard. The curriculum development teams at SEP have primarily relied on the categories and concepts included in the reference framework, rather than the study results. This can be attributed to the limited number of national or regional studies on this topic, impeding researchers and educators from having comprehensive national references on the performance and achievement of students in civic and citizenship education (Luna, Meléndez, & Sánchez, 2019). The challenge lies in finding diverse approaches to utilize ICCS results as a robust evidence-based source of information that can strengthen citizenship development and support new government policies promoting equity, inclusion, and quality in education.

The current six-year term government (2019-2024) aims to materialize its educational policy through what is known as the New Mexican School (NEM) project, placing girls, boys, and young people at the forefront of education. This policy prioritizes civic, ethical, and socio-emotional education, as well as critical thinking while fostering recognition and respect for nature, cultural diversity, and human dignity. It is imperative for NEM to address the diverse realities of Mexico by delivering quality education with equity and inclusion.

Some of the results derived from national reports are directly linked to the civic education curriculum and the overall educational program. There is an opportunity to reconsider citizenship education within the curriculum at this political juncture. Analyzing information based on the perspectives of teachers, subject matter experts, curriculum designers, and government policies holds great importance for curricular analysis and reorientation. The performance results of students in studies such as ICCS serve as valuable evidence and input that should be considered in order to strengthen citizenship education across the curriculum, thereby enhancing overall educational quality.

Analyzing the implications of this work in different countries, including Italy, is of paramount importance. This presentation aims to spark a discussion on how policymakers can effectively utilize ILSA-derived data to develop and enhance curriculum reforms. By sharing experiences and insights from Mexico, we hope to inspire policymakers and researchers in other countries to leverage ILSA data in their own educational contexts, leading to more effective citizenship curriculum reforms that foster engaged and responsible global citizens.

***Keywords:*** civic and citizenship education, curriculum reforms, curriculum development, educational policy, large-scale assessments

# How motivational constructs interact to predict elementary students' reading performance

**Beti Lameva – Zaneta Chonteva**

Description. Reading is one of the most important abilities students acquire as they progress through their school years (Kintsch, 2013; Mullis et al., 2017). More than any other skill, reading proficiency is important to effectively navigating the school curriculum, shaping each individual's trajectory through life, and actively taking part in broader society (Martin et al., 2011). PIRLS (The Progress in International Reading Literacy Study) was first conducted in 2001 as a follow up of the IEA's 1991 (International Association for the Evaluation of Educational Achievement) Reading Literacy Study and is one of the core studies of IEA that provides internationally comparative data on how well children read; and offers policy-relevant information for improving learning and teaching. Conducted once in every five years, PIRLS evaluates the reading comprehension development of fourth graders (Mullis & Martin, 2019).

Effective teaching of reading literacy includes understanding the importance of positive attitudes toward learning and fostering their development among students. Many studies have shown that students' motivation to learn is related to higher achievement, but when making decisions to improve learning and practice, it is important to recognize that cultural influences and gender may also play a role, and establishing links between achievement and motivation are thus especially complex. The context of the reading situation can support the construction of meaning by promoting engagement and motivation to read, but the context also can place specific demands that might not support the construction of meaning (Christianson & Luke, 2011; Lorch et al., 2011; Miller & Faircloth, 2009; Taboada et al., 2009). This research focuses on motivational predictors of children's reading performance using the data from PIRLS 2021 and explores an alternative perspective on the complex association between motivation and reading performance. In particular, drawing on the multidimensional and multifaceted characteristics of domain-specific motivation, this study investigates the interactions among different facets under four motivational dimensions in relation to reading performance of North Macedonian students. Numerous research studies confirm the role of students' motivation as significant predictor of their achievements in tasks for which they have to read and understand the text, that is, to discover, select, interpret, integrate and evaluate different information in the process of solving (Andersen & Nielsen, 2016; Good & Lavigne, 2017). The model for measuring motivation is theoretically elaborated and defined according to social cognitive theories of achievement motivation (Self-determination theory), and it is determined by four dimensions (factors): Competence, Autonomy, Relatedness and Value of reading. The four factors are hypothesized to covary and interrelate into a higher order factor named Motivation for reading (Ryan & Deci, 2017; Deci & Ryan, 2000).

Object and research hypothesis. The research aims to define the model and the role of motivation to read in the international assessment program of students' reading achievement (PIRLS 2021) in their fourth year of schooling. The sample consists of 2929 primary students from 150 primary schools in North Macedonia. In the research framework for conducting the PIRLS 2021 testing, four scales were introduced to measure the construct Motivation for reading, that is why the research is divided into two parts. The first part focuses on testing the hypothesized validity of the four constructs related to motivation for reading, while the second part of the research aims to determine whether these constructs predict the achievement of tasks in verbal content comprehension tests among 4$^{th}$ grade students in the Republic of North Macedonia, covered by the testing within the international research study PIRLS 2021. The average score for the North Macedonian students in reading is 442 and it is below the international average, so the findings and recommendations of this research would help educational policy makers to design programs aimed at emphasizing the importance of students' motivation in order to achieve higher academic results.

Method(s) employed. The research used scales that refer to the constructs related to motivation for reading. They are an integral part of the questionnaire intended for students, as part of the research procedure prescribed by the protocol of the international research study PIRLS 2021 (Mullis & Martin, 2019). Four scales from the student's questionnaire applied in PIRLS 2021 were used to assess students' motivation to read. The scales assess the following dimensions: Competence (Students Confident in Reading scale), Autonomy (Students Engaged in Reading Lessons scale), Relatedness (Students' Sense of School Belonging scale) and Value of reading (Students Like Reading scale). To determine the hypothetical (construct) validity of the proposed factor model (first order and hierarchical models), confirmatory factor analysis was used in the statistical package SPSS-Amos (Statistical Package for Social Sciences-Analysis of a Moment

Structures). Linear regression was used in order to determine which motivational factors are significant predictors of the students' reading achievement in the PIRLS 2021 testing by using the IEA International Database (IDB) Analyzer software.
Results achieved. Results from the confirmatory factor analysis show that the hypothesized 4-factor model has an acceptable fit to the data from this research and it can be considered when determining the optimal approach to designing reading lessons. In the final hierarchical model, the following values were obtained for the selected model suitability criteria: CMIN/DF = 2.96; TLI = .91; CFI = .92; SRMR = .04; RMSEA = .04 with 90% CI [0.042, 0.049]. Results from the linear regression analyses showed that all four motivational factors are significant predictors of students' achievement of the tasks in the tests of understanding verbal content. Total percentage of explained variance is 18% ($F(4, 1978) = 106.530$, $p < .01$).

***Keywords:*** PIRLS, student achievement, motivation to read, self-determination theory

## Preparing Teachers for Assessment: Evidence from Three International Large-scale Assessment Programs

**Serafina Pastore – Davide Azzolini – Sonia Marzadro**

Introduction. Schooling policy, over the last two decades, has been run through with neoliberal ideology (Lewis & Lingard, 2015; Smith, 2016): the widespread of the new public management paradigm, the liberalisation and privatisation of education in many countries have permeated different aspects of schooling (from parental involvement to teacher and student conceptions and practices). The launch of reforms in teacher education (e.g., in the US, UK, Australia, Singapore, Hong Kong, or within the EU context, in Germany and Finland) to improve teaching and student learning nationwide raised the standards for education quality (Fullan et al., 2018; Hardgreaves, 2020; Zhao, 2016). More specifically, teaching quality has been recognised as the most important policy lever for educating the future citizens for the knowledge society. Assuming the importance of learning achievement as the main goal of most education reformers, different countries embraced the use of large-scale assessment (LSA) programmes, sometimes very different in terms of principles and methodologies (Black & Wiliam, 2007), in order to improve national curriculum and educational quality (Verger et al., 2019; Yue-Yi, 2021). The interest reserved for student learning assessment, as well as the growing emphasis on standardised tests and high-stakes assessment, have led national school systems to diversely frame assessment in order to reply both to accountability requirements and educational improvement (e.g., Australia, Canada, South Korea, Hong Kong) (Smith, 2016). In this perspective, a great amount of attention has been globally directed to school achievement, as well as to teacher assessment practice, use of assessment data and teacher preparation for sound assessment (i.e., assessment literacy).
Given the compelling evidence that assessment represents a key leverage point for improving student outcomes (Andrade & Heritage, 2018; De Simone, 2020; Klinger & Rogers, 2011) improving teacher assessment literacy represents one of the most pressing and contested contemporary educational policy and practice issues. The efforts made by policymakers, teacher educators, and administrators clearly show that the intersection of teacher education and assessment practice matters in fostering the quality of instruction in national school systems. Assessment literacy is an important component of teacher professionalism. However, the increased importance accorded to assessment is not always linked to effective professional development paths that can encourage a systemic review of teacher assessment practices and conceptions (DeLuca & Volante, 2016; Klinger & Rogers, 2011; Will et al., 2019). Teachers struggle to transfer what they have learnt in professional development paths in the real context of the classroom. Moreover, the unprecedented times of the Covid-19 pandemic have made more evident how teachers, despite the broad literature on educational assessment, struggle to navigate old and new instructional circumstances in their assessment practice.
**Research aims**. On the backdrop of comparative educational research on teacher professional development and teacher assessment literacy, the following questions frame the present study:
- Which variations, across grades and countries, can be identified in (i) teachers' training on assessment; (ii) assessment strategies; and (iii) assessment practices?
- Are today's teachers more frequently trained on student assessment?

- Do they use a wider spectrum of assessment practices than teachers in the past?

The paper provides an extensive cross-national comparison, in the European context, on the evolution of these aspects over about 10 years.

Research method and data analysis. To this end, the study exploits data from three TALIS waves from 2008 to 2018, as well as three PIRLS and TIMSS teacher questionnaire data spanning from 2006 to 2019 (Tab. 1). The three surveys (i.e., TALIS, TIMSS, and PIRLS) allow comparing three main indicators related to teacher training, conception and practice of student assessment, which were collected with very similar and hence comparable questions (Tab. 2). More specifically:

1. The indicator on training comes from a question asking to what extent teachers study assessment methods as part of their formal education and/or training;
2. The indicator on assessment strategies measures how much importance do teachers place on the different sources to monitor students such as short regular written test or assessment of ongoing work;
3. The indicator on practices measures how often teachers use different methods of assessing students' learning.

Results. Given the limited availability of comparative information on teacher training and practices in assessment, this study offers an ideal opportunity to investigate how research education and accountability demands interact and the consequences of that interaction for school systems, teacher education, and, ultimately, teacher practices. Through this cross-national and trend analysis, this study intends to contribute to the understanding of current global dynamics that affect teacher assessment practices and identify future policy directions and research agenda on teacher education and educational assessment.

Based on the comparative research on teacher training and practices in the educational assessment domain, this study findings represent a useful step forward in mapping out the adjustments, in theory and practice, necessary to support teacher assessment literacy.

***Keywords:*** international large-scale assessment, teacher assessment literacy, professional development, comparative educational research

# THEME 5. EDUCATION AND SOCIAL MOBILITY

**ORGANIZER: INVALSI**

**COORDINATOR: LORENZO MARAVIGLIA**

**24TH NOVEMBER: 2.00 P.M. – 4.00 P.M. {ROOM 1 – RESEARCH SESSION 11}**

---

## The Impact of Primary School Math Skills on Academic Achievement in Lower Secondary School

**Chiara Masci – Mara Soncin – Tommaso Agasisti – Melisa Diaz Lema**

The importance of early learning and academic achievement in primary school for long-term academic success has been widely acknowledged in literature (Morgan et al., 2016; Pappas et al., 2019). Indeed, some studies indicate that mathematical achievement during the early years of primary school serves as a reliable predictor of students' subsequent performance (Garon-Carrier et al., 2018) and academic challenges that emerge early tend to persist throughout the school path (Watts et al., 2014). National assessments in Italy reveal that a considerable percentage of students face academic challenges in mathematics. According to the Italian National Evaluation Committee for Education (INVALSI), proficiency in mathematics was observed in only 56% of 8th graders in A.Y. 2022. This study aims to uncover the most relevant math skills in primary school, as extracted from standardized tests in grades 2 and 5, that significantly impact proficiency at the end of lower secondary school.

The analysis utilizes data from three different national cohorts of primary school students who were tested in grade 2 (school years: 12-13, 14-15, and 15-16) and grade 5 (school years: 15-16, 17-18, and 18-19). It matches the students' patterns in four prevalent areas of math literacy (relationships and functions, data and predictions, numbers, and space and figures) with the likelihood of being a student at-risk at grade 8 (school years: 18-19, 20-21, and 21-22). In this respect, at-risk students are defined as those that do not reach the minimum proficiency in the test (i.e., whose competence level is equal to or less than 2) and thus are at-risk of educational poverty and implicit/explicit dropout.

To explore students' patterns on math literacy during primary school, the study relies on a two-step latent class analysis, a mixture modelling technique that allows to group students into latent classes on the basis of their observed patterns of responses in the mathematics test. While for the prediction of at-risk students at the end of lower secondary schools, the study compares multilevel statistical and machine learning models. In this second part, models are evaluated in terms of goodness of fit, predictive power and interpretability.

The study makes a twofold contribution. Firstly, it provides a national-level modeling of pupils' math skills in primary school over time, offering a comprehensive understanding of the specific math literacy skills that significantly influence academic achievement in lower secondary school. Secondly, the findings enable the identification of potential areas of improvement at the primary school level, allowing educators and policymakers to develop targeted interventions that enhance specific math skills and promote long-term academic success for students.

***Keywords:*** math skills, primary school, latent class analysis, at-risk students

## Geographical cleavages in educational outcomes: Educational achievement and social inequalities in Inner and Central areas in Italy

**Andrea Pietrolucci – Moris Triventi – Nazareno Panichella – Stefano Cantalini**

Several empirical contributions documented the relevance of the North-South geographical cleavage in students' educational outcomes as a key characteristic of the Italian educational system (Bratti, Checchi, and Filippin, 2007; Checchi and Braga, 2010; Argentin et al., 2017; Falzetti and Sacco, 2021). In this literature, however, little attention has been paid to the analysis of the role played by the divide between Central and Inner areas in the stratification of students' educational outcomes. Indeed, municipalities located in Inner and Central areas show substantial differences on a set of contextual characteristics which

are likely to influence educational outcomes, including public and schooling services, public transport, availability of digital services and infrastructures, and access to cultural heritage (Lucatelli, 2015). Analysing gaps in educational outcomes between Central and Inner areas would allow us to get a deeper understanding of territorial inequalities in educational opportunities in Italy and to inform policy interventions aimed at reducing these inequalities. Additionally, we are interested in whether this geographical dimension contributes to moderating socioeconomic inequalities rooted in the family: is the relationship between students' social background and achievement heterogeneous across central and marginal areas?

The main aim of this paper is therefore to examine whether students attending school in Inner areas show different educational performances compared to students attending school in Central areas while also assessing whether the Inner/Central areas gap in the educational performance varies across school grades and geographical areas. Furthermore, we are interested in evaluating whether geographical mobility from Inner to Central areas in upper secondary school is likely to influence schooling performances. To this purpose, we use data from INVALSI-SNV on the student population in the 5th, 8th, 10th, and 13th grades for school years between 2011/2012 and 2021/2022. As main dependent variables we rely on educational performance as measured by standardized tests scores in Italian-reading and Mathematics. Scores are standardised by year and grade and are thus interpreted in a relative way, as one student's position in the cohort-year specific achievement distribution. As main independent variable we define the location of schools at municipality level (*Comune*) by distinguishing between schools located in Inner areas and in Central areas by using the classification provided by *Strategia Nazionale per le Aree Interne* (SNAI). To examine gaps in educational performance between students attending schools in Central and Internal areas we perform a set of linear regression models separately for each grade. We also include a set of controls to account for composition effects, including students' gender, Socio-Economic and Cultural Status (ESCS) index, school cohort, and macro-area. We also specify an interaction term between, on the one hand, Inner/Central areas and, on the other hand, macro-geographical areas (North, Centre, and South and Islands) and student ESCS. To evaluate whether moving from Inner to Central areas in upper secondary school is likely to influence schooling performances, we rely on a difference-in-differences approach. For this latter analysis, we restrict the analytical sample to students who attended school in Inner areas in both 5th and 8th grade. Next, we observe the same units in 10th and 13th grade, and we include those students who are observed to attend school in Central areas in the treatment group (*moving*), and those students attending school in Inner areas in the control group (*not moving*). Finally, we estimate the Average Treatment Effect (ATE) by subtracting the difference in standardized tests scores in Italian and Mathematics between treatment and control groups in 8th grade, with the difference between the two groups in 10th grade.

Preliminary results document slightly lower test scores in both Italian and Mathematics for students attending school in Internal areas compared to those attending schools in Central areas, with larger gaps observed for Italian-reading tests. Moreover, the size of the gap in test performances appears to increase over grades, being larger in 10th and 13th grades as compared to 5th and 8th grades. While in primary and lower secondary school, the gap in Italian-reading tests (adjusted for the set of controls specified in the model) amounts to 0.025 standard deviations, in upper secondary school this gap reaches 0.05 standard deviations. Gaps between Inner and Central areas in Math test scores follow a similar pattern although they are less pronounced. Furthermore, the Inner/Central areas educational gap is likely to vary across geographical areas. In 5th and 8th grades the geographical divide shows multiplicative effects on the Inner/Central areas educational gap: as compared to the North, we observe a larger Inner/Central areas gap in the Centre and in the South. Results for grades 10th and 13th document however a different pattern. While the Inner/Central areas gap in the South is similar to the one observed in the North, this gap is larger for students attending upper secondary schools in the Centre. Finally, the geographical divide between Inner and Central areas is likely to moderate the relation between students' Socio-Economic and Cultural background and educational outcomes. The advantage in terms of educational performances associated to higher ESCS appears to be slightly smaller in Inner areas as compared to Central areas. This result holds for scores in both Italian and Mathematics and across each grade. Preliminary results from the difference-in-differences approach document a positive effect of the mobility from Inner to Central areas at the upper secondary level on schooling performances. The ATE of moving from Central to Inner areas amounts to 0.06 standard deviations for Italian-reading tests, and to 0.09 standard deviations for Mathematics test scores.

In conclusion, this work documents that, over and above the classical North-South cleavage, the Inner-Central areas divide is also likely to play a role in generating inequalities in educational performances in

Italy and illustrates how educational gaps between Inner and Central areas may have heterogenous patterns across geographical areas, school grades, and students' socio-economic background. Furthermore, it indicates that geographical mobility from Central to Inner areas in secondary school may represent a relevant mechanism which is likely to reinforce geographical inequalities in education.

***Keywords:*** educational opportunities, educational performances, geographical inequalities, central and Inner areas, INVALSI

## The Voice of Students, Parents, Teachers and Principals: The Hermeneutics of Data for the Promotion of a Transformative Educational Alliance from the Bottom-Up

**Antonio Piscopo**

Teach For Italy, as part of the global Teach For All network, focuses its activities in Italy and 60 other countries around the world on fighting educational poverty and injustice.
The lever through which TFI intervenes and strives to contribute in this regard is through young talents from diverse academic and professional backgrounds, united by a deep conviction in the potential of schools to facilitate and create conditions for overcoming educational poverty and promoting social mobility.
As various studies, including the World Economic Forum 2020, have shown, Italy is a country where not only is the social mobility index low, ranking it among the lowest out of the 82 countries examined, but schools themselves are also considered "social silos." Additionally, the persistent gap between the North and South of the country, as highlighted in a recent publication by ISTAT (Regional disparities in the National Recovery and Resilience Plan: Ten objectives for the South of Italy, January 2023), not only shows no signs of diminishing but is further exacerbated by phenomena such as significant youth emigration.
Teach For Italy does not have a universal recipe to revive social mobility or bridge the North-South divide, nor is it the sole entity working towards these goals. The uniqueness of Teach For Italy's approach lies in seeing schools as fundamental to the harmonious and equitable development of the country and overcoming disparities. In summary, Teach For Italy's work can be summarized as follows:

- Innovative and structured selection of teacher fellows based on interpersonal skills, soft skills, ability to work with students, motivation for social impact, etc.
- Intensive summer training for teacher fellows before entering service, with continuous training throughout the academic year.
- Professional growth support for teaching fellows through instructional and pedagogical coaching during two years of teaching experience and an alumni programme afterward
- Constant exchange of best practices with partner organizations within the global Teach For All network.
- A Monitoring, Evaluation and Learning system to promote a learning organization approach within Teach For Italy's organizational context and as a tool for self-monitoring, particularly in the educational relationship between teaching fellows, their students, and parents to foster educational alliances.

Teach For Italy's Monitoring, Evaluation, and Learning system was previously presented in the context of the INVALSI Seminars. In 2021, the focus was on the entire Monitoring, Evaluation, and Learning system as a general approach and its potential for all stakeholders to extract value from the data they generate for self-reflection and participative change processes. In 2022, the emphasis was on how Teach For Italy's MEL places structural inclusion of student and parent voices at the core through a decentralized evaluation approach, serving as a tool for both innovation and pedagogical effectiveness, as well as co-responsibility and the creation of a genuine educational alliance among teachers, students, and parents based on advanced community dynamics based on the hermeneutic practices of community-generated data.
The objective of the 2023 intervention is to address two aspects:
I) Providing the community of experts gathered at the seminar with an overview of Teach For Italy's MEL Framework, including data collection, sharing, feedback, and collective interpretation practices, as well as the specifically developed technological systems (Teach For Italy's Data and Learning Hub) to make data interaction as immediate and accessible as possible for teacher fellows, students, parents, and school leaders. It will also illustrate how the self-monitoring data generated by Teach For Italy's MEL aims to be intertwined with INVALSI data, particularly for schools where Teach For Italy fellows operate.

II) Sharing with the audience the political-philosophical framework underlying Teach For Italy's MEL, which attempts to engage with the ethical fracture in the relationship between subjectivity and their ecosystem that emerged at the beginning of the modern era in Europe and how at the same time propose itself as a remedy to that fracture. This part of the intervention will explore the complex factors that have led to political, organizational, and social approaches, allowing for significant civilizational progress – including the introduction and establishment of widespread education. However, these approaches have also left behind some key tenets of humanistic thought, such as the centrality of the individual as a generative subject responsible for their own destiny (homo faber). Within this context, the practice of "evaluation" emerged as society transitioned towards larger, urbanized communities integrated into production processes, reflecting and incorporating the gradual weakening of achievements made during the time of humanism, such as the sense of responsibility or self-belonging. The emergence of 'evaluation' as a procedure aimed at providing meaning to performance metrics and guiding decision-making processes within a hierarchical structure took place in the following centuries. This mostly unidirectional evaluation approach, with its inherent power asymmetry, represented in the physical and metaphorical distance between the places of the decision making and the places where the actual action takes place, originated an ethical fracture, which traditional evaluation systems exacerbated rather than caused. The philosophical analysis of Feuerbach, focused on the essence of Christianity, becomes the basis for the reflections of philosophers such as Marx, Stirner, Adorno, Horkheimer, and Heidegger, delivering to all of them a partially common semantic field, when describing the phenomenon of "alienation" and the detachment between human beings and their ability to belong to themselves ('authenticity, in the heideggerian expression).

This part of the intervention will attempt to shed light on this ethical fracture, emphasizing how the MEL designed by the author aims to address and heal this fracture through practices rooted in Gadamer's hermeneutics and Habermas' ethics of communication, while also attempting to overcome certain aspects of postmodern skepticism.

The intervention will also provide a description of the model and methodological approach that, during the first three years of Teach For Italy's activity, generated not only data useful for overall organizational evaluation but also for a didactic-educational project that is not only related to the "curriculum" dimension for Teach For Italy's teacher fellows but also to the generation of a healthy co-responsibility among students and parents regarding the educational project.

To frame our approach and its specific characteristics, as well as its evolutionary path, here is an overview of the guiding principles underlying Teach For Italy's Monitoring, Evaluation, and Learning system and its tools.

MEL System: Guiding Principles

Data collection is just a means that, when used effectively, helps us better understand what we do and generates questions about our activities.

Monitoring and evaluation have a direct influence on what is measured and evaluated. That is why we want our "MEL" system to define us as an organization, as a platform for the learning of all stakeholders involved in Teach For Italy's activities, and as an accelerator of impact.

Impact dimensions:

Our MEL system is applied to three impact dimensions: our fellows, our students, and the Italian educational ecosystem. In each dimension, MEL aims to involve the most important stakeholders as active participants in a broader learning process.

Teaching Fellows:

We want our fellows to become the best teachers in Italy during the two-year program and transform into leaders and change agents in the future. For our fellows, MEL is intended to be a platform for reflecting on their personal and professional growth and their impact, promoting:

- Self-reflection
- Impact-oriented teaching skills
- Participatory learning approach

Students

We want our students to feel co-responsible for their own growth and educational journey. MEL aims to listen to their feedback, reinforce their voices, and develop their metacognitive skills through:

- Understanding their development
- Empowering metacognitive skills
- Promoting student leadership

Systems

We want Teach For Italy to promote systemic change in the fight against educational inequalities by involving key actors within the educational ecosystem around the student, both at the local and systemic levels. Our MEL aims to actively engage these ecosystem actors, fostering dialogue through:

- Active dialogue with school leaders
- Dialogue and engagement with parents
- InterInterconnessione between the didactic and extra-didactic dimensions (educational alliances within the community)
- Promoting a shared spirit in tackling educational inequity challenge

Tools:

Our MEL system aims to promote a strong culture of a multilevel "Learning Organization" by employing seven tools to be implemented from the Summer Training School and throughout the school year, addressing all three impact dimensions:

- Monthly Reflections: "L as Learning Together". Teach For Italy staff and fellows gather once a month to reflect on the challenges they face, the impact they are making, and how they work together. During these monthly sessions, data provided by MEL is discussed. At the end of each school year, all aggregated data is studied and reflected upon in a workshop attended by all teacher fellows and Teach For Italy staff.
- Fellow Self-Assessment. Fellows are asked to reflect on their own development and their role as fellows through a survey focused on the Leadership Development Framework.
- Observation Cycle and Growth Conversations. Tutors observe fellows during their teaching activities, based on the Student Rubric. Separate Growth Conversations focus on fellows' leadership development and provide space for individual reflection.
- Student Growth Monitoring. Fellows understand and reflect on their impact through regular monitoring of their students' development in three areas related to Teach For Italy's Student Vision: Personal Leadership, Academic Achievement, and Active Citizenship.
- Student Constituent Voice. Fellows learn from their students through a circular feedback system consisting of two stages: 1) students provide feedback to fellows through a short anonymous survey, and 2) the survey results are returned to the class and discussed together.
- Parent Voice. Starting from the academic year 2021-2022, Teach For Italy piloted a system to include, structure, and strengthen parent voices in two classes. This tool aims to synthesize categories found in "Student Growth Monitoring" and "Student Constituent Voice." Similar to the student feedback process, data collection is followed by plenary sessions where parents and teachers (and, where possible, even students themselves) interpret the data and engage in discussions about the educational project.
- Teach For All Student Survey. This extensive survey is administered internationally by almost all organizations within the Teach for All network. It aims to learn what our students think about their relationship with fellows and our approaches. The Teach For All survey is conducted twice per school year.
- School Leaders' Survey and Feedback. This survey is conducted once at the end of the school year. It provides an opportunity to listen to and collect school leaders' perceptions of our fellows' work, Teach For Italy's program in general, and its impact on students and schools. The survey aims to involve school leaders in the learning and improvement process of our activities and program, enabling structured dialogue with these fundamental actors in the education system. At the end of each school year, a feedback meeting is organized to share the results of our observations on the impact of the program on students and fellows.

***Keywords:*** citizenship, teachforitaly, teachforall, education, school as learning organization

# Students' outcomes in four-year courses: an analysis of data from State exams and INVALSI tests

**Paolo Davoli – Francesco Orlando**

Introduction. In the 2018/19 academic year the Ministry of Education, University and Research MIUR (now the Ministry of Education and Merit, MIM) has proposed a national experiment of four-year upper secondary education programs, as opposed to the regular five-year length of the Italian upper secondary education. This initiative was introduced through Decree n.567 on August 3, 2017. The call for national experimentation pilot was specifically targeting schools offering lyceum and technical programs (thus excluding vocational programs). A total of 192 classes were authorized to participate in the pilot program, with 144 allocated to lyceum and 48 to technical schools. Among these, 127 classes were in state schools, while 65 classes were in private ones. Participating schools were given the autonomy to make adjustments to the national curriculum based on ministerial guidelines. These guidelines emphasized the need for an innovative educational approach that would ensure students acquire in four years the skills expected in the five year programs. In June 2022, the classes which started the pilot program in the academic year 2018/2019 concluded the experimentation by graduating the first cohort of students. These students completed both the regular INVALSI standardized test of grade 13 (which in their case was "grade 12"), and the State Exam, likewise their five-year peers.

Data and methodology. Sixteen schools in Emilia Romagna were chosen to participate in the experimentation starting in 2018, even though only eleven took then actually part in the pilot program, belonging to different tracks (linguistic lyceum, scientific lyceum (ordinary, applied sciences, sports), Information Technology (in the only one technical school). To the best of our knowledge, we are the first to quantitatively analyse the outcomes of this reduced-school-length experimentation in the Italian case. We analyse about 4,000 students in those eleven schools, either in a four-year or in a five-year program, who were assessed at the end of their second year of high school, during the State exams, and by the INVALSI end-of-high school tests. Using descriptive statistics techniques, supplemented by qualitative interviews with State Exams boards of examiners, we merge several datasets:

- information about end of school grades, from the State exams and from the standardized tests INVALSI 2022 (grade 13 and 12) of the students at the end of the five-year and four-year programs;
- grade 10 school assessments and INVALSI tests 2022 for students from both pathways;
- 8th grade data (first cycle exams and INVALSI tests 2017 e 2018) for several of the participating students observed in their last school year in 2022
- 5th grade results for students attending the second year of high school in 2021/22 (unfortunately it is not possible to relate them to grade 8 INVALSI tests, not performed in 2019/20 due to Covid pandemics);
- background data, such as ESCS and parental education levels from INVALSI datasets.

Research hypothesis. The data allow addressing the following research questions, of great relevance for the professional community.

- Due to the Education Ministry requirements, the four-year programs had to be equivalent to the corresponding five-year ones in terms of learning objectives. Is there any evidence to support whether this goal has been achieved?
- Did the implementation of innovative planning, active teaching methods, workshops, technology use, CLIL, and elective courses for orientation purposes have an impact on students' skills?
- The reduction of one year of the school path generally resulted in a higher weekly workload for students. Does this mean that only students with higher academic performance and family support were selected for the pilot or did they self-select into this paths?
- Are there effective teaching and educational management practices that can be transferred to the regular five-year programs?

Results. The data analysis is still ongoing, but several significant findings have already emerged.

1. Results of the 2022 State exam. In the written composition section of the Italian exam ("first test"), four-year students achieved similar scores to their five-year peers. In the "second test" that focuses on the specific fields of study (such as maths for scientific paths), the scores suggest that the four-year students generally attained equal or better skills compared to their five-year peers. There was also a difference in scores in the oral part of the exam, with teachers attributing this evidence to the four-year programs' emphasis on transversal and connection skills due to different teaching.

2. The entry conditions in 2018 varied among the different schools for four-year students who took the 2022 exam. In some cases, there were higher entry conditions for four-year students as compared to their five-year peers(e.g., higher grades in the first cycle exams or school-based selection), while in other cases entry conditions were similar for both groups (and there was one instance where four-year students had lower entry skills compared to their five-year counterparts. Therefore, the four-year students' outcomes do not appear to be solely influenced by a possible selection or self-selection.
3. The better performance of the four-year students does not appear to be attributed to commission compliance or favourable bias (only internal commissioners and external president, in 2022). In fact, in cases where the examination commissions had more discretion in making choice (such as awarding *cum laude* or discretionary bonuses up to 5 points out of 100), the behaviour of four-year programs commissions was similar to, or even more restrictive than, the commissions for five-year programs.
4. Grade 12-13 INVALSI results. The outcomes of Italian tests appear comparable between the four-year and five-year programs, with a higher concentration of four-year students in the intermediate levels 2-3-4 (out of 5 levels). In Math tests, four-year students performed better than five-year students, with fewer students in levels 2 and 3 and more in levels 4 and especially 5. This difference is less pronounced for scientific lyceum tracks, while for linguistic lyceum and for the unique technical tracks the results for four-year courses were even better. Better performances were also recorded for four-year students in the English tests. These overall results align with the findings from the State exam scores.
5. According to the teachers involved in the four-year experimentation, the teaching approach is characterized by significant attention to methodologies such as interdisciplinary approach, group work, flipped classroom, formative assessments and tutoring, with technologies playing a crucial role. It is reported that these diverse methodologies are essential for organizing the curriculum (based on five years) within a compressed time frame. The design work for teachers is relevant, promoting greater collegiality and interdisciplinary planning, which requires specific motivations and stability among teachers in their classrooms.

***Keywords:*** four-year programs, state exams, standardized tests, student outcomes

# THEME 4. GENDER STEREOTYPES

**ORGANIZER: INVALSI**

**COORDINATOR: ELLEN CLAES**

**24TH NOVEMBER: 2.00 P.M. – 4.00 P.M. {ROOM 2 – RESEARCH SESSION 12}**

---

## The impact of anxiety on girls and boys differences at the INVALSI: A multilevel path analyses with meta-analysis study

**David Giofrè – Tommaso Feraco – Enrico Toffalini – Sara Caviola**

Introduction. Literature suggests that, on average, females outperform boys in school achievement. A host of research, however, indicates that the opposite might happen in math, although the effect size is limited and heterogeneous across contexts. Examining INVALSI data, Giofrè et al. (2020) found that gender differences in math follows a geographical gradient in Italy, being larger in northern regions and smaller in the southern ones. This pattern is explained only in small part by economic factors and by average performances. This result suggests that other factors, such as socioemotional ones, might somehow moderate the effect. We in fact know that math, and the general school performance, is affected by several factors, including for example self-esteem and anxiety. As for anxiety, this is of utmost importance also because there is a host of research indicating that girls present higher levels of academic anxiety.

Object and hypotheses. The aim of this study is to understand the impact of anxiety levels on both math and reading performances using INVALSI data. Following the main literature outcomes, we hypothesized that girls would have higher performances in reading, while boys would have higher performances in math. At the same time, we also expected higher levels of anxiety in girls. The novelty of this work is to focus on whether differences in both academic domains would have been eventually reduced (in mathematics for girls) or were amplified (in reading for boys), after considering the impact of anxiety.

Data. Data from the INVALSI were analyzed for this study. Since 2006, the INVALSI has collected data on math and reading abilities for the Italian population of 2nd, 5th, and 8th graders. We focused on the representative samples (i.e., students belonging to classes that were followed by members of the INVALSI committee, which guarantee higher reliability of the data and lower cheating possibilities) of 5th graders from the academic year 2011/2012 to the academic year 2016/2017. We decided to focus on 5th graders because an additional questionnaire measuring students' academic anxiety was administered to children. The selected sample counts 146,227 students in total (73,670 males), with the smaller sample size being 19,383 (academic year 2014/2015) and the larger sample size being 29,332 (academic year 2011/2012). 72 more students (0.05%) were available, but were not considered because they did not identify as male or female.

Method. Multilevel path analyses were used to analyze the data. Schools were treated as a clustering factor to account for dependencies in the data, also allowing us to examine heterogeneity. The first model included only the effect of gender on the two outcome variables: math and reading abilities, whereas the effect of anxiety on the outcomes was set to zero. In the second model the complete pattern of relations was specified: gender predicted both the outcome measures and anxiety, with anxiety directly predicting math and reading performance. In this case anxiety was set free to mediate the effect of gender on math and reading performance. We then estimated how the effect of gender on math and reading performance changed after anxiety was entered in the model. Standard errors were obtained with multilevel bootstrapping with 1,000 iterations. Finally, estimated effects calculated for different years were combined via random effect meta-analysis to offer general estimates not disaggregated for years.

Results. Concerning the first multilevel path analysis model, girls showed lower performances in math (with betas ranging from -.21 to -.13 in each year) while at the same time having higher performances in reading ($.06 \leq \beta \leq .19$). When looking at the second model, however, the pattern of relations was somehow different: the disadvantage of females was lower in math ($-.14 \leq \beta \leq -.08$), while their advantage in reading was even higher ($.13 \leq \beta \leq .25$). Indicating that the effect of anxiety was considerable and was having a stable effect on the results. Both math ($-.24 \leq \beta \leq -.13$) and reading ($-.22 \leq \beta \leq -.14$) were affecting the level of anxiety, which was higher in girls compared to boys ($.30 \leq \beta \leq .36$). Meta-analytic synthesis suggested that the average effect of gender on math was -.17, which was reduced to -.11 after accounting for anxiety (meaning that about a third of the effect was explained by the effect of anxiety). As for the effect on reading, the

pattern, favoring girls, increased from .13 to .20 (also in this case indicating that the effect was amplified by about a third by the effect of anxiety). Effects remained statistically different from zero, suggesting that although anxiety might explain a large portion of gender difference variance in math, about two thirds of the variance was not explained by anxiety.
This study might have important theoretical clinical and practical implications. From a theoretical point of view we demonstrated that anxiety is an important determinant of academic achievement. This is particularly relevant since data are on a large and representative sample of Italian children. Also, this pattern was stable across several different academic years. From a clinical and practical perspective these results showed that anxiety can in fact hinder the potential of several children, who are underperforming in both math and reading despite having good skills. Future studies should probably focus on specific training aimed at lessening the anxiety of children, and in particular of girls, for which the impact of anxiety seems to be noteworthy.

***Keywords:*** gender, sex, mathematics, reading

## Gender stereotypes and children's performance in elementary schools

**Piera Bello – Annalisa Cristini – Federica Origo**

Introduction. The belief that some domains are for males and others are for females is normally due to cultural stereotypes, i.e., to preconceived opinions that derive from processes of oversimplification and overgeneralization (Bordalo et al., 2016). A widespread cultural stereotype is that girls are worse than boys in science and math and better than boys in arts and humanities.
It has been demonstrated that such a stereotype, when part of a teacher's beliefs, can have important consequences on children performance in middle schools and consequently on children's high school choice (Carlana, 2019).
While there is evidence that even in primary schools, boys perform better than girls in maths (Colella, 2014; Matteucci and Mignani, 2021; INVALSI, 2022) no much evidence exists on the role played by teachers' gender stereotypes at this school level (the only exception is Sule et al, 2018). At this regard, an analysis is interesting on several grounds. On the one hand, the early school years provide the foundation for children's self-confidence and the competence beliefs that children develop in the first three years of the elementary school are crucial predictors of subsequent choice and performance (Wigfield et al., 1997). On the other hand, it not yet clear the extent to which the gender differences in competence beliefs that children express in the early years in school can actually change in time, in response to new experiences with peers and teachers (Eccles et al., 1993).
Object and Hypotheses. The objective of the paper is to contribute to this literature by investigating the effect of primary school teachers' gender-role stereotypes on children's grades. While Sule et al. (2018) focus on explicit stereotypes, we collect information on both implicit and explicit gender stereotypes of elementary school teachers.
In light of the existing evidence, our hypothesis is that stereotyped teachers negatively affect the relative performance of girls in maths as compared to boys. However, since teachers and parents are the predominant role models in early childhood, we expect the effect to be even larger than in high school.
Data. Our analysis covers 40 state schools of the province of Bergamo, in Northern Italy, 196 teachers, and 2627 students. The teachers in the sample were asked to take the Gender-Science Implicit Association Tests (IAT), which measures the strength of teachers' implicit stereotype according to which girls are generally associated with Arts and Humanities and boys with Maths and Science. We use the IAT score obtained as a proxy for implicit measure of stereotypes.
Moreover, teachers were asked to answer a questionnaire including questions of the European Value Survey related to gender, which we use to measure explicit stereotypes.
We then associate teachers to the class or the classes in which they teach and link each student in the classes to the standardized blindly-graded results of the INVALSI tests taken in the teaching years 2021-22 and in 2018-19.
Methods. In order to analyse the effects of teachers' stereotypes on student performance, we exploit two identification strategies (in line with Carlana, 2019). The first one, by using class FEs, leverages the variation

in gender gaps observed among students within the same class, while the other one, by using school FEs, exploits variation in the performance of students of the same gender across classes assigned to teachers with different stereotypes.
Results. Our results provide important insights on the effects of teacher stereotypes on student performance. First, we document a detrimental impact of teachers' stereotypes on the maths performance of girls. In grade II, pupils in classes assigned to maths teachers with a 1 standard deviation higher IAT score exhibit a 0.257 standard deviation larger gender gap in maths performance. In grade V, the substantial magnitude of the coefficient is confirmed but it is no longer statistically significant. No effect is found for literature teachers, in line with Carlana, 2019.
Moreover, by exploiting school fixed effects, we investigate the effects of stereotypes separately on boys and girls. We show that stereotyped teachers have a positive effect on boys' maths performance, while harming the girls' one. However, the negative influence on girls is much more pronounced than the benefit seen in boys.
Additional evidence shows that the background of teachers, specifically the type of high school they have attended, plays an important role in shaping the effect of gender stereotypes.
Finally, we analyse whether explicit gender stereotypes have a similar effect on the gender gap in student performance. While the same detrimental effect is found for maths, the performance of girls in literature improves if the teacher is stereotyped.
Our results are robust to several robustness checks, such as the use of a dummy variable instead of the continuous IAT score variable, the exclusion of outliers, an alternative specification with interaction between the variable of interest and a dummy for grade II.

***Keywords:*** gender stereotypes, implicit association test, student achievement, INVALSI test scores

## Bridging the gender gap in mathematics: examining STEM enrolment rate by gender in Higher Education, in Italian Provinces

**Giorgio Bolondi – Clelia Cascella – Federica Ferretti – Ugo Rizzo**

Introduction. Identifying factors explaining gender differences has been considered as a priority for both the national and the international policy and research agenda. The current paper aims to contribute to the existing debate by investigating the possible relationship between gender differences in secondary education and the probability of enrolling in a STEM (Science, Technology, Engineering, and Mathematics) major. As it is well-known, the number of in-demand jobs in the STEM fields is rapidly increasing to the point that the World Economic Forum (WEF), in 2021, estimated that gender parity was farer to reach than expected because of women's low interest in STEM subjects/majors. Previous studies have already shown that girls' intention to enrol in a STEM major is largely affected by their achievement in mathematics and science in secondary school (e.g., Delaney & Devereux, 2019): the better the girl achievements, the higher the probability they enrol in a STEM major. The first objective of this paper is to replicate these results at the level of Italian provinces. We therefore test whether the average gender gap of NUTS III regions explain the probability of females residing in those regions to enrol in a STEM major. Female underachievement in mathematics has been explained by a number of different factors, such as social conditioning and stereotype threat (Guiso et al. 2008, Nollenberger et al., 2016; Nollenberger & Rodríguez-planas, 2017; Noseck et al., 2009; Penner & Paret, 2008; Pope & Sydnor, 2010), and thus by the division between *gender identity* and *mathematics identity* (e.g., Radovic et al 2017). Gender stereotypes are considered to shape students' development and understanding of their own gender identity, thereby learning from the context what is deemed more appropriate to do, think and feel depending on their gender. In this research strand, more recent studies have shown that gender stereotypes are socially, culturally, historically and thus geographically located and, therefore, that results based on nationally aggregated data can be incomplete, if not totally false, thus potentially misleading or making policy interventions completely ineffective, if not dangerous (eg. Cascella & Pampaka, 2020, Cascella, Pampaka & Williams, 2018). In light of these results, as with previous studies (e.g., Cascella et al 2022), in the current paper we took student's place (Butler & Sinclair, 2020) as a socio-cultural and historical concept as it refers to the communities, and the locales students live in, where students participate in culturally mediated and gendered practices (e.g.,

Bourdieu, 1990) thus learning and *practicing* the meaning of 'being a girl'. In the present study, we thus hypothesise that girl intention to enrol in a STEM major is linked to 'contextual' factors that (i) affect girls' achievement in math and science (especially) in secondary education, and (ii) can create a contradiction between 'feminine identity' and 'mathematical identity'. In particular, we hypothesise that gender-biased environment (for example, in terms of females' lower participation in politics and/or in the job market, as well as the higher proportion of women in roles/jobs stereotypically considered as appropriate to woman) can explain girls' absence of their interest toward STEM subjects and, therefore, their underachievement (especially in secondary education) and their loose of interest in STEM majors as they are preparatory for jobs deemed not suitable for women. Compared with previous research, the current research contributes to knowledge by locating the exploration of the hypothesised relationship (between personal and contextual factors in explaining the probability girls have to enrol in a STEM major) at the provincial level. The second objective of this paper is to investigate the role of contextual factors in moderating the effect of the gender gap in maths on the probability of enrolling in a STEM major.

Data, Methods and Expected results. We investigate our research questions by building an original panel dataset at the Italian provincial level (NUTS III) from 2018 to 2022. We combine the information provided by the Ministry of Education on the number of enrolments in STEM versus non-STEM faculties by province of residence of the students, with the maths gender gap calculated on the year before the enrolment thanks to the data provided in the INVALSI (Italian national institute for the evaluation of the educational system) surveys. INVALSI surveys Italian students' population for mathematics and reading at various grades: from 2018 it started to surveys also grade 13, that is the last year of secondary school. By averaging mathematics scores at the province-year level we firstly aim at investigating the following empirical model:

(1) $Enrolment\_rate_{pt} = \beta maths\_gendergap_{pt-1} + \gamma \boldsymbol{X}_{pt-1} + \alpha_p + \tau_t + \epsilon_{pt}$

Where represents the share of female students residing in province *p* at time *t* that enrols in a STEM course where *t*represent the academic year following the secondary school graduation year; measures the gender gap in maths score collected from the INVALSI surveys, averaged at the provincial level, for year *t-1*, that is the students grade 13 academic year; therefore represents our main parameter of interest; *X* represent a set of control variables at the provincial level collected mostly from ISTAT, and and represent respectively province and time fixed effects, that serve to control for time-invariant idiosyncratic features of Italian provinces, and for unobservable heterogeneity which is constant across cross-sectional units, but is time variant.

Once identified the (expected) effect of the gender gap in mathematics on the probability of enrolling in a STEM major, we turn to the investigation of the moderating effect of contextual factors. We will therefore identify some principal provincial-level factors expected to moderate the relationship investigated in equation (1)

(2) $Enrolment\_rate_{pt} = \beta maths\_gendergap_{pt-1} + \gamma \boldsymbol{X}_{pt-1} + \delta\, maths\_gendergap_{pt-1} \times \boldsymbol{X}_{pt-1} + \alpha_p + \tau_t + \epsilon_{pt}$

In equation (2) our parameter of interest is represented by that identify the moderating effect of the context on the relationship between gender gap in maths and enrolment rate in STEM. In so doing we aim at contributing to explaining the role of social conditioning and stereotype threats on the probability of female students to enrol in STEM courses via the performance in maths test in secondary school.

***Keywords:*** gender gap, STEM enrolment rate, gender identity, mathematics identity, higher education

## Peer interaction and gender stereotypes in primary education: A literature review

**Catalina Miranda**

Gender equality is a demand that persists over time and is sustained as a global challenge (UN, 2020; UNFPA, 2020). In this sense, we seek to achieve gender equity in different dimensions, such as in social interactions and educational contexts, in order to reduce the stereotypes that are widely embedded in society (Hsu et al., 2023) and bias everyday spaces (Flores et al., 2023). In relation to social interactions, it is observed that parents and peers have the greatest influence on how children internalize gender norms and attitudes as they grow up (UNESCO, 2022). In education, specifically in schools, gender stereotypes can translate into gender-based violence (e.g., psychological violence), which affects children's well-being and learning.

(UNESCO, 2022). At a general level, the index of gender social norms, which includes dimensions and indicators of the political, educational, economic and physical integrity spheres, shows that almost 90% of people have at least one sexist prejudice, and these are present in both men and women. (Hsu et al., 2023). Given the demand for reducing gender stereotypes in different social spheres, there is an emerging interest in focusing attention on early education and the social interactions generated in the classroom. School institutions are configured as a two-sided coin. On the one hand, they are meeting spaces that positively impact the student body, since they provide forums for discussion and dialogue and interactions among peers (Treviño & Miranda, 2021) and counteract problematic behaviors (Varela et al., 2020). On the other hand, school education operates differently for each child, and specifically, for each gender (p.e. Follegati et al., 2022) reproducing gender stereotypes.

Stereotypes are cognitive schemas used to process information about others. These stereotypes reflect beliefs, characterizations of social groups and predetermined roles. This construct shows a quality attributed to a person, but also the expectation of a behavior (Dovidio et al., 2010). A distinction can be made between descriptive and prescriptive stereotypes (Gill, 2004; Rudman & Glick, 2001). The first describes those who belong to a social group (e.g., children are aggressive) and the second describes the behavior that a person of a certain social group should have (e.g., children should be aggressive). Both may be implicitly present in social interactions within the school classroom.

Gender stereotypes vary according to the developmental stages of the child and thus, the form of social interaction with peers (e.g., La Freniere Strayer & Gauthier, 1984). Stereotypes, in the first instance, are transmitted in the home (Hernández et al., 2020), and the student upon entering school and interacting with peers expresses in the form of gender segregation (i.e. the individual meets with a person of the same gender and not opposite) (Fabes et al., 2015) such stereotypes acquired and spontaneously articulated since childhood (Maccoby, 1990). As the child grows older, these associations become more sophisticated and even at the age of five or six these stereotypes are rigidly maintained in the child's life (Martin, & Ruble, 2010). Along these lines, it is hoped that the school, as the years go by, will make these cognitive schemes more flexible and generate spaces for collective coexistence.

In the face of gender stereotypes (cognitive schemas) that are developed over time (cognitive development theory) and exposed in social interactions (intergroup theory), it is relevant to investigate their permanence throughout their school trajectory. Cognitive development theory indicates that advancing or limiting cognitive skills affects the construction of social categories and their meanings (Bigler & Liben, 1993). This theory goes beyond the interaction between people and contextual factors, it also relates the way of processing, inferring and remembering information according to the individual's cognitive schemas. Thus, the cognitive development theory refers to the way of understanding how gender stereotypes are formed (Bigler & Liben, 1993). In other words, this theory allows us to understand how stereotypes are formed and thus how they are expressed at a later age. From this last point, the importance of the theory of intergroup development emerges, which allows explaining stereotypes through multiple simultaneous characteristics (e.g., gender, race, sexual orientation, among others) that converge in the interaction between peers (Bigler & Liben, 2007).

Thus, gender stereotypes have repercussions in different areas of life and one of them may be academic performance or relationships between actors within the school. However, in order to address these challenges in the educational trajectory of students, it is first necessary to recognize the gender stereotypes present in early education and particularly in the interactions between peers in the classroom. Therefore, the present study posed the following review question: What does the empirical evidence on gender stereotypes represented in peer interactions in primary education report?

To address the review question and ensure the reproducibility of the study, the PRISMA methodology was used to operationalize and systematize the information resulting from the search and selection of manuscripts (Page et al., 2021). The search was performed through the Web of Science database (SSCI y A&HCI) included the following syntax: TS=("gend*" OR "stereotypes" OR "roles") AND TS=("interaction*" OR "play" OR "behavior") AND TS=("peer*" OR "classmates") AND TS=(classroom OR school*) AND TS=("elementary school" OR "primary school") NOT TS= ("early childhood" OR "preschool" OR "college" OR "higher education" OR "tertiary"). The result was 344 articles. The second filter was by time period (2012-2023), languages (Spanish and English) and empirical articles. The selection was reduced to 194 papers. A third filter was the WoS subject categories (Psychology, Education and Social Sciences). The number of articles selected was 180. Subsequently, exclusion criteria were established for the review of abstracts, firstly, and secondly, for the review of full articles. These exclusion criteria were according to the educational stage of the student, purpose of the study and educational actor. From the first review

(abstracts) 81 articles were retained and in the second review (full text) 25 articles were selected for analysis.
The selection criteria returned 25 research studies, which were examined according to their focus of interest, sample, location, method and instrument, and finding. Subsequently, a thematic analysis characterization was performed to expose similarities and differences among the findings. For this processing of results, a five-phase thematic analysis was applied (Nowell et al., 2017), i) data familiarization: organization of the selected documents according to their focus of interest; ii) generation of initial codes: use of code matrices; iii) search for themes: connection of concepts and themes by category; iv) definition of the name of the categories; and v) production of the report or detailed description of the coding process according to category and analysis.
At a general level, most of the articles are quantitative; with geographic variety, but centered in the United States and Europe; and focused on students from 5 to 12 years of age. Also, two macro categories were distinguished: 1) expectation of inclusion or exclusion from a peer group according to gender stereotypes and 2) friendship preferences according to gender stereotypes.
In the first category, 15 articles are grouped and the information is obtained through surveys or interviews. Two patterns of results were found. The first, children relate to peers of the same gender, which indicates that they maintain a process of inclusion of peers of the same gender in play dynamics and classroom activities. A second pattern of results corresponds to the exclusion of students of the opposite or different gender than the student reporting their relationship with other peers. This exclusion manifests itself in teasing or concern for not being chosen in game dynamics.
The second category is made up of 10 articles and the data are reported through surveys or interviews. Only one pattern of results was evident. Gender segregation in terms of friendship preference is only observed in males but not in females. This pattern of results includes articles that show that students choose peers of the same gender and with stereotyped behaviors, while girls do not differentiate their friendships according to gender, but according to stereotypes. That is, girls relate to peers of the same gender and do not differentiate between girls with stereotypical and non-stereotypical behaviors, but do indicate that they are not friends with men who express themselves according to male stereotypes.

***Keywords:*** gender stereotypes, early education, peer interaction, school, literature review

# THEME 6. THE EVALUATION OF PUBLIC POLICIES IN THE SCHOOL SYSTEM

**ORGANIZER: INVALSI**

**COORDINATOR: ANDRES SANDOVAL HERNANDEZ**

**24TH NOVEMBER: 2.00 P.M. – 4.00 P.M. {ROOM 4 – RESEARCH SESSION 13}**

## Effects of Workload Allocation per Course on Students' Academic Outcomes: Evidence from STEM Degrees

**Carmen Aina – Koray Aktas – Giorgia Casalone**

When preparing for university exams, students have to allocate their time efficiently to ensure that they study a certain amount of material proposed by their course teachers. Several studies investigated the impact on students' outcomes of the number of years of schooling at different educational levels (Card 1999; Pischke 2007; Cappellari and Lucifora 2009; Krashinsky 2014; Marcus and Zambre 2018), of the amount of coursework during college degrees (Arteaga 2018), and of the organization of academic calendars (Bostwick et al. 2022). However, no evidence exists on how workload allocation across university courses within degree programs influences college students' performance. Notably, the literature on higher education lacks evidence indicating the optimal organization of the study time.

Let us consider a simple situation wherein a college student takes two mathematics courses—say, mathematics I and II; each course is associated with five university credits (ECTS). Accordingly, they sit two final exams to successfully complete these courses at the end of the semester. In a different situation, a student takes a mathematics course with 10 credits and sits an overall examination; hence, they have to study more course material per exam. A priori, the effects of the increased workload per course are ambiguous. On one hand, studying for the unified exam can help students enhance their knowledge of the subject, while sitting fewer exams can also mean less stress for students. This can foster non-cognitive skills, as students must learn to better manage their time via self-imposed deadlines. On the other hand, evidence reveals that students procrastinate and therefore benefit more from externally imposed deadlines (Ariely and Wertenbroch 2002). Tuckman (1998) considers frequent testing a motivational equalizer. When students sit frequent exams, they receive greater feedback (De Paola and Scoppa 2011). When students sit unified exams, they must study and process a great deal of information for one exam (i.e., task juggling), which can culminate in worsening academic performance. The results of an experiment by Buser and Peter (2012) demonstrate that individuals who multitask and are free to organize their own schedules perform worse than those who are forced to work on tasks sequentially.

This study documents the first evidence in the literature on how workload reallocation among university courses, which increases the average workload per course, affects college students' key academic outcomes. Accordingly, we exploit the exogenous variations in the number of courses, holding constant the total number of university credits students must achieve to graduate across degree programs and academic years. Further, we investigate graduates' early labor market performance before and after the change in degree programs' structure. Our identification strategy builds on the exogenous variation in the number of courses (i.e., exams) induced by a nationwide tertiary education reform in Italy. The Ministry Decree 155/2007 (known as Mussi's Reform) sought to standardize degree programs' organization across country and reduce the time taken to complete a degree. To homogenize degree programs and avoid the excessive fragmentation of curricula, the reform sets a maximum cap for the total number of courses required for degree completion (for instance, 20 courses to complete a three-year bachelor degree) without altering the number of credits required to achieve the degree (180 credits for three-year bachelor degree). To comply with the reform requirements, degree programs having more courses than the maximum number set by the law had to change their curricula, predominantly by integrating two or, in some cases, three courses into one. Hence, the workload intensity per course increased mechanically.

We use the administrative dataset of Università del Piemonte Orientale (UPO), which contains all the details regarding the students' enrollment records and performance during their degrees, as well as information regarding these students' pre-enrollment characteristics. At the UPO, the reform affected degree programs differently. For example, students enrolled in the bachelor STEM programs required an average of 33 courses to attain their degrees and eight courses in their first year; after the reform, these numbers fell to 20 for degree completion and five in the first academic year. However, the number of credits required to

graduate (180 credits) as well as their distribution across years (about 60 credits per year) did not change. By contrast, the number of courses required for bachelor's degrees in health care sciences organized by the medical school remained unchanged because the course numbers were already at the level required by the reform.
Therefore, a quasi-experimental environment was established following the reform's implementation. The law compelled all universities to revise their degrees by the beginning of the 2009/10 academic year. In our setup, STEM degrees reorganized their course schedules since the 2009/10 academic year. In the difference-in-differences framework, we consider the three-year degree programs in STEM subjects as treated units, whereas the three-year degree programs offered by the medical school are considered the control group. We estimate the effect of the reallocation of workload among university courses—and the consequent increase in the average workload per course—by comparing the differences in the outcomes of the treatment and control groups before and after the reform's implementation. Throughout the paper, we present results on varied academic outcomes, focusing on the effects on first-year dropout and degree completion rates. These two academic outcomes are the most policy-relevant indicators of higher education efficiency (Aina et al. 2022). Additionally, they are not subject to endogeneity problems as they are available to all students.
Our working sample covers the academic years at first enrollment (i.e., matriculation) from 2002/03 to 2014/15, which allows us to comprehensively investigate the validity of the parallel.
Our results demonstrate that after the introduction of unified exams, students in STEM degrees are, on average, up to 21.5 percentage points (p.p.) more likely to fail all their exams in their first year, which increases the first-year dropout probability by 18 p.p. We provide further evidence that the effects are more pronounced for high-procrastinators, suggesting that students fail to effectively organize their study time for exam preparations. The increase in the first-year dropout rate translates into a decrease in the graduation rate by approximately 18 p.p. Moreover, the results are robust to controlling for student composition through the exact matching of key pre-enrollment characteristics. However, we do not observe any significant effect on the other outcomes related to degree completion—namely, time to graduation (measured in months) and final graduation marks. Our heterogeneity analysis reveals that high-ability students are significantly less affected by the new curricula with respect to their dropout decisions at the end of the first year. Overall, the difficulties of new exam schedules and improved student composition of degrees after the first year, seemingly, cancel each other out—in terms of their effects on final graduation marks and time-to-graduation. Finally, we examine the early labor market performance of students who achieved their degree after the reform using the survey data of AlmaLaurea on Italian graduates' employment conditions one year after graduation. We observe sizable effects of unified exams throughout their degrees on the employment rates of graduate students. The effect that we estimate for graduates with STEM degrees is an average 13.4 p.p. increase in employment rate, which becomes a 24 p.p. increase when we control for the students' composition by performing exact matching among graduates based on pre-enrollment characteristics. This indicates that reduced class size and increased peer quality after the first year play key roles in explaining the effect on labor market outcomes.

***Keywords:*** number of courses, college dropout, graduation, European credit transfer system, credit per course, higher education

## The Effect of School Closures on Students' Academic Performance in Italy

**Josep Amer-Mestre – Sara Flisi**

Introduction. Following the outbreak of the COVID-19 pandemic in early 2020, more than 60 million students in the EU were temporarily kept at home when schools were fully closed as part of the policy response to limit the spread of COVID-19. The share of students affected by full school closures varied over time, in line with the cyclical pattern of the pandemic, with more than 90% of students sent home during the first wave in spring 2020. Subsequent waves saw lower shares of students affected, peaking at around 35% in January 2021. While full school closures became less frequent during subsequent waves of the pandemic, many schools still operated at limited capacity. In April 2021, for example, around two-thirds of students attended schools that were only partially open.

A range of national studies shows large variations in the impacts of school closures on learning progress. This reflects considerable cross-country variation in the intensity of the pandemic, length and extent of school closures adopted, different modes of distance or hybrid learning adopted, and also the type, scope, and timing of measures adopted to mitigate learning loss.

This paper investigates the impacts of the pandemic and in particular of physical school closures, on academic performance in one of the countries that was most severely hit by the pandemic, i.e. Italy.

Institutional background and existing literature

Italy was the first EU country to experience the outbreak of the COVID-19 pandemic, and the first one to impose a national lockdown. This included the closure of commercial establishments, the prohibition of events and exhibitions, the limitation of individual mobility, and the closure of schools at all levels, as well as the large-scale institution of home-based work ([2]). During the 2019/2020 school year, the country registered one of the longest periods of school closures in Europe.

In November 2020, to address the second wave of the COVID pandemic, the Italian government introduced a new policy whereby different levels of restrictions were applied depending on the local epidemiological situation and risk level. More in detail, based on a set of indicators, Italian regions (and autonomous provinces) were classified into four areas – red, orange, yellow, and white – corresponding to different risk scenarios, for which specific restrictive measures were foreseen. The indicators and the resulting classification were reviewed weekly. As a result, the time spent by each region/autonomous province in each risk zone could vary greatly. Different levels of restrictions implied also a different combination of in-person and online learning at different education levels. This heterogeneity represents the basis for the identification strategy for the study.

A few studies already attempted to measure potential learning losses produced by school closures during the pandemic in Italy. Contini et al. (2022) combine data from the standardized national tests administered by INVALSI (the National Institute for the Evaluation of the Education and Training System) at grade 2 with results from a novel standardized assessment administered by the research team at the end of grade 3 in the province of Turin. By comparing the pre-COVID and the COVID cohorts (sitting the grade 3 tests in April 2019 and October 2020 respectively), they find a large mean negative impact on pupils' performance in mathematics. Borgonovi and Ferrara (2023) use INVALSI data for the whole country to provide a comprehensive assessment of the medium-term effects of pandemic-related disruptions on the mathematics and reading achievement of students who completed primary school and lower-secondary school in the summer of 2021. By comparing this cohort with the pre-COVID one, completing the cycle in 2018/2019, they find drops in math achievement at both levels and a more mixed picture for reading. Bazoli et al. (2022) also study the impact of school closures in Italy, focusing on grades 5, 8, and 13, by comparing the performance in such grades before and after the pandemic using a matching approach. Contini et al. (2023) study the impact of a full year of the COVID-19 pandemic on school performance, focusing on students at the end of upper secondary school, who are therefore about to enter the labour market or start university without having had the opportunity to recover; they find that the pandemic had a huge negative impact on students' performance in mathematics and reading.

Data. The study is based on two main data sources. First, the study relies on anonymized administrative data collected by INVALSI (National Institute for the Evaluation of the Education and Training System) to measure students' academic performance.

The study uses the data available for the above-mentioned school grades for the years up to 2022. No INVALSI examination was administered in 2020 due to the lockdown measures put in place in the spring of that year. The school year 2021/2022 is also of particular interest to our study; recent meta-analyses (Betthäuser et al., 2023; Di Pietro, 2023) suggest that the negative impact of the pandemic on students' achievement in the first months of the pandemic normally persisted in the following year. This study investigates this aspect as well.

Second, the study uses a dataset containing information on the colour zone that was in place in each region and day for the 2020/2021 school year. This dataset can be expanded to also include further relevant local information on the spread of the virus within each region.

Methodology and preliminary results

We rely on a difference-in-difference strategy to compare students' test scores who, due to the existent regulations, were exposed to a higher number of days of online learning with those who enjoyed more in-presence school.

Descriptive evidence shows that higher grades were more severely hit in the first period of the pandemic, and are still not showing clear signs of recovery. The scores of the post-pandemic cohorts in these grades

are considerably below 2019 levels in both 2021 and 2022. INVALSI data indeed shows that grade 13 was the one showing the highest decline in test scores between 2019 and 2021, and virtually no recovery in 2022. A smaller but still clear decrease in test scores was registered for grade 8 as well, for both subjects.
The decrease in academic performance in lower grades was stronger in mathematics, while gaps in reading appeared only in 2022. Both grades of primary school show a small decrease in test scores in 2021, which further widened in 2022. In reading, the scores of students in 2021 are even above those of pre-pandemic years, while they are lower in 2022.
Preliminary results of the difference-in-difference strategy show a clear deterioration of the standardized test scores of 13th and 8th grade in 2021 in both Italian and Mathematics when compared to those obtained by students in 5th grade. Scores of 13 and 8 graders were 42% and 19% of a standard deviation lower in Italian, and 22% and 10% lower in Maths in 2021, respectively. In 2022, students in grade 13th were still scoring 18% of a standard deviation below their younger counterparts in 5th grade. However, 8th graders in Italian as well as both grades in Mathematics obtained similar scores on average as those obtained by 5th graders.
Overall, there seem to be signs of recovery after the initial negative impact of the pandemic on students' performance. These results are in line with the evidence provided by the literature; recent meta-analyses indeed suggest that the negative impact of the pandemic on students' achievement in the first months of the pandemic normally persisted in the following year.

***Keywords:*** school closures, online learning, covid19

## Teachers and educational inequalities reproduction: how can we reduce tertiary effects? Evidence from an RCT combined with a factorial survey experiment

**Elisa Manzella – Gianluca Argentin**

Introduction. The Italian educational system, like other Western countries, is characterized by a high degree of educational inequalities (Schizzerotto & Barone, 2006) and tracking in secondary school is an important factor of their reproduction (Gasperoni, 1997; Checchi & Brunello, 2007; Triventi et al., 2016). Indeed, upper secondary education track choice has a significant impact on students' further academic and occupational careers. Access to tracks differs by socioeconomic and cultural background, conditioning the choice (Gasperoni, 1997). Information gap plays a crucial causal role in this process, as shown by a previous RCT (Barone et al., 2017).
In order to get a conscious educational choice, it is required an effective school guidance, a key factor for the success of students' educational paths, especially the ones from lower family social background. Hence, an effective school guidance may also increase equity of the entire school system. In Italy a national school guidance system is still lacking, with the exception of a (not mandatory) teachers' advice. This is provided to each student at the end of the lower secondary education and is formulated by teachers. Usually, the advice is focused on the choice between the different tracks of high school. The guidance advice represents the only device operating at national level, although it is highly heterogeneous in its implementation. In addition, being the advice not mandatory (despite used a screening criteria by some high schools), its influence on students' actual choices seems various and unequal.
Previous quantitative studied showed the existence of robust associations between students' origins and their teachers' advice (Checchi, 2010; Bonizzoni et al., 2014; Argentin et al., 2017), but they tell us only few about underlying factors the advice; on the other hand, previous qualitative studies that analysed the advice's formulation processes, highlighted teachers' considerations on extracurricular and ascribed elements (Romito, 2016). These studies shade light on the mechanisms underlying teachers' advice formulation, but they exploiting only partially the existing variability, being focused on local contexts. Overall, previous studies present two crucial additional limitations: i. they did not directly face the issue of causality and ii. they did not provide proposals for remedial interventions.
Aims. The aim of this study is causally test whether raising teachers' awareness about their biases when formulating the advice and the related role in inequalities reproduction may reduce inequalities among students. In order to test this hypothesis, we try to increase teachers' awareness on these issues and to see whether their advice towards upper secondary school change consequently.

Data and Methods. In order to pursue the aim, we implemented a randomized control trial to assess whether teachers' biases may be reduced raising teachers' awareness by a professional training. We randomly provided to half of our sample of lower secondary school teachers' the opportunity to attend a professional development course focused on school guidance and related inequalities.
We aspect that, after the intervention, students assigned to treated teachers will display a school guidance advice less strongly associated to their ascriptive characteristics (social background, gender and migratory background), after controlling for students' performance. Moreover, we aspect too that students assigned to treated teachers will display lower association between their social background (parental education and social class) and at student level, we aspect that students assigned to treated teachers will display a school guidance advice, at the end of s.y. 2021/2022, less strongly associated to their ascriptive characteristics, after controlling for students' performance. Moreover, we aspect too that students assigned to treated teachers will display a lower association between high school track enrolment and their social background (parental education and social class), after controlling for students' performance.
We recruited 196 schools and randomly assigned them to treatment (102) or control group (94), blocking the randomization on macro-region, urban size, school size (total of 12 blocks). The treatment, developed starting from a previous work by Gerosa and Argentin (2022), consisted of an online light touch professional development targeted to teachers called to formulate guidance advice and aimed at: i) making them aware of the processes of reproduction of inequalities at stake; ii) provide some recommendations about how to neutralize these processes. The intervention was delivered though a dedicated web platform, lasted about 5 hours and was delivered before teachers' formulation of guidance advice in s.y. 2021/22.
Thanks data linkage with administrative data, coming from INVALSI and the Ministry of Education - National Students' Registry, we are able to assess the impact of our training on: i. the net association (meaning after controlling for students' performance) between students' ascriptive characteristics and the advice provided them by their teachers; ii. students' educational choices in current school year (s.y. 2022/2023), to assess whether the advice was followed and with which results in terms of education pathways.
The administrative data contain information on the 19.136 students of the third class of the s.y. 2021/2022, belonging to our 196 randomized schools. More precisely, the data include students' academic performance, their final exam grade, INVALSI tests scores (Italian, Mathematics and English), the guidance advice formulated by the teachers, the actual enrollment choices for the a.y. 2022/2023 and ascriptives and socioeconomic students' characteristics. This dataset allows to assess the advices' bias for the schools and, comparing treatment and control group, to estimate the intervention effect. Similarly, we can compare students' actual choices in upper secondary school tracks between treatment and control group, assessing the intervention impact on their enrolment.
Results. By analyzing the effects of the intervention in the real world, a positive trend that aligns with expectations is detected. The information treatment generated among teachers a reduction in the bias in the guidance advice, related to students' ascriptive characteristics. The difference in the advices' formulation goes in the expected direction with a reduction in the bias by family's education and social classes, or by migratory background, also counteracting some spontaneous tendency detected in the control group. The treatment group tends to reduce the bias by sending students from more disadvantaged social backgrounds less frequently to vocational schools and also sending students from higher social classes less to classical and scientific high schools.
On the overall sample, real world effects are small, are limited to the extremes of school tracking distribution (higher lyceum and vocational track) and remains statistical uncertain. Nonetheless, he effects are larger on students displaying intermediate final grades, the ones where previous studies detected larger bias.
Conclusions will be drawn on the relevance of teachers' awareness for future actions to counteract social inequalities in education.

***Keywords:*** upper secondary school choice, educational inequalities, teachers, RCT

# THEME 11. EQUITY BETWEEN LOW ACHIEVING STUDENTS AND THE EXCELLENT ONES: BOTH SIDES OF THE SAME COIN

**ORGANIZER: INVALSI**

**COORDINATOR: PAOLO BARABANTI**

**25TH NOVEMBER: 9.30 A.M. – 11.30 A.M. {ROOM 1 – TEACHING SESSION 1}**

## Equity, quality and inclusion in schools in the autonomous province of Trento: examples of good practices

**Anna Maria Ajello – Paolo Barabanti – Chiara Tamanini**

*Introduction.* Since the early years of INVALSI tests, students of the autonomous province of Trento have kept good performances. Outcomes in primary school are always in line with the national ones; results of lower secondary schools are equal or slightly higher than the ones of the neighboring provinces (Argentin, Tamanini, Vergolini, 2021). Levels of students in the upper secondary school show, especially for grade 13, a positive gap also in comparison with those of many regions of the North of Italy. Therefore there is a growing positive trend in results from primary to upper secondary schools.

Reflecting on data about schools of the second cycle of education, it is necessary to take into account the educational reform, unique at national level, which introduced in the autonomous province of Trento a dual system in 2011: the professional institutes were eliminated and merged into the technical institutes path and the system of vocational education and training (VET) helped both students with weak results complete compulsory education and students interested in professional qualifications undertake suitable paths to enter the world of work or to be adequately accompanied to the higher technical education.

VET is a training system – part and parcel of the second cycle of education – which has positive effects on fighting early school leaving and in promoting inclusion; students attending VET participate in the INVALSI surveys of grade 10, showing lower results than those of high schools and technical institutes, but better than those of VET in other regions (Tamanini, Oliviero, Covi, 2020). Another continuous element, highlighted thanks to INVALSI data, is the positive situation about equity, especially for schools of the first cycle of education, and the decreasing hidden drop-out rate. Finally, a particularly intriguing result is that the autonomous province of Trento in the post-pandemic survey of 2021 managed to preserve the same – if not higher – levels of 2019, therefore not presenting the so-called learning loss matter (Ricci, 2021).

*Object of research.* This paper tries to analyze the positive performances in INVALSI tests of the students of the autonomous province of Trento also achieved by the ones with learning fragility and difficulties, to then detect possible factors and measures aimed at tackling them, anticipating actions which have taken a pioneering character on the national level.

The reasons for this substantial temporal stability in learning levels are certainly complex and multidimensional and may mostly be investigated in a first snapshot by quantitative data; further inferences relating to the school policies of the autonomous province of Trento, to the structural reforms and to the interventions on teaching and training as elements which might be instrumental in those results.

As an autonomous province, more resources have certainly been available than in other regions, but it is relevant to highlight that it has been decided to invest them heavily in the world of education. It is also hypothesized that this focus on the education, which probably has historical roots in the Austro-Hungarian tradition, has made it possible to anticipate and pragmatically address some problems and critical moments in the school field.

In this way, a positive attention to education and to quality of learning by school principals, teachers, parents and non-profit associations was supported and promoted, as well as a focus on the livability and sustainability of school structures. These reasons, in addition to the presence in the area of some organizations specifically dedicated to educational research and experimentation (IPRASE) and system evaluation (Provincial Law 7 August 2006, n. 5), could have contributed to keep good learning results.

*Data and method.* Empirical evidences on quantitative data refer to INVALSI data; qualitative insights were then done by focus groups, regarding specific aspects that emerged as worthy of further analysis.

The research questions are as follows:

- What previous measures and actions have led to an absence of learning loss after the Covid-19 pandemic? Is there a possible connection with the focus on digital training started around five years before the national digital plan?
- May strong INVALSI results in English test have been kept by the promotion of bilingualism begun in 2006 and strengthened by "Trentino Trilingue Plan" started in 2014 and gradually confirmed with some necessary corrections?
- How, more generally, may the teachers' training activity offered in the autonomous province of Trento about inclusion and equity be linked to positive results in INVALSI tests?

*Main results.* From the analysis of the ways in which the school system of the autonomous province of Trento has faced some scholastic problems, it is possible to highlight some indications on good practices implemented to promote equity, quality and inclusion.
Even with some inevitable elements concerning problems, guidelines that can be improved and a peculiar institutional and territorial situation, some past and current practices in Trentino schools are to be indicated so that they also might be interesting for other contexts.
The identification of educational choices and operational strategies highlighted by the interventions of school leaders involved in the focus groups indicate methods of intervention that can also be useful in a more general perspective in a phase in which Italy is engaged in the use of a large amount of resources to improve the education system as a whole.
In addition, a set of reflections on the possibility of investing the available resources in a positive and constructive way in the school environment, enhancing research, training, experimentation and evaluation are proposed.

***Keywords:*** Autonomous Province of Trento, INVALSI tests, learning loss, best practices, equity

## Fragile and high performing male and female students in coexisting environments: how do they differ in characteristics? A case study based on the goals of the Agenda 2030: providing quality education and achieving gender equality

**Alessia Cividin**

Since the publication of the 2030 Agenda for Sustainable Development in 2015, a strong focus on equity in learning has been observed, particularly in the first cycle of education. This focus is based on Goal 4, which aims to provide quality, equitable, and inclusive education and learning opportunities for all, and Goal 5, which aims to achieve gender equality and empower all women and girls. The purpose of this research is to highlight the results of a case study conducted among male and female students who took the INVALSI exams for the fifth and eighth grades in 2022 and 2023 at a lower secondary school in the province of Trieste. The study aims to investigate students with high-level fragility and top performers who achieved levels 4 and 5 in the standardized INVALSI tests during the 2022/23 school year. The analysis illustrates the main characteristics of these students, considering variables such as gender and participation in educational paths aimed at promoting gender equity and gender equality, particularly referring to projects implemented during the 2022/23 school year. The methodology used in the study defined three groups of students, and for each of these groups, the individual results and the adherence to gender-related educational paths were identified. The population considered in this case study includes primary school students and lower secondary school students from non-sample classes who participated at INVALSI tests in May 2022 and 2023, specifically the fifth-grade classes and the eighth-grade classes. Students who were absent from one or more tests were excluded from the investigation. The focus of the study is on the analysis of data related to the students' fragility in the INVALSI tests for Italian and mathematics. Fragility is defined based on the performance levels achieved by the students. Students who reached level 1 or level 2 in the tests were considered fragile, while achieving level 3 indicated a risk of fragility. No fragility was attributed to reaching level 4, level 5, or level 6. To investigate the fragilities in the population under consideration, various variables concerning the students, provided by the school, were used along with several indices reflecting well-being at school, teacher-student ratio, test anxiety, and peer relationships. These indices were constructed using information collected through institute projects aimed at improving the well-being

of the student population, in which male and female students participated during the 2022/23 school year. Through the analyzed data, the research identified the conditions in which students exhibit fragility and those in which they are highly performing, as well as potential factors and characteristics influencing these conditions. This document aims to draw attention to the importance of gender equality in quality education, the significance of considering sex and gender in standardized assessments, and how the theme of gender equality has become integrated into educational practices. Despite the progress shown by the data, achieving gender equality remains necessary. The study highlighted conditions that can lead to project choices aligned with these objectives.

***Keywords:*** Agenda 2030, fragility, top performer, gender, gender differences

## From SES to training orientation: a regional experience of equity and improvement

**Ornella Campo – Giorgio Cavadi**

Recent studies have recognized the school as a crucial role in the transmission and construction of a cognitive heritage of knowledge, as well as in the development of socio-emotional skills (SES) (Corcoran et al., 2018; Greenberg et al. 2003, )in a perspective of maximizing the potential of each student.
With the issuing of the recent Guidelines for Guidance - D.M. 328/2022- the educational orientation assumes a central value in the school context, understood as a complex and unified set of activities and actions capable of developing, through learning processes, the empowerment of the person, especially in the direction of the construction of the self and the definition of an autonomous project of life.
Considering these reflections, the Regional School Office for Sicily - in collaboration with the Foundation for the School of the Compagnia di San Paolo di Torino - has started a training course on socio-emotional skills and educational orientation, addressed to School Managers and Teachers of all levels.
The governance organisation was entrusted to a "Regional Support Unit" with the task of:
-defining the accompanying actions of the Guidelines for guidance to ensure coordination at territorial level with the actions provided by the Ministry of Education and Merit;
-managing organizational and implementation aspects in support of educational institutions; developing guidelines for the construction of the vertical orientation curriculum;
-planning of activities of connection in the transition School/ University and organizing formative/ informative moments addressed to the school staff.
Object and hypothesis of research. The construction of the vertical orientation curriculum stems from the need to ensure regional governance that can support schools in the implementation phase of the "Guidelines" and avoid the overlapping of actions and interventions with inevitable repetition and discontinuity in the design and implementation of the 30-hour training modules to be carried out in all secondary school classes of grade I and II, as required by the aforementioned "Guidelines".
In the construction of the vertical orientation curriculum, we have tried to define for each grade and order of school - also starting from kindergarten - the objectives in terms of learning orientation that, of course, can be replicated in the various school orders.
The support team will identify a sample of secondary schools of I and II grade that, from 2023/2024, will experience in teaching practice the orientation curriculum to verify - through the data of the tests INVALSI - in the transition from grade 8 to grade 10 how guidance teaching has influenced the choices and quality of the subsequent school career. Through the comparison of INVALSI results in grades 8 and 10, the outcome of the career for both excellent and" fragile" students will be verified.
Key elements of the vertical orientation curriculum. Goal of competence: to develop the ability to build a project of life as a dynamic process, functional adaptation to the needs of learning subjects that change in different stages of life, ensuring continuity in processes.

1. Orientation learning objectives expressed by school order (with a matrix that can be regional).
2. Identification of disciplinary objectives with orientation, at school level, by the teachers of the different departments.
3. Contribution of training agencies of the territory, belonging to the productive world at the territorial level.
4. Contribution of the student component to Grade I and II secondary school.

5. Contribution of representatives of universities and AFAM.

Data used. Throughout Sicily, 30 courses have been organized by the polo schools for training, in which about 20 teachers of each order and grade for each course have participated. Ex ante monitoring data are being acquired. The impact in terms of improvement of learning outcomes will be verified through the use of the results of the tests INVALSI

Expected results

- Reduction of dispersion through more conscious and targeted choices of the subsequent course of study and, in general, of the choices that concern the personal life project.
- the configuration of the school as a training centre in the surrounding social context, recognising the central role of guidance responsibility; both to ensure the achievement of educational success at school and out-of-school by the majority of pupils and to limit the occurrence, among young people, of situations of hardship, marginalization or self-exclusion.
- Promote a process of self-knowledge and self-awareness.
- Increase the level of student awareness of the variables involved in the educational and professional choices (resources and personal characteristics to refer to in the present, to design their own future).
- To promote skills enabling young people to develop appropriate decision-making processes.
- To identify appropriate disciplinary knowledge to support and accompany the development of self orientation.

***Keywords:*** equity, excellent students, early school leaving, learning weaknesses

## Gifted Students In The Education System: Identification, Challenges And Strategies. For A Mixed But Complementary Use Of INVALSI Data And Qualitative Tools

**Brunella Fiore – Paolo Barabanti**

*Introduction.* Gifted students have an exceptional cognitive potential and they may show extraordinary abilities in one or more fields of knowledge or in other domains, including artistic or non-cognitive ones.

The topic of giftedness is experiencing a renewed interest among researchers, teachers and policy makers beacuse there has been a grawing interest in the matter of the merit of individual students and in the ways to help them enhance their talent as a collective capital good with an impact not only on the individual but also on the society as a whole. However, there are still significant challenges for the education system in order to properly identify gifted students, according to shared scientific standards and increasingly distant from particularisms and arbitrary considerations, and to adopt effective strategies to value their talent.

Since several researches highlighted the fact that giftedness is not necessarily accompanied by high academic performance, it is therefore not always enough identifying giftedness with top performing students in a specific cognitive standardized test. In particular, when teachers do not have suitable tools and identification pattern to properly recognize giftedness, the risk of school-leaving and drop-out becomes greater. It therefore becomes crucial trying to rely on suitable methods to read data, to detect pupils with giftedness and to focus on strategies to enhance this type of students within the classroom context.

From 2012 to 2018, a ministerial technical-scientific committee worked to issue some guidelines, within a regulatory framework which, at that time, was being rapidly defined. However, the arrival of the pandemic interrupted the ongoing process of debate and discussion.

More recently, ADI (Italian Teachers and School Principals Association) has started a review of international experiences, with a specific attention to what OECD proposes in 2022 on giftedness.

*Research object.* This paper aims to explore the issue of giftedness in the education system, encouraging ideas and comments on both the latest research and best practices about the identification of this kind of students. Its objective is to provide a comprehensive overview of the strategies for identifying and measuring the excellent talents of gifted pupils, based also on the use of international and national data.

In particular, it focuses on the specific challenges that arise in the education context about high potential pupils, such as the need for personalization of learning paths, managing students' and teachers' expectations and fostering an inclusive school environment and setting.

*Data and method.* After having reviewed the national and – especially – international literature on the subject, in particular from a psychological and sociological point of view, methods and ways through which

INVALSI returns tests data to schools will be analyzed to provide useful tools for properly reading them and for identifying particularly top performing students.
*Main results.* The International literature review show that the experiences among countries are progressively differentiating: in particular, in East-Asia there is a significant invest on the issue with a specific attention to the economy growth of each Country; there are also some important experiences in Europe, albeit within contexts that are often more reluctant and which tend to bring the giftedness only in the fields of social inclusion and of the special education needs. In particular, some studies of ECHA (European Council for High Ability), in Netherlands, are highliting the spreading of this matter also in Europe.
Italy, however, seems particularly absent and late on that, even if there are some centres that involve young researchers and academics to study the phenomenon in a sociological-psychological way.
Schools, with the exception of the few ones linked to academic centres in specific networks, are mostly absent though. In particular, there is a lack of interest among school principals' and teachers' organizations; these school professionals are sometimes expressing some perplexity if not hostility, too. At the moment, there is a draft law about the recognition of pupils with high cognitive potential, recently re-proposed in the current legislature.
The use of national data, such as the ones deriving from INVALSI tests, may right away provide a solid basis for identifying and enhancing the excellence of students in the Italian school system, promoting a more equitable and satisfying educational path for all gifted students. Italian schools might, actually, already use some tools, including INVALSI data, to preliminarily identify the levels of excellence (the microdata return specific data for each student) which then require, for example, a qualitative integration by customized grids and rubrics.

***Keywords:*** Gifted students, mixed methods, INVALSI data, talent, excellent students

## Excellence And Fragility In Mathematics Learning: What Methodological Approach Can Guarantee An Equitable School System

**Carmina Laura Giovanna Pinto**

Introduction. In 1997 Jacques Delors in 'In Education a Treasure' stated: 'More than ever, the fundamental role of education seems to be to give individuals the freedom of thought, judgment, feeling, imagination that they need in order to be able to develop their talents and to remain in control of their lives as far as possible...'. Recognising and promoting talents by giving them the chance to develop into excellence should be the first educational objective, regardless of entry conditions and as a prerequisite for any educational intervention, expressing itself also in the case of any fragility detected in learning. From this point of view, the equity of the school system is declined in a series of interventions aimed at filling any fragility highlighted at entry, so as to enable everyone towards personal progression towards excellence.
Purpose and research hypothesis
This paper answers the research question: how do we interpret the concept of equity when there are excellent students in a class and at the same time students who show fragility in learning in mathematics?
In the name of equity, does the school system respond in a way that raises the average class level or in a way that lowers the average class level? One answer could be the personalisation of the educational pathway without 'the fear of daring', offering students with learning weaknesses opportunities to compare and collaborate with excellent students, so that the 'floor effect' does not occur, but on the contrary the 'ceiling effect'. It is necessary to offer challenging opportunities that allow each student the vision of new opportunities for personal growth, also by personalising the educational pathway so as to allow each one to discover his or her own talent, to allow his or her personal fulfilment through the achievement of his or her own excellence. With this interpretation and management of excellence and fragility in the classroom, the equity of the school system, is the expression of a responsible and aware education that offers everyone the chance to assert themselves and improve over time.
Data used. This study analysed INVALSI mathematics data from the school year following the evolution of excellent pupils (score not less than 230) and pupils of the same class fragile (score not more than 170) from primary second grade a.s. 2015/16 to primary class V a.s. 2018/19 and of excellent pupils (score not

less than 230) and pupils of the same class fragile (score not more than 170) from secondary second grade a.s. 2015/16 to secondary class V a.s. 2018/19.
Method. From the INVALSI mathematics data from the second primary class a.s. 2015/16, data of excellent pupils (score not less than 230) and fragile pupils from the same class (score not more than 170) will be extrapolated. With the identification of the pupil codes, these will be searched among the INVALSI mathematics data for class V primary a.s. 2018/19, comparing the outcome results and their evolution also in relation to the evolution of the outcome compared to the average of the whole class. The same method of work will be carried out in the case of secondary school.
Results. The aim is to analyse, by following the evolution of the pupil's school career by means of his or her identification code, whether certain levels of excellence/ fragility have been maintained over time and whether they have changed towards an improvement or worsening compared to the average level of the class to which they belong, thus highlighting any correlation with the class context and giving an account of how in the Marche Region the ethical value of equity in the school system has been interpreted in the three-year period 2015/16- 2018/19, also in relation to the occurrence of conditions characteristic of the "floor effect" or "ceiling effect" in the classes under study.

***Keywords:*** equity, mathematics, excellent students, fragile learning

## How To Read The Educational Fragility ... To Intervene

**Francesco Mammarella – Ettore D'Agostino – Roberta Franchi – Carla Lavista**

Introduction. The Omni-comprehensive Institute of Città Sant'Angelo is divided into primary school, lower secondary school and high school with five courses: human sciences, linguistics, human sciences, with economic and social option, applied sciences and sports. The study, starting from the INVALSI data on fragility relating to the 2022 survey and from the school results of the 2021/2022 school year, aims to interpret two polar phenomena such as the phenomenon of educational fragility and that of excellence in students who however, have emotional, social and relational weaknesses, regardless of their level of learning. To this end, the scholastic career of these students was monitored and analysed, in progress and at the end, stimulating and encouraging their "immersion" in design paths and formal, informal and non-formal learning environments that could enhance awareness of their best skills and for which, therefore, psychological well-being becomes a strategic lever to improve learning.
These interventions were oriented both on individual classes (6 third year high school classes in their passage therefore from the compulsory two-year period to the final three-year period) and on individual students, noting the correlated empirical evidence. The aim is to tend educational action towards a fair school system that integrates merit and inclusion, development of talents and support for fragile students at risk of dropping out of school, in a dynamic of research and constructive and comparative exchange.
Object and research hypothesis. The object of the research is represented by the process of identifying and implementing types of interventions that can help support the educational growth of students, both of those who start from previous learning disabilities, and of those who instead show excellent results in standardized and scholastic surveys used for research. The research hypothesis consists in verifying whether targeted interventions implemented by the school supported by the educating community can help compensate for any fragility at several levels and if overcoming such fragility (emotional, relational, social) also leads to improvements in the levels of educational and learning fragility. At the same time, the work intends to analyse how much this process of fragility rebalancing also affects those who show excellent results in the educational process since, as psychologist Michael Pluess points out, the healthy development of an individual necessarily involves a continuous reorganization of one's global functioning , in order to increase the level of adaptation to environmental demands, adaptation which is expressed in terms of greater autonomy and capacity for emotional and behavioural regulation.

Data used. The data used during the analysis were: 1) the results of the INVALSI surveys for the year 2022 and the consequent data relating to fragility in the disciplines being surveyed; 2) the assessments of individual pupils expressed by the Class Councils at the end of the 2021/22 school year; 3) the surveys of emotional, relational and social fragility in students both with learning difficulties and with excellent results during the October 2022 Class Councils, meetings with the Institute psychologist, the bimonthly monitoring of the class coordinators and the results of the Team Emergency; 4) the assessments of individual pupils expressed by the Class Councils at the end of the 2022/23 school year.
Method. The methodological approach used in this work was that of research-training and metacognitive teaching through three major interventions: 1) PCTO third class: the PCTO courses of the third classes of the high school entitled "Going to school as teachers" were carried out in the lower secondary school and in the primary school of our institute, being an Omni-comprehensive institute, and in the nursery school of another institute, according to various actions such as, for example, the observation of the teaching and learning process and the analysis of the path with the employer to allow students to know the value of being reflective professionals, having the future in their eyes, aware that the work of teachers and ATA staff requires time, study, innovation and above all attention in everything that is done and towards each student; 2) Activation of NOPs linked to the sphere of the development of autonomy and self-knowledge such as Erasmus in action, L'Angolino in stampa, InAcquediverse, Sportivamente implicate, Diversimente insabbiati, Civil…mente, Mathematical animation. Each NOP stimulated the dimension of sociality and hospitality, through immersion in learning contexts Community and a re-reading of one's own experiential and emotional life; 3) Motivational meetings: the meetings were held by Sports teachers and were based on Motivational Coaching activities with small groups of a maximum of 4 students in order to work on each one in a personalized way and have a picture at the end of the meeting cycle precise of those towards whom the school had to subsequently direct further actions aimed at safeguarding the right to study, the educational success of each student, the enhancement of excellence. We started from the general analysis of the class and then moved on to a careful and systematic monitoring of the situations of individual students to analyze how much fragility of various kinds reduces in a more or less accentuated way the ability to maintain a sense of wholeness and internal centering.
Results. At the end of the interventions (with further available data such as the disciplinary assessments of the 2022/23 school year, the Self-assessment Unit, with the support of the teachers who have followed the INVALSI training course "Communicating and managing relationships in the evaluation process", analyzed the path of the students who had been identified as relevant for the purposes of the research and therefore their school career was monitored and analyzed against the background of the three major interventions in which they had been included: PCTO, NOP and motivational meetings , in order to outline which interventions had substantially contributed to overcoming the emotional, relational or social fragility found at the beginning of the school year, if this overcoming had led to changes in the students' educational path and in what terms these changes had manifested themselves.

***Keywords:*** equity, excellent students, early school leaving, learning weaknesses

# THEME 9. ITALIAN, MATHS AND ENGLISH: TEACHING PROCESSES THROUGH INVALSI DATA AND TOOLS

**ORGANIZER: INVALSI**

**COORDINATOR: ROBERTO CAPONE**

**25TH NOVEMBER: 9.30 A.M. – 11.30 A.M. {ROOM 2 – TEACHING SESSION 2}**

---

## The language of INVALSI Mathematics tests

**Ottavio Giulio Rizzo**

INVALSI Mathematics tests are mainly "world problems," i.e., "a text (typically containing quantitative information) that describes a situation assumed familiar to the reader and poses a quantitative question, an answer to which can be derived by mathematical operations performed on the data provided in the text, or otherwise inferred" (Greer & al., 2002).

World problems strive to make a mathematical question more relevant to the student (and they do not always manage to do so, given the ontological category of "[senseless] mathematical problems," cf. Zan 2017) but the addition of a natural language layer could introduce a further obstacle to comprehension if the language is not clear enough to the student. This is particular relevant from an inclusiveness perspective, since disadvantaged students will often not have Italian as mother tongue (be it a local dialect or a foreign language).

On the other hand, it is clear the need that "an individual»" — and henceforth a student — "should be able to reason mathematically, understand mathematical proof and communicate in mathematical language." The language of mathematics, that is, has to be clear to students.

Clearness of language has many different components, amongst which a relevant one is vocabulary: do students actually understand the meaning of all the words in the problem? Indeed Ferrari (2021) warns that "parts of learning difficulties in Mathematics are of linguistic origin".

De Mauro (2019) recognizes a "fundamental core" of 2000 Italian words, which any functional speaker of the language can be supposed to know; 3000 "high frequency" words, which occur often in the spoken or written language; 2500 "high availability" words, that albeit not appearing frequently are usually known to native speakers (e.g., "pepper"). Following De Mauro, we call these 7500 words the "basic" Italian vocabulary, to which we have to add 33,000 "common" words (and some 220,000 regional, poetical, specialistic or obsolete words). The Italian school system does not have an official set of words that have to be known at a given grade, but the introduction to De Mauro's fundamental core points out that it is not unreasonable to consider the basic 7500 word vocabulary as commonly known at grade 8.

Toth (2021) analyses the word frequency of basic and common vocabulary in INVALSI Italian tests at grade 5, 8 and 10 and finds that word frequency does not explain the difficulty level: this is partially surprising, and Toth suggests that further inquiries are needed on students' vocabulary to understand the phenomenon.

Rizzo (2022) shows that the same phenomenon holds for the G8 Mathematics tests. We confirm, in this work, that the same holds for G2 and G5.

- In grade 2, 98% of words belongs to the basic vocabulary in the first ventile by difficulty (6%–26% of correct answers), 96% in the last ventile (85%-93% of correct ansers), while 94% of words of all items belong to the basic vocabulary.
- In grade 5, 95% of words belongs to the basic vocabulary in the first ventile by difficulty (7%–29% of correct answers), 97% in the last ventile (85%-96% of correct ansers), while 94% of words of all items belong to the basic vocabulary.

As source data we used the corpus of INVALSI Mathematics tests of grade 2 and 5. Collected and sorted words were stemmed using the SnowBall (Porter 1980) library; automatically sorted amongst fundamental, high frequency, high availability, common, and specialistic vocabulary. What remained was sorted by hands or cast away as non words (e.g., ABCD in "the square ABCD")

The analysis of words let me reconstruct a "basic mathematical vocabulary" for primary school: this vocabulary was based on Baruk (1992) and supplemented with all words not belonging to the basic vocabulary which I determine to be "mathematical" (e.g., "decina", i.e. "tens").

***Keywords:*** primary school, Mathematics, Italian

# A new idea of CLIL with INVALSI maths questions

**Stefano Babini – Ivan Graziani – Chiara Cataldo**

Many Secondary and Primary school maths books often show English questions that are not always so well written. Besides, some weird INVALSI examples can be found in the last pages of many chapters. As a consequence, we suggest some CLIL activities, providing a special maths.

CLIL (Content and Language Integrated Learning) is an acronym that was born in Europe in the early 90s. At that time, European Institutions such as the European Council and the European Commission gave life to a process using foreign languages as means to convey other subjects in the School system. *"The need to find solutions was pointed out: the aim was reaching higher levels of foreign languages' knowledge. The traditional language teaching was not enough anymore"* (Balboni, 2015) The so called "Buona Scuola" Law (L.107, 2015), article 7, defines as its main educational goals *"The development and enhancement of linguistic competences through Italian and English as well. Moreover, this process may concern other European languages, even through CLIL method "*

When a student faces a subject in English, he has to:

- learn English better, that is to say, CLIL has to improve his competences of the foreign language;
- understand the teacher during his/her lessons, when he/she describes, assumes, explains, defines, comments, compares, calculates (as far as Maths is concerned). The student has to read texts that are different from the ones he/she is used to.
- fulfil learning tasks and all the related cognitive process;
- create written and/or oral texts (answers, ratios, synthesis, descriptions, explainations, definitions) using maths through a new language.

This represents a linguistic and cognitive challenge for the student (and for the teacher as well, who has to deal with the process). Therefore, when a maths teacher wants to plan a CLIL activity he/she has to be able to adapt his/her own didactic attitude and to put an innovative method into practice, which includes some strong points ,such as:

- creating an active, stimulating and entertaining learning environment;
- adopting task-based activities, peer-education and cooperative learning;
- showing maths contents in a real and visual way, even using multimedia support;
- creating didactic material (between English and maths);
- introducing and authentic, integrated and formative evaluation.

A strong motivation is essential to plan a CLIL activity: getting a satisfactory outcome may be hard without this initial boost. Maths teachers must use an inspiring and engaging didactic method to show their contents using English. For instance: teachers should provide useful input, explaining abstract elements with examples, using different materials, in order to make students understand mathematical terms in English (using maps, charts). In addition to this, students could browse webistes providing interactive resources.

The Khan Academy's operating system (www.khanacademy.org/) has been paramount to us: it has helped us to join together English and Maths and we used it especially in the creation of Maths and Science CLIL units. Salman Khan was its founder, he is an American with Bengali roots. This teacher started his activity nearly by chance: he started to upload his maths lessons on Youtube since his cousins needed private lessons but they lived too far from him (Graziani & Babini, 2018).

The CLIL methodology stimulates the language proficiency (in terms of lexicon, fluency,communicative effectiveness). CLIL also improves the soft skills (taking part in a discussion, asking questions, expressing a personal opinion), cognitive skills and an autonomous reasoning. In conclusion, this methodology promotes students' self esteem and self confidence, since they begin to think that they can use a foreign language as an operative tool.

This method consists of teaching different school subjects in a foreign language. This could promote both the learning of the language as well as the learning of the different subject matters.

For this reason and thanks to our English colleague, we came up with the idea of playing with this aspect. We presented some maths exercises taken from the INVALSI database, and we translated them into English.

We asked ourselves: Is English really helpful or is it a further obstacle for maths young learners? Is the familiarity with Italian language really so beneficial?
We believe that using multiple ways may turn out learing maths advantageous. As a matter of fact, sharing different points of view, adopting a cooperative learning, splitting students in little groups: all these means might facilitate peer education and a deeper understaning of certain concepts.
Understanding maths content is crucial for its learning and CLIL may be even harder because:

- A student can reach its knowledge through the language;
- A language may "shape" contents and ongoing ideas. A foreign language makes all this even harder, if we do not facilitate its comprehension.

As a consequence, we looked for INVALSI questions that could be easily translated into English and that did not require difficult explanations in a foreign language.
We chose the questions to be translated, thanks to the different search mode provided by the Gestinv 3.0 Operating System (www.gestinv.it). We chose the 8 and 10 degrees, too.
Providing these maths exercises in English has been a way of testing competences and knowledge of our students in both the subjects. This process also enabled them to better understand the studied topics.
Furthermore, facing these particular tests, the students had to try to answer properly, after having grasped the real meaning of the question.
We also wanted to see whether our students looked for shortcuts (like hasty readings or a search for keywords) making the usual mistakes, or if they were led to read more carefully, thanks to the foriegn language.
Interestingly, we noticed that our target students made fewer mistakes when they had to read English texts. This is because they paid more attention to their reading comprehension since when they read in their own mothertongue, they are more likely to be distracted.
This allowed us to inform them on how important it is to well understand a text, even though we refer to maths questions.

***Keywords:*** CLIL, INVALSI questions, learning, reading comprehension

## Storytelling mathematics with INVALSI questions

**Ivan Graziani – Stefano Babini**

Learning in mathematics is a complex and time-consuming process. To foster it, it is important to characterise it with varied and meaningful experiences, so that it is then truly lasting and competent.
Telling stories is a natural competence of human behaviour and doing so in mathematics is very important, in order to contextualise theorems and other contents in a temporal and human way, which acquire greater value in this way.
Narrative methods, typical of storytelling, are increasingly proving to be more effective than other forms of communication. Indeed, storytelling owes its strength to its ability to engage the recipient, to make the abstract concrete and to be remembered over time. Narrative involves the recipient because, if we focus on the dynamics of reception, we can say that: a narrative is a text that takes you by the hand and accompanies you to the end (Perissinotto, 2020).
The story of mathematics is also important because it gives a cultural depth to the subject, becoming an exceptional tool for communicating content, methods and contexts, gradual difficulties, participation and even motivations.
Knowing that there are 'histories' of mathematics can make us realise that

- this discipline, like the others, has not always been the same, from the time of Euclid or Leonardo to the present day, but has evolved over time, with stumbles and magnificent leaps forward.
- They have also developed in a physical space. They have their own geography, their own places, and have spread thanks to the contamination between people, ideas and knowledge.
- there is a social and cultural context of reference within which the facts, and thus also the insights of the 'scientists', of yesterday and today, take place.
- Mathematics is an expression of the human soul, conceived by the minds of men and women, undoubtedly talented, with their merits and flaws. And this is not to reduce the narrative to anecdote,

but certainly to intrigue, motivate and bring the student closer to scientists as people and not as superheroes.

- mathematics was born out of both concrete needs and the inherent need in human nature to ask questions, to formulate problems and try to find solutions, to prove claims.
- there are no absolute truths, as the discovery of irrational numbers or that electrons are not just the smallest parts of matter, for example, show us.

It is also very important to approach mathematics through active methodologies, such as storytelling, also approached through cooperative work in small groups, so as to foster peer dynamics that help students better understand certain concepts. This has also long been suggested by the National Directions for the first and second cycle, the Guidelines and the various documents produced by the Ministry in recent years.
For older students, it will be easy to show them and convince them that "the natural way to conceive a demonstration is to construct it as a story, and to understand it, to listen to it as a story (Lolli, 2018).
If we then want the teaching-learning process to bear the 'hoped-for' fruits, we cannot neglect these aspects in order to maintain this pair, because there can be no real teaching if this does not lead to real learning in students.
Furthermore, it is important that the activities, which we want to propose to the students, are truly challenging and that they are interested in them, in order to be able to involve them and make them feel truly at the centre of their training and learning process.
We thought of storytelling through facts and events related to mathematics.
Denis Guedj in his book "The Parrot Theorem" (2000), wrote: "From time to time, a name would be heard echoing: Thales, Pythagoras, Pascal, Descartes, but it was only a name, like that of a cheese or an underground station. There was not even any mention of where and when a certain episode had taken place: the formulae, the demonstrations, the theorems ended up on the blackboard as if no one had created them, as if they had always existed, in the same way as mountains or rivers.... And it got to the point that theorems had a timeless air... Mathematics was neither history nor geography nor geology. But then what was it?"
One of our aims was to dispel the myth of superb, abstract and self-enclosed mathematics and, by recounting the various characters, places, events and its importance for human civilisation, to give it a more tangible context through history. A way to see, for the first time, mathematicians not as abstract characters or superheroes, but as people immersed in their time.
Together with the students, we talked and researched a few stories about mathematics, Euclid, Pythagoras, Thales and Nicolas Tartaglia, Peano, and others.
The various anecdotes create memorable points, typical of episodic memory, around which mathematical concepts are reconstructed, linked to the stories in a strong way. This is memorised in depth.
We have told the students about some particular and central episodes in the history of mathematics, which are rich in didactic clues for both primary and secondary schools, and we have revealed some interesting aspects of the lives of some authors who are mostly known for their theorems or theories.
This was mainly to give mathematics a more appealing and attractive look even to those pupils who were less motivated to get involved and learn this discipline.
We therefore thought of searching among the INVALSI questions for concepts, theorems or ideas related to the various authors whose stories they had tackled.
The questions were researched together with the students, under our supervision and with our advice, using the Gestinv 3.0 platform (www.gestinv.it), among the tests released for grades 5, 8, 10 and 13.
The aim was to be able to address the topics presented from the mathematical stories to the various and areas of mathematics, to test the skills acquired by the students, trying to get them to acquire the knowledge and skills necessary to better learn what they have studied.
Another important aspect was that, while dealing with the stories, the students also had to try to answer them correctly and possibly argue their answers.
This activity was also interesting in making the INVALSI tests more real and contextualised, bringing them closer to the students, to get them used to tackling them, in a natural way, without resorting to exercises in preparation for the tests.

***Keywords:*** History, mathematical tales, INVALSI tests, learning

# The INVALSI data from fifth-grade primary school as a tool for evaluating a methodology based on games for teaching mathematics

**Valentina Vaccaro – Maria Francesca Ambrogio**

Abstract: Studies have identified a double paradox that characterizes the teaching-learning process. On one hand, for students to learn, they must accept breaking the instructional relationship and establish a direct relationship with knowledge (paradox of the act of learning). On the other hand, if teachers explicitly declare what they want to achieve, they will not be able to reach it (paradox of the act of teaching). In the Theory of Situations (Brousseau, 1998), overcoming this paradox is achieved through the establishment of an a-didactic situation that creates a specific condition of the instructional relationship between teacher, student, and environment, inherent to knowledge. A situation is considered a-didactic with respect to a specific knowledge if it provides the conditions for the student to experience it independently of instructional instances. The actions performed by the student, the answers provided, and the arguments presented must, therefore, depend on their (not entirely explicit) relationship with knowledge, understood as knowledge to be acquired or used to provide answers.

The action research involving a mathematics education researcher, a teacher, and her students from first to fifth grade of primary school was initiated in the school year 2017-2018. Based on the presented studies, the researcher and the teacher rethought the planning of didatics activities. The teacher, starting from the study of the Mathematics Framework proposed by INVALSI (INVALSI, 2018) and the processes required by the items of grade 2 and grade 5, devised, selected, and built strategy and logic games in continuous collaboration with the researcher, to stimulate and develop in her students the knowledge, skills, and competencies suggested by the National Guidelines (MIUR, 2012). The use of strategy and logic games in the classroom offers the opportunity to observe and evaluate the students' learning process. The use of games transforms the classroom into a laboratory where "devolution necessarily takes place because the teacher stimulates and disappears, leaving the child with great responsibility. Their involvement leads them to experience firsthand, taking risks" (D'Amore, 2005). It is necessary to carefully choose the games to be used, ensuring they meet some established criteria from practice, such as being quick and accessible to all and, of course, requiring logical operations (Bolondi & Vaccaro, 2017).

The research was born with the aim of integrating an innovative methodology into daily teaching, which is often only sporadically used within the teaching-learning process: the use of mathematical games as a tool to acquire lasting competencies. Before applying this methodology, it was necessary to develop an awareness of using games, which are often seen as distant from the school context, as a learning tool. The idea of introducing games from the beginning of the journey with a first-grade class has proven to be extremely effective in fully integrating them into instructional activities. The success of this approach led to continued work in this direction in subsequent classes until the end of the educational cycle. Materials useful for recording key elements of the learning process were adapted, rethought, and created to evaluate the learning process (Vaccaro & Ambrogio, 2021). The research was initiated before the pandemic, and despite the difficulties associated with distance learning, the methodology was maintained throughout the duration of the cycle.

In this study, we use the INVALSI data from fifth-grade primary school to assess the benefits derived from this methodology and introduce improvements. The analysis focuses on the methodological innovation caused by the introduction of strategy and logic games throughout all stages of the teaching-learning process, rethought from an educational perspective. We relate the methodologies and tools used in the educational action to the results obtained by the students involved in the study in the national INVALSI assessment of grade 5 in the academic year 2021-2022.

From the data analysis, a clear improvement was observed compared to the results from the school year 2020-2021 for the fifth-grade students. The experimental section achieved a score of 217.3 on the national scale in the school year 2021-2022, representing an increase of 26.5 points compared to the school's average score in the previous year. The increase is also significant when compared to the school's average score obtained by students of the same cohort in the last assessment before the pandemic (school year 2018-2019); the increase is indeed 22.3 points. Such comparisons are possible since the scores expressed on a scale of 200 for each subject from the academic year 2018-2019 are directly comparable to those in subsequent assessments.

In Table 1, we compare the results obtained by the students involved in the study in the national INVALSI assessment of grade 5 in the school year 2021-2022 with those obtained by the same students in grade 2 in

the academic year 2018-2019 (Vaccaro & Ambrogio, 2021). In Table 2, we compare the increase obtained by comparing the results of the students involved in the research with those obtained by grade 5 students in previous years. We decided to compare the increases to reduce the noise created by the increase in grade level and the measures used during the pandemic, which undoubtedly influenced the performance of students across Italy.

Tab. 1: Diachronic-longitudinal comparison of the percentages obtained in grade 2 and grade 5 divided by areas and dimensions with reference to the school average.

| | Grade 2 – 2018-2019 | | Grado 5 – 2021-2022 | |
|---|---|---|---|---|
| | Class | School | Class | School |
| Numbers | 65,8 | 63,4 | 68,6 | 55,2 |
| Space and Figures | 67,5 | 66,2 | 59,9 | 52,6 |
| Data and Previsions | 61,7 | 56,3 | 58,2 | 54,6 |
| Relations and Functions | --- | --- | 49,9 | 46,7 |
| Know | 73,5 | 69,5 | 61,2 | 50,8 |
| Solve problems | 56,3 | 53,6 | 58,5 | 54,4 |
| Argue | 56,8 | 60,7 | 76,5 | 65,4 |

Tab. 2: Comparison of increments between grade 2 and grade 5.

| | Class-School Increase Grade 2 | Class-School Increase Grade 5 | Increments' difference between grades |
|---|---|---|---|
| Numeri | 2,4 | 13,4 | 11 |
| Spazio e Figure | 1,3 | 7,3 | 6 |
| Dati e previsioni | 5,4 | 3,6 | -1,8 |
| Relazioni e Funzioni | --- | 3,2 | 3,2 |
| Conoscere | 4 | 10,4 | 6,4 |
| Risolvere problemi | 2,7 | 4,1 | 1,4 |
| Argomentare | -3,9 | 11,1 | 15 |

The obtained results are encouraging, and the increase in argumentation shows how innovative methodologies and assessment tools can impact the acquisition of high-level competencies. An analysis of individual items will be necessary to investigate the mechanisms used by students to maintain control over the entire process of verifying acquired knowledge, skills, and competencies.
The external evaluation, made possible thanks to the national assessments conducted by the INVALSI Institute, allowed analyzing the effects of the experimentation without the interference of the teacher-researcher in the verification and assessment process. Full access to the INVALSI tests for the second and fifth-grade classes, the framework, and all materials and data provided by the Institute, facilitated a thorough analysis of the similarities between the games proposed during the school journey and the questions in the assessments. The amount of data collected, both aggregated and related to individual items, was crucial for a critical review of the errors made by students. Particular attention will be given to the

analogies and differences between the types of games proposed and the items in the standardized test, considering the methods of verification and evaluation used during the learning process.

***Keywords*:** mathematical games, mathematics education, large scale assessment

## Serious Games and INVALSI items: an idiographic study of an innovative mathematics teaching approach

**Roberto Capone – Eleonora Faggiano**

The focus of this contribution is a research concerning the use of digital serious games in secondary schools whose teaching content is based on INVALSI Mathematics questions. The study, which is idiographic in nature, aims to analyse individual case studies in which the latest INVALSI results showed low participation and achievements in mathematics below national standards. This research is part of a larger international project, MaTIn4MER (Methodological and Technological Innovations For Math Education Resources), with which the interchange between the world of research, the world of business and the world of education is intended to be put into practice. The project aims to contribute to ongoing international research on the integration of new technologies in mathematics education, investigating how the interactions between student, teacher and knowledge change with the introduction of digital resources in teaching practices. The first phase of the project sought to understand how the profile of the 21st-century mathematics teacher is perceived in terms of technological, didactic and methodological competencies.
We decided to focus on the teachers' idea of the use of gamification for mathematics didactics and their view of the possible integration between innovative teaching practices, research in mathematics education, and corporate interest in producing games for educational purposes. The analysis of the teachers' training needs, carried out through questionnaires and individual interviews, was structured with reference to the theoretical frameworks of Technological Pedagogical Content Knowledge (TPCK) and Metadidactic Transposition. The aim of the investigation was to identify the aspects to be leveraged on in order to personally involve some teachers, not only in experimenting but also in designing serious games for the teaching and learning of mathematics. The research hypothesis is that the involvement of teachers in the design of serious games constitutes a fundamental step in their teacher training in relation, in particular, to the integration of innovative technologies in teaching practices. We then chose to focus on contents closely connected to the INVALSI tests with a twofold purpose: on the one hand to show how the use of serious games can foster the acquisition of knowledge and key competencies in mathematics, and on the other hand how such acquisition should not be based on a sterile "training" in the tests. We assume, in fact, that the involvement of students in problem-solving activities through digital serious games, which can also promote a good emotional disposition towards the discipline, can improve teaching and learning processes in mathematics.
In the second phase of the project, therefore, groups of teachers in Italy were involved in the design of a digital serious game to be developed in collaboration with the researchers and the company. An initial version of the game was implemented by the company and in the third phase of the project it was tested in the schools and classes of the teachers involved in the design. After involving the students in the use of the game, the improvements in terms of the acquisition of mathematical skills, and the parameters of motivation, engagement and participation were analysed.
The serious game under investigation is based on maps, the storytelling of which takes us back to ancient Greece, to Aeneas' journey from Troy to Rome. In the various stages, the player has to overcome obstacles of all kinds that are reproduced through the resolution of some INVALSI items. Examples of the use of INVALSI mathematics items for the implementation of some scenes of the game will be shown.
The study was successful from the point of view of both mathematics teaching and learning. This was detected through a test administered to the students and a questionnaire administered to both students and teachers. The results of the final test and the questionnaire seem to show that serious games encourage a better emotional disposition towards mathematics and improve the teaching and learning process of the discipline. In addition, serious games seem to enable teachers to reflect on their teaching methods in order to foster more effective teaching of mathematics, and involvement in the design thus seems to contribute positively to their professional development.

***Keywords:*** serious game, gamification, interdisciplinarity, large-scale assessment

## The performance of students attending the Liceo Matematico in INVALSI tests

**Francesca Coppa – Alessandro Gambini – Stefania Gubbiotti – Davide Passaro**

Introduction. The project called "*Liceo Matematico*" (LM) was established in the academic year 2015-16 at the Department of Mathematics of the University of Salerno and quickly spread throughout Italy. The *Liceo Matematico* is characterized by the introduction of at least one additional hour of mathematics per week, conducted with a laboratory approach. The additional hours are dedicated to interdisciplinary activities designed and implemented in close collaboration between schools and universities. Currently, there are 26 universities involved with over 150 schools participating in the project.
For a more detailed description, please refer to *Bernardi et al. 2022* and the website https://www.liceomatematico.it/.
The Italian Mathematical Union (UMI) established the UMI Group on LM in 2020, with more than 120 UMI members participating. The presence of UMI strengthens the national dimension of the project.
Objective and research hypothesis. In 2022, a research group was formed consisting of statisticians, experts in mathematics education, and secondary school teachers. The aim was to analyze the results of the INVALSI tests of students enrolled in LM sections and compare them with the results of students enrolled in the same schools but in non-LM sections.
The members of the research group are: Claudio Bernardi, Alessandro Gambini, Stefania Gubbiotti, Francesco Saverio Tortoriello, Francesca Coppa, Federica Ferretti, Maria Flavia Mammana, Davide Passaro, Francesca Tovena.
For the schools that participated first in the LM project, a five-year cycle has been completed. Therefore, it was decided to analyze the results of the INVALSI tests in Italian language and Mathematics, limited to these schools.
In particular, the research aims to answer the following questions:

- Is it true that LM students have better average results than their colleagues in the same schools who do not participate in the project?
- Is it true that LM students show improvement both in the transition from lower secondary to upper secondary education and throughout upper secondary education?

Note that a positive answer to the first question could be expected: it is likely that students who participate in the LM project, by accepting to do extra-curricular hours, tend to have, on average, good results. The second question is much more delicate.
Therefore, data regarding the results of grade 8 (referring to lower secondary school) and grades 10 and 13 (referring to the second and fifth years of upper secondary school) were requested from INVALSI in order to longitudinally analyze the performance of students participating in the LM project.
Data used. The research is based on an exploratory and knowledge-based investigation, in which the reference population consists of all secondary school students belonging to the cohort that took the INVALSI tests in grade 10 in 2019 and grade 13 in 2022, and for whom the result of the grade 8 test (taken in 2017) was also available. The data provided by INVALSI were integrated with information obtained directly from school principals regarding the participation in the LM by individual sections. The study involved a total of 2,588 students from 11 schools in 5 different provinces (RM, CT, CE, AV, LT).
Method. Although the analyzed data belong to a reasoned (non-probabilistic) sample, and therefore the analysis is descriptive in nature, the effects of participating in the LM can be clearly identified based on the Warm's mean Weighted Likelihood Estimates (WLE) scores obtained in the Italian language and Mathematics tests (according to the Rasch model). The division of these scores into two classes (LM and non-LM students) was considered, and the trend over the three grades under consideration was analyzed.
It is believed that the obtained sample, although partial, allows for a good approximation of the characteristics of the reference population. In any case, to compensate for any distortions, all sections of the schools considered in the specific cohort of interest were included in the study, using the non-LM classes as the control group.
Results. The scores obtained in the INVALSI tests in Italian language and Mathematics show encouraging results in favor of the LM. In particular, the research has shown that, despite starting with relatively uniform

results at the end of lower secondary school, there is a clear difference in the WLE scores favoring LM students in the subsequent grades.
The data analysis shows that the effect attributed to participating in the LM project on the mathematics score is positive: all other factors being equal, an LM student has an advantage over a non-LM student. The same applies to the score in Italian language, which is consistent with the interdisciplinary nature of the LM.

***Keywords:*** teaching practices, learning, educational research, mathematics education

# THEME 8. LEARNING TO LEARN IN SCHOOLS, IN ADULT EDUCATION AND IN VOCATIONAL EDUCATION AND TRAINING

**ORGANIZER: INVALSI**

**COORDINATOR: CRISTINA STRINGHER**

**25TH NOVEMBER: 9.30 A.M. – 11.30 A.M. {ROOM 3 – RESEARCH SESSION 14}**

## Educational planning and planning to intervene in the ability to learn to learn

**Franca Da Re**

The European Recommendation of 2018 on key competences for lifelong learning, as well as the one on the analogous subject which had preceded it, of 2006, indicated, among the essential competences that every citizen should achieve and develop throughout life, the Personal and social competence and the ability to learn to learn.
In the 2018 version, this competence brings together both personal and social skills and those of learning to learn, considering them all as a single asset of the person.
The research conducted by INVALSI and coordinated by Cristina Stringher among teachers from six different countries has shown that the ability to learn to learn does not depend only on components relating to self-regulation, metacognition, the ability to select, understand, organize and reprocess information, but also to dimensions related to motivation, self-confidence, interest, significance attributed to materials and experiences, attribution styles of learners, as well as their socio-cultural background.
This would justify the choice made in the Recommendation to consider the ability to learn to learn as related to other personal and social skills.
Object and research hypothesis. The research object is the didactic planning and design related to the development of personal and social competence and the ability to learn to learn, in the Indications, in the national guidelines, in the school curricula and in the teaching practices. Both National Guidelines and guidelines and school curricula rarely describe learning outcomes related to the competence in question, such as citizenship competence and entrepreneurial competence, because both national guidelines and school curricula generally address more of the skills of a more cultural nature, which have direct references to the study disciplines. The research hypothesis is that a setting of school curricula on European key competences and a consequent precise description in terms of knowledge, skills, specific competences of the learning outcomes relating not only to the five competences of a more cultural nature, but also to the other three, it would make it possible to plan more rigorously effective educational paths for the development of personal, social and learning to learn, citizenship and entrepreneurial skills.
Data used and method. The method used is observation in the field, over many years of analysis of national documents, school curricula and teaching practices in both first and second cycle schools, as well as teacher training paths and interlocutions direct with them in the field of training, didactic consultancy and external evaluation of schools.
The data used derives from the educational offer plans of the institutes and from the school curricula. The generally observed results show that personal, social and learning-to-learn skills, as well as entrepreneurial skills are rarely punctually defined and described in curricula. As a result, there is no real agreement in teachers' colleges and class councils on the nature of these skills, nor on how to help pupils develop them. In the curricula and programming of class councils and of individual teachers observed in the field, these skills are generally stated without following protocols to pursue them, which instead happens more punctually for the cultural skills that refer to the disciplines. The most virtuous teaching practices generally take charge, in relation to the ability to learn to learn, of study skills, self-regulation, the ability to orient oneself in texts and to report information, but these skills are almost never related to motivation, interest, the significance of information, the social dimensions of learning. In other cases, teaching practices limit themselves to providing students with information on some ways of dealing with study texts, organizing information, taking notes, without even addressing the correlated topics overall, such as strategies and cognitive styles and learning and metacognition. Confirmation of this is also given by the difficulty encountered by schools in compiling the Self-assessment reports in relation to the results area relating to key competences, which does not have indicators provided by the Ministry, unlike the other areas. Schools found it extremely difficult to find indicators both in terms of outcomes and related processes. The

interviews conducted among teachers during the external evaluations of schools showed that generally the dimensions connected to non-cultural competences are not systematically examined in the curricula and Class Councils and that the actions are more often carried out by individual teachers, each with their own techniques and strategies not shared collectively.
Conclusions. In conclusion, it is believed that a precise description of personal and social skills and of the ability to learn to learn starting from the content of the European Recommendation in school curricula which is subject to intersubjective agreement between teachers would help to plan precise specific teaching paths and at the same time within the different disciplines to help students develop these skills and to allow them to be observed and evaluated.

***Keywords:*** learning to learn, personal and social competence, ability to learn to learn

## The assessment of personal, social and learning to learn competence. The testing of a formative assessment dispositive

**Arduino Salatin – Andrea Giacomantonio – Fabrizio Giovannini – Marta Santanicchia**

Introduction. From September 2018 to March 2023, the Italian National Institute for Public Policy Analysis (INAPP) promoted two research projects. Their first aim was the development of an authentic assessment dispositive for personal, social and learning to learn competence and three other key competences (citizenship, entrepreneurship, cultural awareness and expression competences) addressed to vocational education and training (VET) learners. The second objective was to set guidelines and other resources to support a full integration of key competences, their assessment dispositive as well as its axiological, methodological and organizational logic in VET curricula.
In the first research, the INAPP collaborated with a temporary grouping of enterprises (RTI) composed of the Salesian International Higher Institute of Educational Research (ISRE, lead partner), Italia Forma (IF) and Central Training School (SCF); in the second one, the RTI was composed of ISRE, IF, SCF, Italian Center for Salesian Women's Works – Vocational Education and Training (CIOFS-FP), National Center for Salesian Works – Ongoing Vocational Training (CNOS-Fap), Intellera Consulting and PTSCLAS.
We are going to outline some of the essential features of the dispositive and the fundamental steps of its design.
The objective and the conceptual terminology apparatus. The objective has already been stated. In brief, it consists in the design and testing of an assessment dispositive for 4 key competences, including personal, social and learning to learn competence. At its core, a key competence has been regarded as an articulated and partially coordinated set of enduring internal dispositions (Bourdieu, 1980, 1997). These are habits; they seem to constitute an individual's character through repeated practice (Pellerey et al., 2013). Furthermore, a key competence likely manifests itself in performance, constantly intertwining with technical-professional competences and basic symbolic skills. In a transactional perspective (Dewey, Bentley, 1949), the distinction between these three factors - key competences, technical-professional competences and basic symbolic skills - does not have any ontological, but a functional value for the research and assessment. Consistent with the same perspective, we hypothesize that each key competence is constituted by the transaction of four dimensions: intellectual, methodological, strategic and social dimensions. The four dimensions are supposed to mobilise resources belonging to knowledge cores, that are relevant on epistemological, axiological and existential levels (Giovannini, Santanicchia, 2023). In the era of the hegemony of subjective reason, this conjecture represents an attempt to find a balance between the latter and objective reason (Horkheimer, 1947). From this articulated set of hypotheses, we can infer that the dispositive unfolds its potential when used as a formative assessment tool (Hattie, Clarke, 2019).
Research design, population and instruments. The design of the research is configured as a "Teacher Professional Development Research" (Asquini, 2018): a growing number of teachers, in fact, collaborated in the design and revision of the dispositive. The two researches were developed through 6 work phases. Six administrations were carried out involving convenience samples of varying sizes and properties. In the first administration (November-December 2018), approximately 245 students from 13 classes of 8 Vocational Training Centres (VETCs) in northern and central Italy participated; in the third one, prior to the outbreak of the pandemic, 1,396 students from 65 classes of 42 VETCs distributed throughout the country

took part. The articulated nature of these 4 competences required a mixed-method assessment design (Trinchero, Robasto, 2019). Consistent with the authentic character of the assessment (Wiggins, 1993), the main tool was the reality-based assessment. No new reality-based assessment tools were proposed to VETCs; rather, instructions, suggestions and exemplary cases were provided to modify those already envisaged. We also designed the relevant observation rubrics (Castoldi, 2018; Tenam-Zemach, Flynn, 2015). In these rubrics, the dimensions of each competence were operationalised by identifying the topics and the indicators. The indicators were divided in two cluster: specific to each competence or transversal to the 4 competences. The dispositive also provided for the administration of highly structured assessment tools that were changed in the course of the researches. In the last phase, the following instruments were administered: a reading comprehension test, the Learning Strategies Questionnaire (Pellerey, 1996), a revision of the citizenship questionnaire used in the 2016 ICCS IEA survey (Schulz et al., 2018a, 2018b), a problem-solving test and a student questionnaire. While renouncing stimulus invariance, therefore, it was assumed that valid and reliable information on key competences could be gathered.
Results. Many are the research findings. We present only three of them briefly.
The first concerns the informative potential of the reality-based assessment about personal, social and learning-to-learn competence. In the second administration of 2019 (September-November), the reality-based assessment was carried out by 215 pupils belonging to 13 classes in as many CFPs. They achieved an average of 60.6% of the total. The informative potential of the dispositive thus appears to be medium to high. The second result concerns the relationship between transversal and specific indicators: the personal, social, learning to learn competence appears more transversal than the other three key competences. The third result concerns the perception of the assessment dispositive by the teachers involved in the second project - surveyed in 2023 through the focus group. Let us summarise only the main statements:

1. the dispositive seems sufficiently sustainable;
2. it appears as a formative assessment tool;
3. teachers appreciate the distinction between specific and transversal indicators;
4. teachers note the relationship between the assessment dispositive and the teaching and educational tools used to facilitate the development of key competences.

In a broader perspective, the research offered a stimulus on the possibility of developing and assessing key competences - whose strategic value is strongly linked to their metacognitive character. The research also identified at least two further areas of investigation - also through two peer learning activities, in which researchers and socio-institutional stakeholders took part. The first area concerns the relationship between system and formative assessment; the second one is validation and the social recognition of learning outcomes related to key competences in a lifelong learning perspective.

***Keywords:*** formative assessment, key competences, reality-based assessment, INAPP

## Promoting learning achievements of prospective teachers through gamified cognitive activation

**Andrea Tinterri – Anna Dipace**

Introduction. The "new normal" scenario of education after the COVID-19 pandemic poses a great challenge of methodological innovation for educational institutions. In higher education, for instance, the problem of student dropout in the first year has been exacerbated in the context of the deregulation of teaching and learning in physical distance due to the pandemic (Bernardo, Castro-Lopez and Diaz Mujica, 2022). To face these challenges, educational institutions at all levels are under pressure to adapt their teaching procedures, integrate technologies in teaching and update the skills of teachers (Eradze et al., 2023). Cognitive activation (CA) describes the mental stimulation of learners to engage in deeper mental investigation of the subject matter within the learning context (Groß-Mlynek et al., 2022). Such reflective practice is especially relevant in the current educational scenario, as today's students, and tomorrow's professionals, need to be responsible for their own learning and lifelong growth (UNESCO, 2020). Digital technologies can facilitate its use and enhance its effects and at the same time stimulate an increase in so-called feedback literacy (Trinchero, 2018). At the University of Foggia, digital tools such as Kahoot! (Bienvenido Huertas, Rubio Bellido & León Muñoz, 2023) have been used since academic year 2020/21 for formative assessment and

student engagement in different teaching and learning contexts, from remote teaching during the pandemic, to hybrid teaching during the post-pandemic transition, and in current blended and traditional forms. In those scenarios, Kahoot! was used to create quiz competitions among students at the beginning of each lessons using multiple choice questions (MCQs) as a tool for formative assessment and cognitive activation (Dipace & Tinterri, 2022). This contribution presents a pilot study aimed at assessing the impact of digital CA practices using MCQs on the perceptions, learning results and academic behavior of prospective teachers at the University of Foggia.

Description and goals of the study. The study had a semi-experimental design. The research activity involved students attending two courses, one from the Education and Formation Sciences degree course and the other from the five-year degree course in Primary Education Sciences. The sample was of 83 students for the first course and 132 in the second one. At the beginning of the course, students were divided in three groups. "Makers", after an initial asynchronous training on MCQs, were tasked with designing each week six MCQ: those questions would then be graded by the course tutors and a selection of them was used for CA. The second group, named "Users", participated in the CA activity at the start of each lesson but did not contribute to creating MCQs. Finally, "non-participating (NP)" where those students that either did not follow courses in presence or choose not to engage in CA activities. Each week, the MCQs used for CA where also uploaded on the course platform to allow NP students to self-assess their learning. The final examination at the end of the course was the same for all groups. Examination of the training impact was defined according to Kirkpatrick's four levels (1994) and concerned the levels of reaction (students' perception concerning the satisfaction with the training, engagement with the learning material, and perceived relevance of the activity) learning (students' performance at the final examination) and behavior (student's decision to attend exams). The hypothesis was that engaging in CA activities – both simple participation and creation of MCQs - could positively impact learning at different levels. Thus, the research questions were defined as follows:

RQ1: How did the role played by students impact their perceptions concerning the CA activity?

RQ2: Did the role played by students differently impact their learning outcomes?

RQ3: Did the role played by students differently impact their decision to attend exams?

Material and Methods. To investigate student perception, two questionnaires were administered at the end of the course to participants (both Users and Makers) and NP students. The questionnaires were an adaptation of the one by Brazeal et al. (2016), based on the 5 criteria of formative assessment identified by Black & William (2009). They included both Likert items, using a 1 to 5 scale, and open-ended questions and were administered through Google Form to students at the beginning of second semester activities. The instrument was created to investigate students' perceptions with respect to the use of cognitive activation for the following dimensions of the teaching-learning process in academic instructional settings: motivation, metacognition, study method, sense of self-efficacy, and sociality. To investigate students' learning achievement, we took into consideration final examination grades. Finally, we examined students' participation to exam sessions as a measure for students' behavior. Data was collected through the University platform (ESSE3) and Microsoft Excel. Statistical analysis was performed using Jamovi software.

Preliminary results. Analysis of the study results is still in progress; the following are preliminary results concerning students' reaction to the activity (RQ1). In total, 73 students who did not participate in the research activity (NP) and 100 among participants (both users and makers) answered the post-hoc questionnaire. Participating (P) and non-participating (NP) students did not differ significantly in terms of sex, residence (most of them are commuters in both groups), and previous academic achievement; however, among NP students there were more adult students (18/73 were >30 years old, compared to 12/100, $X2 = 14.9$, $p=0.002$). Among NP students, 22/73 did not frequent classes at all, whereas the other frequented at least in part to in-presence activities. Among P students, 75 of them would have frequented classes anyways, while for 19 students CA played a role in their participation. When asked why they did not participate in the CA activity, NP students cite lack of time (2.88 out of 5 on average) rather than not being interested (1.08/5) or the activity being too complex (1.04/5) and they considered the asynchronous activities very useful for their preparation (4.48/5). In terms of expected learning achievements, P and NP students had similar expectations concerning the final grade ($X2 = 1.19$, $p=0.880$), but there was a significant difference in the rapport between expected and actual results (Mann-Whitney $U = 2104$, $p<0.001$) with NP students achieving, on average, worse results than expected and P students results in line with their expectations. Obviously, the significance of these observation depends on the rest of the undergoing analysis; however, data suggest that gamified CA activity could play a role in motivating students to participate in presence and help them obtain a more realistic assessment of their preparation.

***Keywords:*** cognitive activation, formative assessment, technology-enhanced learning

## The contribution of learning orientations to school success and well-being. Results of studies conducted with students and teachers

**Vettori Giulia – Bigozzi Lucia – Giuliana Pinto**

Introduction. Study difficulties and school dropouts are crucial indicators for the efficiency of educational institutions and the well-being of students and teachers. There is a pressing need to identify the factors that contribute to sustain learning and learning to learn (Ajello, 2018; Stringher et al, 2021) considering the low levels of student learning performance in many OECD countries (OECD PISA 2016). Different motives may sustain students to persist at school and using specific strategies (Biggs, 1976; Entwistle, 1977; Marton & Säljö, 1976). Learning conceptions, intended as what learning means and how it occurs from the point of view of learners, are an important psychological construct for teaching and learning (Vermunt & Donche, 2017) given their influence on study approach and school outcomes. Representations that are born and develop in the intersection between the learner's self and the knowledge process including beliefs about the cognitive and metacognitive (e.g. memorisation, concentration, self-monitoring), sociocultural (e.g. learning in groups, learning as individual work), emotional-affective and motivational (e.g. anxiety, satisfaction, confidence in one's own abilities) dimensions of learning. In this presentation, the main findings of the research line carried on by the Florence research group about students and teachers' conceptions of learning and the influence that students' learning orientations exert on students' well-being and motivation will be presented.
Objective and hypothesis. Our research line aims to (1) identify learning conceptions in secondary school students and pre-service teachers by adopting a person-oriented approach, (2) investigate the pattern of relations between academic performances, conceptions of learning, and university pathways (on-time, delayed, and dropout) with a cohort retrospective study, and (3) assess the multidimensionality of students' learning orientations through the use of the self-report questionnaire "LO-COMPASS: Learning Orientation-Cognition Metacognition Participation Assessment" and their association with school achievements.
Date used. The target population is secondary school and university students of more cohort study samples and more than two hundred pre-service teachers. This large population allows us to shed light on significant school transitions.
Method. Conceptions of learning in pre-service teachers, secondary school and university students were investigated by using the validated self-report instrument "Learning Conceptions Questionnaire" (LCQ; Perez-Tello et al., 2005). Furthermore, students school performances were collected, and learning orientations assessed trough the self-report questionnaire "LO-COMPASS: Learning Orientation-Cognition Metacognition Participation Assessment". Factor analysis, cluster analysis and regression were conducted.
Results. Findings showed the existence of superficial and deep learning conceptions in secondary and university students (Pinto et al., 2018; Vezzani et al. 2018) and pre-service teachers (Vettori, et al., 2019), suggesting different mindsets and motivational dispositions to learning. The results from our studies have revealed the predictive power of the conceptions held by secondary school students with respect to academic performance school performance (Pinto, Bigozzi, Vettori, & Vezzani, 2018). Alongside the direct role assumed by conceptions in predicting performance, the results support in recognising conceptions a mediating role in fostering the relationship between self-regulation and academic success (Vettori, Vezzani, Bigozzi, & Pinto, 2018). Furthermore, retrospectively dropout university groups were significantly more likely to hold a superficial conception of "Learning as a reduction of deficit knowledge and passive-receptive role" than on-time participants did (Vettori et al., 2021). Finally, LO-COMPASS (Vezzani et al, 2023; Vettori et al., 2022; Vettori et al., 2020) allowed to trace two profile of learning orientations significantly related to school performance (students Dragged by the current vs. students At the helm). The research findings will be discussed by paying attention to practical and research implications.

***Keywords:*** learning to learn, conceptions of learning, learning orientation, self-regulation, participation, assessment.

# THEME 1. A DIFFICULT EDUCATION: CHALLENGES OF THE SCHOOL SYSTEM TO ENSURE SOCIAL INTEGRATION, SKILLS/KNOWLEDGE AND SOCIAL MOBILITY

**ORGANIZER: INVALSI – ESPANET**

**COORDINATOR: EMMANUELE PAVOLINI**

**25TH NOVEMBER: 2.00 P.M. – 3.30 P.M. {ROOM 4 – RESEARCH SESSION 15}**

---

## The uphill battle: The amplifying effects of negative trends in test scores, COVID-19 school closures and teacher shortages

**Letizia Gambi – Kristof De Witte**

Introduction & research hypothesis. Despite the pivotal role of education into fostering long-run economic growth and welfare (Hanushek & Woessmann, 2015), multiple (interrelated) challenges put education systems under pressure. Many Western education systems face a declining trend in education outcomes, high early school leaving rates, increasing teacher shortages, low student motivation, and progressing inequality in test scores. Furthermore, both learning and teaching have come under additional strain since the COVID-19 pandemic, with growing evidence showing that school closures exacerbated some pre-existing weaknesses in education systems (Betthäuser et al, 2023; Moscoviz & Evans, 2022; Patrinos et al., 2022). To the extent that these challenges reinforce each other, the decrease in test scores will accelerate over time (i.e. our research hypothesis). This paper investigates whether schools are able to reverse this uphill battle. In particular, this paper examines the evolution of standardised test scores in multiple subjects at the end of primary education in Flanders in the aftermath of the COVID-19 school closures and the increasing teacher shortages (i.e. the absence of teachers due to unfilled vacancies). We used a rich dataset with administrative data and standardised test scores from private Catholic schools representing the largest primary education provider in the Flemish region of Belgium. Using the exact same curriculum-based test since 2019, we show how learning deficits accumulate over time due to existing negative trends in a school system and the intensifying effects of school closures.

Data. The dataset for the present study consists of panel data at the school level (2015-2022) from the Flemish community of Belgium. The dataset combines: (1) data on standardised tests annually administered in the last year of primary school (grade 6) by the network of Flemish Catholic schools in Belgium (Katholiek Onderwijs Vlaanderen[1]), available between 2019 and 2022 in the exact same test version[2]; (2) administrative data provided by the Flemish Ministry of Education comprising several dimensions of the school environment; (3) administrative data on summer schools, available for the years 2020 and 2021; and (4) administrative data on teacher shortages at the school level, available for the years 2021 and 2022. Our analysis focuses on two main outcome variables: standardised test scores and inequality measures. In addition to the test and inequality data, the analysis gradually controls for multiple dimensions of school environment that earlier evidence has found as being associated with student performance: teacher characteristics, school district composition, and characteristics of both grade 6 and the entire school.

Methods. We estimated the evolution in school outcomes and inequality measures (both at the level of the school $s$) by means of a fixed effects panel regression, using the following model specification:

$$y_{s,t} = \alpha_s + \beta(Change\text{-}in\text{-}2020)_{s,t} + \gamma(Change\text{-}in\text{-}2021)_{s,t} + \delta(Change\text{-}in\text{-}2022)_{s,t} + \theta X_{s,t} + \epsilon_{s,t} \quad 1)$$

for schools $s$ = 1,2,..., N and time $t$ = 2019, 2020,..., 2022.[11] The analysed outcome variable (i.e. test score in each subject and inequality measures) of school $s$ at time $t$ is denoted by $y_{s,t}$, while $\alpha_s$ denote the school fixed effects. The school fixed effects capture the time-invariant unobserved heterogeneity due to, for example, differences in school policy, school management quality, school location and school reputation. As time elapses, it becomes more difficult to distinguish the COVID- 19 impact from the trend in education outcomes, and vice versa. Therefore, in the present analysis, we do not aim to make this distinction and focus on the total change in learning outcomes since 2019. The estimated coefficients capture both the effects from the COVID-19 school closures and the changing test scores over time. In model (1), the outcome $y_{s,t}$ is regressed on three dummies: *Change-in*-2020, *Change-in*-2021 and *Change-in*-2022. Each dummy takes on the value of 1 in the year of reference (2020, 2021 and 2022, respectively), and 0 otherwise. Thus, the coefficients $\beta$, $\gamma$ and $\delta$ represent the main coefficients of interest, reflecting the yearly change in test scores

(and inequality measure) relative to pre-pandemic levels. Finally, in the regression we cluster the standard errors $\epsilon_{s,t}$ at the school level, accounting for serial autocorrelation in the error terms.
Results. The results indicate a significant decrease in test outcomes since the start of the pandemic, although further evidence is provided of the presence of a negative trend in the test scores evolution already before 2020. Comparing the standardised test data at school level before and after the school closures suggests that in three years' time the grade 6 education outcomes decreased, on average, by -0.47 SD in the Dutch language, -0.31 SD in the French foreign language,
-0.12 SD in math, -0.11 SD in science and stayed constant in social sciences. At student level, the estimates present the same significance and direction of the sign, although with lower magnitude (as the individual variation in test scores is larger than the variation across schools). The learning deficits seem to accelerate over time, partly driven by the weakening performance of the best students, and exacerbated in schools with high shares of teacher shortages. Indeed, compared to pre-pandemic test scores, the test scores of the best-performing students in a school decrease their test scores most in recent years as compared to pre-pandemic levels. As for remedial actions in the aftermath of the pandemic, summer schools seem to mitigate part of the learning deficits of those schools that undertook a summer program during both summer 2020 and 2021. However, the increasing number of unfilled teacher vacancies work in the opposite direction. We show that teacher shortages could exacerbate the educational outcomes of students. Specifically, we find that one percentage point increase in unfilled vacancies (as percent of total teachers as FTE) is associated with a decrease of -0.04 SD the Dutch language, and -0.05 SD in math proficiency. Furthermore, we find that schools with higher shares of teachers with less than 10 years of experience show larger average attainment deficits while the opposite direction can be observed in schools with higher shares of teachers with more than 26 years of experience. Finally, we observe that the within-school inequality in test scores has not been reduced since the start of the pandemic, while the increase in between school inequality in test scores observed in 2020 has slowed its pace in both 2021 and 2022.
Overall, the results illustrate that learning deficits and declining trends in an education system are hard to counter. If students do not sufficiently master the curriculum from earlier grade years, teachers have to devote time to lecture or rehearse the curriculum from earlier grade years (and spend less time on new learning material), resulting in accumulating delay in learning. Moreover, as classes are becoming increasingly heterogeneous in terms of pre-knowledge, teachers tend to focus more on students' weaknesses rather than students' strengths (Siegle & Powell, 2004) such that everyone reaches at least the basic levels. In the Flemish education system, this is even an obligation by law as the curriculum goals present minimum targets that have to be reached by a majority of the class. However, the focus on the weakest students distracts instruction time from the strongest students, which are left behind. Consequently, their test scores decline, resulting in a further slide of (average) test scores.
[1] The Flemish education system is highly decentralized. It is organised by three educational providers: community schools (directly provided by the Flemish Government), grant-aided provincial/municipal public schools, and grant-aided free schools (publicly funded but privately-run). The network of Catholic schools (Katholiek Onderwijs Vlaanderen) belongs to the latter group, providing education to about 65% of students in primary education in Flanders, being by far the largest education provider.
[2] There are no central examinations like INVALSI in Belgium but all education providers have their own assessments for quality assurance purposes. In particular, the network of Catholic schools administers yearly standardised formative tests to students in their final year of primary education (grade 6).

***Keywords:*** school closures; teacher shortages, summer schools, standardised tests, inequality, primary school

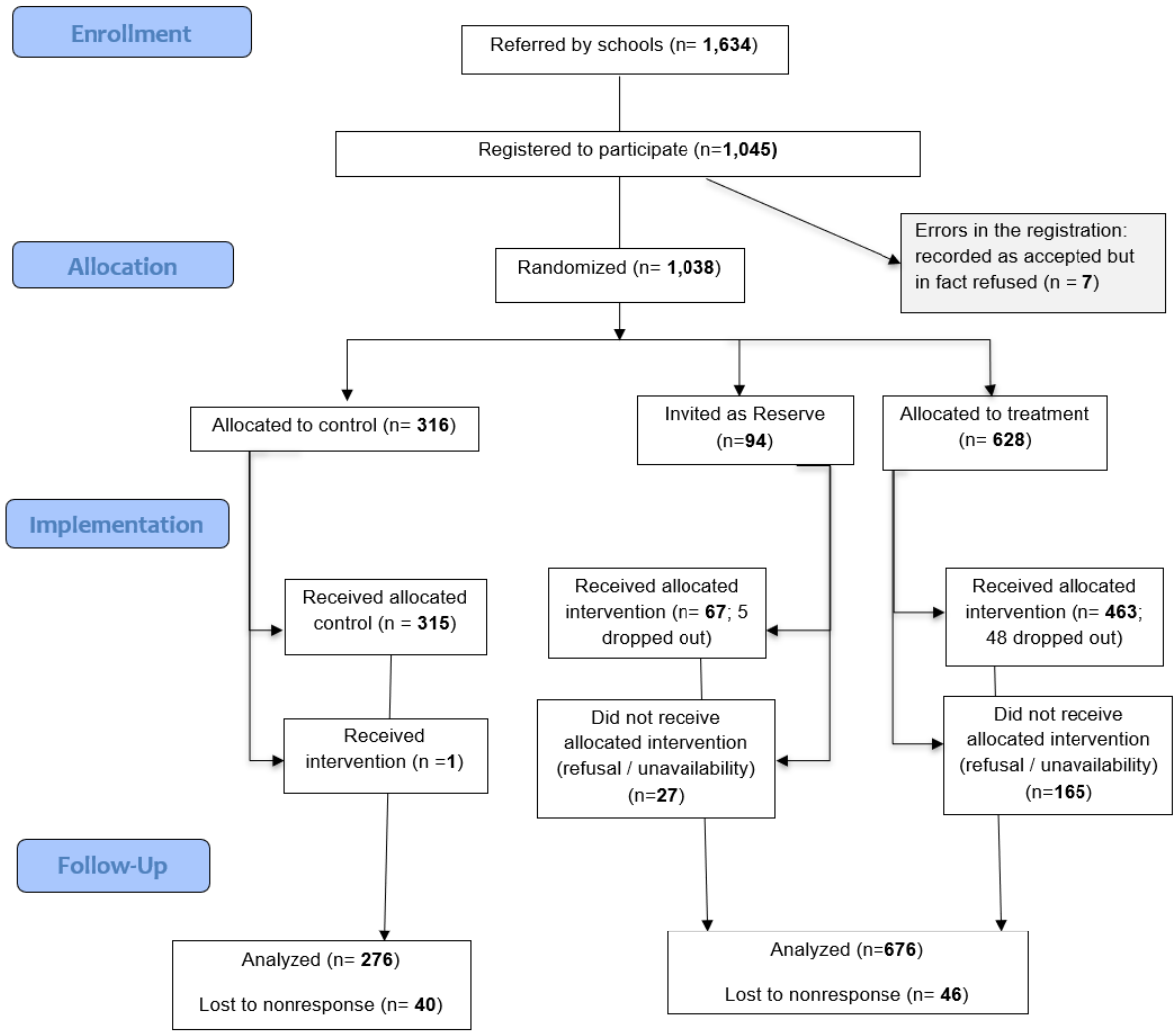
## Starting the school year on the right foot. Effects of a summer learning program targeting low-performing students in Italy

**Davide Azzolini – Martina Bazzoli – Sergiu C. Burlacu – Enrico Rettore**

Purpose. Our study sets out to extend the evidence base on summer learning loss and the effectiveness of summer learning programs to a context where these topics have never been studied before, i.e. Italy. The study employs a randomized controlled trial to estimate the effects of a summer learning program targeted

to low performing students (i.e. *Arcipelago Educativo*) on student learning achievements in math and reading. The study was preregistered at the AEA RCT registry (AEARCTR-0009958).
Setting. Italy is a country with lower than OECD average test scores, relatively high levels of inequality (OECD 2019), and a strikingly low proportion of tertiary graduates (OECD 2023). Despite its 12 uninterrupted weeks of holidays (European Commission 2022), summer learning programs are a rarity in the country. *Arcipelago Educativo* is one of the only systematic attempts to introduce this kind of interventions in the country. The 2022 edition of the program was run in 10 cities all over of the country.
Participants. 17 primary and middle schools referred 1,634 students as in need of learning support during summer. 1,038 subjects accepted to participate in the study. About one third had an immigrant background and a similar share of students was reported to have some special education needs—and thus entitled by law to receive personalized support.
Program. The first component of the program consisted in 88 hours of educational and recreational activities delivered to small groups (10 students each), which were meant to transfer educational contents, mostly in the area of mathematics and reading, as well as to develop sociality through peer education and cooperative learning. The second component was a 12-hours personalized tutoring intervention, which was delivered to groups of two or three students and whose major focus was helping students recovering the learning in areas where they had the greatest gaps.
Design. Right before the summer break, participating students (1,038) were randomly allocated to either the treatment or the control group, following a stratified randomization design, where the strata used were the school and the educational level (primary or middle school). 722 students were allocated to the treatment group (94 of them were initially allocated to a reserve group, but all of them received the invitation to participate only a few days after). The remaining 316 were allocated to the control condition, with the possibility of participating in learning support activities in Fall. T-tests and F-tests were run to check the statistical equivalence of the groups at baseline and after attrition in the late summer test.

Figure 1 Experiment's flowchart




Data collection and analysis. Between late May and early June 2022, before the end of the school year, all participating students took four grade-specific learning tests in Italian (grammar and comprehension) and mathematics (arithmetic and geometry). The tests were created specifically for the study and tried out by some teachers in different classrooms. In addition to the tests, some psychometric scales were also collected. The follow-up tests were almost identical and administered at the end of the summer. Because of organizational constraints, the follow-up tests were administered by the staff of the program during the last week of the summer programs or during the first two weeks of the new school year.

Findings. 530 out of 722 students accepted to join the program. On average, they participated in 61 hours of workshops and 10 hours of tutoring. The evaluation provides both Intent-to-Treat (ITT) and Treatment-on-Treated (TOT) effects of the program. The estimates were recovered through regression models, which accounted for the randomization strata and for baseline characteristics such as gender, age, migration background, grades, school days skipped, need for support in Italian or math/science subjects, and deprivation indices.

Table 1 shows the effects of the four tests and for the average scores in the math and reading competence areas.

Table 1 Effects of *Arcipelago Educativo* on students' learning

| | Arithmetic | Geometry | Math | Comprehension | Grammar | Italian |
|---|---|---|---|---|---|---|
| Control mean | 3.92 | 3.78 | 3.85 | 5.16 | 4.69 | 4.92 |
| ITT | +0.195 | +0.104 | +0.166* | +0.332*** | +0.171* | +0.259*** |
| TOT | +0.267 | +0.142 | +0.226* | +0.451*** | +0.233* | +0.351*** |
| N | 939 | 946 | 947 | 934 | 936 | 947 |

Note: Normalized scores (0-10).

* p<.10, ** p<.05, p<.01

The effects are positive and statistically significant for both Italian and Math. For Italian, the TOT is estimated in .19 SD, while for math the effects are smaller (.11 SD). In particular, the TOT estimate on text comprehension stands out, although the effects are not negligible on grammar and arithmetic (although not significant), while they are much smaller and not significant on geometry. The study shows that the program more than compensated for the (relatively) small summer learning losses estimated in the sample. While students in the control group lost ground during the summer (particularly in mathematics), students in the treatment group in September showed equal or higher levels of learning than they had before the summer.Exploratory analyses also show that the program's effects were largely driven by the effects on primary school students and students with special education needs. Finally, non-cognitive data point to an increase in students' desire to learn but also an increased preoccupation about returning to class and school assignments and tests.
Conclusions. Further research is needed in four areas: the measurement of learning (loss), the disentanglement of the separate effects of the two program components, a closer look into the heterogeneity of the effects, the investigation of medium- and long-term outcomes.

***Keywords*:** summer learning loss, randomized controlled trial, low performing students

## A systematic overview of the effectiveness of teaching action in feedback literacy as a tool for preventing, tackling and reducing implicit and explicit early school leaving factors in the national school landscape

**Rita Marzoli – Antonella Mastrogiovanni – Antonella Vendramin**

Introduction. In the latest report prepared by the Autorità Garante per l'Infanzia e l'Adolescenza (AGIA) on the issue of early school leaving (Autorità Garante per l'Infanzia e l'Adolescenza, 2022), specific recommendations are proposed for countering, preventing and succeeding at school. These recommendations envisage the involvement of several actors on several dimensions that make up the entire reference system of the student. Undoubtedly, the absence of an effective and continuous direction of actions considered as good practices, tested in particular realities, contributes to the 'dispersion' of energies in addressing this issue as a system measure. It is also true that the great variability of territorial and socio-economic conditions makes it impossible to identify general guidelines valid throughout the country. Within this report, specifically in Recommendation No. 4, the AGIA proposes to "invest in a strong renewal of didactics and teaching styles, with a view to facilitating pupils' active participation in learning processes and personal growth, as well as to support pupils' "feeling of welcome" in the educating community, through cooperative work in small groups, peer learning, active citizenship education".
The promotion of effective didactic actions aimed at preventing, contrasting and reducing this phenomenon certainly represents an element that can be used as a system measure, with a view to the activation of training courses dedicated to the teaching staff, especially in those contexts where this phenomenon is most prevalent. The international literature on the study of didactic actions linked to strategies for developing feedback literacy, highlights the increase in students' capacity for autonomy in learning and, more generally, school well-being (Scierri, 2021; Yan & Yang, 2021). However, the TALIS (2018) study shows that Italian teachers, unlike in several other international contexts, make less use of didactic actions involving moments of self-assessment by students and thus, in a broader sense of the formative use of assessment itself. In the present exploratory study, therefore, an attempt will be made to systematise research developed mainly in the national context that reports on the effectiveness of the use of feedback as a formative action in the school context.
Object and research hypothesis. Systematic literature review of empirical studies carried out nationwide that refer to the effectiveness of teaching strategies in the field of formative assessment, with the aim of identifying potential good practices to be translated into national frameworks.
Data. Empirical studies from the systematic review.
Method. We converted the research concept into three key concept groups by employing the Population, Concept, Context (PCC) mnemonic Aromataris (2020) (available from https://synthesismanual.jbi.global. https://doi.org/10.46658/JBIMES-20-01), using it to identify the main concepts in our primary review question. This framework will then inform our search strategy.
Results. The analysis of the systematic reviews will allow the generation of specific information for the construction of a framework useful for the preparation of continuous training actions for teachers, aimed at the prevention, contrast and reduction of implicit and explicit early school leaving.

***Keywords:*** educational research, dropout, systematic review, teaching strategies, feedback literacy

## Hidden drop-out: Secondary education (unseen) failure in pandemic times

**Lorenzo Alderighi – Rosario M. Ballatore – Marco Tonello**

Introduction. The pandemic outbreak in 2020 has resulted in a significant worsening of many aspects of people's lives. Young people, students especially, were among the most affected. From one day to next, many governments decided to close schools and switch from in-person to online remote teaching (Svaleryd & Vlachos, 2022). Largely, this switch occurred in places where schools, families, or both were not prepared to handle with this new learning setting (OECD, 2019). Indeed, many schools and households were not fully

equipped with IT solutions, and, often, teachers did not have sufficient digital skills to provide effective educational input, making learning very difficult.

These disruptions raised a lot of concern about negative consequences on children, resulting in a large number of studies. Most of them analyze the effect of the school disruptions induced by the pandemic outbreak on achievement, making use of standardized test scores and especially focusing on students at the beginning of their educational path. A large consensus has been found on a substantial negative impact from the pandemic, with effects on achievement being strongly heterogeneous across students' and families' characteristics (Betthauser et al., 2023).

Research question. The question of whether these learning losses have translated into a failure to achieve the minimum level of skills required for a given school grade remains unanswered, as this educational failure cannot be seen in standardized test scores alone and is often masked by the formal academic success in passing from one grade to the following. Investigating this aspect is even more important for high school students, as they may have no time to recover. This paper aims to unveil this unseen failure by analyzing an innovative aspect on which the literature has yet not focused: the hidden drop-out of students at the end of secondary school. Different from the usual notion of "school drop-out" (i.e., students who are no longer enrolled in secondary school and have not received a secondary school diploma or recognized equivalent), hidden drop-out refers to students who formally completed secondary school, but did not acquire the skills or demonstrate the competence required for the grade reached.

To this end, we evaluate Italian students in grade 13 during the pandemic, a crucial turning point in each student's life that coincides with the end of high school and the decision of whether to start tertiary education or look for a job in the labor market. While some recent evidence on primary school students points to rapid recovery of learning losses once the usual learning environment has been re-established (Singh et al., 2022). High school students, who have already left the educational system, accumulate losses that, in absence of specific recovery policies, no schooling or labor market institution can remediate. This means that any unfilled educational loss at the end of high school may plausibly affect success permanently in tertiary education, labor market outcomes, or both.

Data. We use the administrative data provided by the National Institute for the Evaluation of the Education and Training System (INVALSI). INVALSI implemented the National Evaluation Program of Students' Achievement (NEPSA) in the 2009-2010 school year with the aim of attaining yearly census data by testing core skills in language (Italian) and mathematics in students in primary (grades 2 and 5), junior high (grade 6), and high school (grade 10). Since the 2018-2019 school year, grade 13 students also take the tests. Importantly for our identification strategy, thanks to an anonymous individual-level identifier, we merged the NEPSA tests of grade 13 students with the grade 10 NEPSA results of the same students. This allows us to collect data about the same student across years and tests.

Together with the standard test scores results and basic demographic characteristics, for students in grade 13 INVALSI also provides a new variable which, for each subject, classifies students into five increasing levels of ability (Desimoni, 2018). This variable is the result of an accurate psychometric analysis that identifies the true student's capacity. In contrast to simple tests, which assign all answers the same relative importance, to obtain the attainment levels the psychometric analysis weights each item used for evaluation according to the skills required to answer it. Level 3 corresponds to the minimum threshold of knowledge and skills that the student should have by the end of high school. Based on this variable, we create our measure of hidden drop-out, which is a dummy that identifies students who fail to reach level 3 both in math and Italian language assessments. Note that all students defined here as hidden dropped-out formally completed high school. In other words, they received their diplomas without possessing the minimum knowledge and skills sufficient to earn them.

Estimation method. To estimate the effects of the COVID-19 pandemic on hidden drop-out we exploit the exogenous variation induced by the pandemic in comparing two cohorts of students: the last of those who completed grade 13 before the spread of COVID-19, and the first who experienced for about one-and-a-half-year period the overall disruption to schooling caused by the pandemic. The effect on hidden drop-out is retrieved by comparing the students in the two cohorts within the same school, and taking into account several pre-determined student characteristics, including prior level of achievement. The baseline regression takes the following form:

$$HDti,k,j = \beta0 + \beta1Ci,k + \beta2Xti,k,j + \beta3X\ t{-}1i,k,j + \beta4At{-}1i,k,j + \eta j + \epsilon i,k,j$$

where: HDti,k,j is a dummy that identifies student (i) of cohort (k) and school (j) that did not reach the minimum level of Italian language and math competences in grade 13 (i.e., hidden drop-out) and Ci,k is a dummy indicating whether the student belongs to the COVID cohort. We also control for a vector of students' and class characteristics (Xti,k,j),including dummies for female, foreign (i.e., non-Italian citizen), grade retained (before grade 10), class size (measured by the number of students in the class, both in linear and quadratic terms), and for school fixed effects (ηj). Crucially for our identification strategy, we exploit the longitudinal dimension of the data by adding to the specification a measure of student's achievement at the previous assessment in grade 10 (At−1i,k,j),as in a value-added specification (Todd & Wolpin, 2003), and the family socio-economic status index measured in grade 10 (X t−1i,k,j).

The main assumption for attaching a causal interpretation to β1 is that – conditionally on predetermined controls - the treatment is independent of potential outcomes of both treated and control cohorts. While we can safely assume that the pandemic affected the COVID cohort in an exogenous way, our estimates also root on the assumption that the two cohorts of students are similar in all observable and unobservable aspects that could affect the human capital accumulation process. Given that the two cohorts of students are really close to each other, we believe that this assumption is basically free of all the concerns regarding cohort-specific demographic composition and trends.

Results. We estimate that the pandemic induced a 8.6 percentage points increase in the probability of hidden drop-out, corresponding to a 49.4% increase with respect to the average probability of hidden drop-out in the pre-COVID cohort. Our estimate corresponds to an annual learning loss of 0.23 units of standard deviations (SD), with respect to the pre-COVID period. Our results lie below the median negative impact of 0.08 SD units found by Hammerstein et al. (2021) or the average loss of about 0.14 SD units that Storey and Zhang (2021) and Betthauser et al. (2023) reported in literature review and meta-analysis of studies related to learning loss due to the pandemic. However, two main caveats are in order. First, our study considers the effects of the pandemic on learning in the wider time horizon of the first and second waves (February 2020-May 2021). Second, usual yearly progress in math and language are declining over time (Bloom et al., 2008), such that the comparison above could be misleading in understanding the learning loss suffered by older students. Indeed, as already noted, the literature has concentrated especially on students below grade 10, to the exclusion of older ones. For these reasons, we believe more appropriate to compare our results with the average natural yearly progress in high schools, which is estimated around 0.20 SD units (Bloom et al., 2008). With respect to this estimate, the pandemic completely destroyed one year of learning progress in high school students.

**Keywords:** Hidden dropout, school failure, secondary education, Covid-19

# THEME 9. ITALIAN, MATHS AND ENGLISH: TEACHING PROCESSES THROUGH INVALSI DATA AND TOOLS

**ORGANIZER: INVALSI**

**COORDINATOR: DANIELE VIDONI**

**25TH NOVEMBER: 4.00 P.M. – 6.30 P.M. {ROOM 1 – TEACHING SESSION 3}**

---

## Promoting algebraic thinking: restructuring INVALSI items to turn them into an opportunity for formative assessment

**Simone Quartara – Francesca Morselli – Elisabetta Robotti – Alessandra Boscolo**

Introduction. This contribution presents a teaching activity, experimented in a first class of a scientific high school, designed around the idea of using INVALSI tests in the classroom adopting a formative assessment perspective, where formative assessment is conceived as a classroom practice in which teacher and students are equally protagonists.

The proposed activity was designed and tested in the context of the community of inquiry DIVA (Didattica Inclusione Valutazione (Formative) Argumentation) at DIMA (Department of Mathematics) in Genoa, a group set up in February 2023 and aimed at reflecting on didactic practice in secondary school. The first theoretical tool shared within the group was Schoenfeld's TRU (Teaching for Robust Understanding) (2018), according to which a 'robust', i.e. effective, lesson must be articulated around the following dimensions: mathematical content, cognitive load, fair access to content, agency (ownership, identity) and formative assessment. These dimensions can be used by the teacher as a tool for observation and reflection on his or her teaching with the aim of planning, evaluating the effectiveness of the intervention and thinking about the next steps in the teaching action.

In this contribution we take the dimension of formative assessment as central, starting from Black & Wiliam's (2009) definition that a practice becomes formative when "evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited" (p. 7).

In order to design activities from a formative assessment perspective, the formative assessment strategies illustrated by Wiliam and Thompson (2007) were taken into consideration: (a) clarifying and sharing learning intentions and criteria for success; (b) engineering effective classroom discussions and other learning tasks that elicit evidence of student; understanding; (c) providing feedback that moves learners forward; (d) activating students as instructional resources for one another; (e) activating students as the owners of their own learning.

According to the perspective adopted within the FaSMEd project (Cusi, Sabena and Morselli, 2017), formative assessment, understood as a teaching methodology, can find its full implementation in argumentative activities, in which students themselves are led to reflect on the teaching-learning processes and to make their thinking visible (Collins, Brown & Newmann, 1989), thus sharing their ideas with the teacher and classmates. In this way, each student becomes responsible for his or her own learning (he or she has to justify the given answers) and the thought 'made visible' can be the subject of feedback from the teacher or peers.

This contribution will show the design and subsequent implementation of activities in which the request to justify answers, first individually and then in small groups, prepares for the fundamental moment of class discussion, thus pursuing the full implementation of all the formative assessment strategies.

The proposed activities have as their learning objective the development of symbol sense (Arcavi, 1994): students must develop a profound understanding of how and when to use algebraic language to represent relations, generalise and demonstrate, an ability to manipulate and interpret symbolic expressions beyond the mere application of procedures and to select the most suitable symbolic representation for the purpose, and an awareness of the fact that symbols can take on different roles in different contexts.

Research object and hypothesis. The research is aimed at developing and testing activities for the development of algebraic thinking designed on the basis of INVALSI items and declined from a formative assessment perspective. The research hypothesis is that formative assessment can become a classroom practice shared with students, through a design and implementation that activates multiple formative

assessment strategies, including those in which students become agents of formative assessment (on themselves and their peers).
In relation to the specific learning objectives related to the symbol sense, the Gestinv database, created within a research project aimed at providing tools and models of action to address how to integrate the results, methods, theoretical frameworks and in general the tools of INVALSI assessments in the local actions of teachers and schools (Bolondi, Gambini and Ferretti, 2017), provided us with information on the most widespread elements of difficulty at national level. Moreover, it allowed us to select tasks consistent with the National Guidelines and significant in relation to specific aspects of the symbol sense.
The proposed activities have as their central core two INVALSI items (D14 G10 of 2010, D17 G8 of 2010), modified so as to include argumentative activities. The alternation of individual work, small group work, class discussion and individual reflection by means of a questionnaire is functional in activating the various formative assessment strategies.
Data. The experiment was carried out in a first class of a scientific high school and took place in May 2023. It should also be emphasized that the systematic use of class discussion characterized the classroom work throughout all the school year. For the research, we have at our disposal the data collected during the experimentation: individual and group answers to the proposed argumentative questions, screenshots of the blackboard shared by the teacher, video recordings of the class discussions, questionnaires (google forms) filled in by the students between lessons.
Method. The collected data are analyzed qualitatively. In particular, group responses and transcripts of discussions are analyzed, highlighting not only the developmental traces of symbol sense, but also the activation of different formative evaluation strategies by the teacher and students.
Results. In the oral communication, the genesis of the assignments from each INVALSI item will be described and the main findings from the analysis of the group work and subsequent class discussions will be discussed. In relation to the research hypothesis concerning formative assessment becoming classroom practice, the analysis of the class discussions will show the activation of multiple formative assessment strategies by the teacher and also by the students.

***Keywords:*** teaching practices, formative assessment, TRU, symbol sense

## Scientific curriculum design starting from maker education experience

**Laura Castellana – Teresa Di Tullio**

Introduction. The national guidelines for Kindergarten and Primary school, published in 2012, are both the didactic-disciplinary map used by teachers to plan their activities and the tool the school uses to design its curriculum that is its own *educational identity card.* The curriculum design is the expression of school planning ability (Cattaneo 2011), given the necessary adaptions to the context and the needs of the pupils. The curriculum often is the summation of the programmes, rather than the tool for defining the articulated process of learning and teaching (De Mauro 2001). In Kindergarten and Primary school the curricula are centered on student or on learning object: the first approach takes into account students' needs, capabilities and objectives and it requires the teacher carefully researches materials suitable for different learning needs; the second approach, by promoting the problem solving ability and creativity development, enables the pupils to use acquired knowledges to find new solutions, to change skills, to improve what ha been learned. How we can choose the right approach to curriculum design? What are the main directions to develop curriculum? What relevance do we attach to teaching and learning? To answer to these questions, the school, designing its curriculum, has to follow three main directions: training direction, by choosing common school projects, integrated intro curriculum; epistemological direction, by INVALSI data analysis, aimed to improve educational activities; management direction, by setting up best practices able to bring changes in teaching and learning.
Object and research hypothesis. The goal of this work is scientific curriculum design driven by STEAM, in order to have both a didactical framework to develop educational challenging activities and an integrated approach to different teachings. Our experiment starts from INVALSI data collected during school year 2021-2022 and it goes on the choice of three activities carried out in primary school, children 8-9 years old: coding, educational robotics, Arduino to teach environmental sustainability education. We carried out an

ecogame to speak about environment and to promote changes in attitudes and behaviors both individually and collectively, by bringing students closer to the robotic world and computational thinking. In this way the students become able to extricate from problematic situations and to build algorithms that provide effective and functional solutions. The students, led by their teachers, have designed an algorithm transmitted, by IDE software, to ARDUINO, that is a small electronic chip based on C++ and Java. The same algorithm has been translated with Scratch, that is a free software visual environment, based on block flow chart.
The students had the opportunity to work on real tasks, by activating designing processes, production and comparison able to bridge the gap between real life and traditional educational path (Dewey, 2004; Doppelt, 2009). When there are multiple solutions to a problematic situation and the learning process develops through trial and error, the ideal conditions are created for the development of creativity and life skills (Sala et al., 2020), useful for managing challenges and changes in personal and professional life (Bocconi, Kampylis & Punie, 2012; Bombieri & Giusti, 2021). The young students, in the guise of makers, actively and experientially built their knowledge through practical activities able to combine manual skills with the exercise of digital skills, aimed at solving open problems inspired by everyday life and the creation of playful and creative artefacts (Repetto, 2020). These are the significat elements of Maker Education which represents the pedagogical background of STEM and STEAM.
Data. The didactic choices were made following the analysis of the INVALSI data for the 2021-2022 school year: the class, that recorded the highest results in the Institute and results higher than the national averages, carried out the activity with Arduino, because it required well-established basic skills and the ability to mobilize knowledge and skills through the use of personal skills able to generalize and to find new knowledge and skills in front of different contexts; the class that recorded lower results than the national averages carried out the activity on the Scratch platform, because the graphic language and the visual approach facilitate familiarization with new educational scenarios that were anything but traditional ones. An important element involved in the experiment was the teachers' ability to get involved, both in the training process and in the implementation of the activity, as these features were aware of promoting a general educational purpose, capable of increasing the pupils' skills.
Method. These activities are often considered for technology insiders and therefore they are reserved to few teachers and the training activities are developed only in few classes. On the contrary, given the significant positive implications of these activities, it is strategic to include them in the school curriculum, because they involve all students, demystify fields of knowledge considered difficult (robotics, mathematics), help students to understand the functioning of life objects and processes and to develop problem-solving skills (Xenos and colleagues 2017). According to Scaradozzi and colleagues (2019), primary school curricula are mainly oriented towards science and mathematics, dedicating little space to information technology, problem-solving and robotics. Although the evident lack of guidelines on how to integrate coding and robotics activities into ordinary teaching and the difficulties in assessing the complex skills stimulated by these new practices (Screpanti, Miotti & Monteriù, 2021), a starting point is to identify nine categories for a quality integrated STEM curriculum (Walker et al., 2018): (1) motivating and engaging context, (2) technology design challenge, (3) integration of scientific content, (4) integration of mathematics, (5) student-centred instructional strategies, (6) teamwork, (7) communication, (8) organization, (9) performance and formative assessment. Thus, the choice of contents cannot be random or improvised but must be functional to the objectives of DigComp. Gamification is strategic, because it is able to help pupils, through physical and mental stimulation, to activate processes of understanding, helping them to acquire disciplinary and socio-emotional knowledge. Moreover, gamification allows to maintain high levels of satisfaction and interest, thanks also to the personalization of the paths (Scratch is optimal as a platform in this) guaranteeing spaces of freedom, personalization, choice, and making progress immediately visible. Any mistakes thus become just an opportunity and an incentive to try again and instead of the traditional evaluation, the scoring systems typical of digital games are inserted.
Results. The results of this work are a radical rethinking of the teaching and learning path in the scientific area, through flexible planning and didactics based on skills capable of effectively integrating technologies, to manage and integrate the complexity of current school contexts. The measure of the didactic value of the experimentation in progress will be the subject of a subsequent investigation and will have as main objective that of analyzing the results that the classes enrolled in the experimentation will obtain at the end of the third and fourth class and in the INVALSI tests when they will be in the fifth class.

***Keywords:*** school curriculum, STEAM, gamification

# At school with Mother Nature: INVALSI tests in outdoor

**Daniela Ruffolo – Carmela Gabola – Antonietta Cammarota**

This contribution proposes a research activity on education carried out by Don Milani-Linguiti Comprehensive School of Giffoni Valle Piana (Salerno) on the use of new educational practices to improve pupils' school performance, starting from INVALSI test results. Considering the high risk of both implicit and explicit school drop-out phenomena in the territory where the school operates, the board of teachers have been thinking on the need to reorganize both school and extra school activities, adopting orientation didactics already with primary school pupils, working on their well-being, sociality, self-esteem, self-determination and their sense of belonging to the community and its places.
Being aware that meaningful learning processes take place not only in school environments but mostly in the community educational spaces, our school has been planning curricular and extra curricular activities in collaboration with the third sector to offer educational opportunities to all pupils, especially to those living in fragile family environments.
The experience hereafter proposed was addressed to a second form of primary school, made up of varying levels and a special need child with communication and relationship issues that affect her expressive speech, relations, learning process and adaptability.
Starting from the analysis of the INVALSI items of the test done by pupils in May, the school planned an orientation campus financed with PON funds, based on group and individual activities according to the laboratory operative approach in an informal context, mostly outdoor.
The orientation campus was held at Casa Natura, a municipally owned building immersed in the green of the mountain which is the symbol of Giffoni, a venue run by the third sector associations that are partners of Edu@ction Valley.
The interventions were planned by several professionals involved in the campus: school teachers, the special need teacher, the expert and volunteers of the Educational Community Edu@ction Valley. The representatives of the territory involved in the activities (a beekeeper, some volunteers of the Environment Educational Centre, an expert in Forest Bathing and others) contributed with their expertise to offer an alternative educational programme fully immersed in the nature, experimenting with the children "new learning experiences" outside the classroom.
The activities that are mainly based on observation and study of the natural heritage of the context (fauna and flora biodiversity) and on sustainability, aimed to develop the pupils' several competences, from the alphabetical-functional one to STEAM, digital, relational, emotional and citizenship competences.
The items of INVALSI ITA and MAT tests were analysed by the teachers and the expert and they were used to plan the educational activities proposed to the children. To develop children's competences several teaching strategies were adopted: reading aloud using more voices, storytelling, onboard diary, organising and managing roles, Forest bathing, using dichotomic tables to observe and identify the natural context, Math games. Though ina a naturalistic environment, technology was not neglected, as mobiles were used to shoot "Casa Natura NEWS", under the supervision of the expert volunteers of the associations.
At the end of the campus pupils were given the INVALSI tests (Italian and Maths) they had already done in May 2023. They did the tests in Casa Natura sitting around a big table. They were asked to do the tests in absolute silence, totally plunged into the sounds and scents of the surrounding Nature, that natural envirnoment they had discovered during the whole campus. The setting in an informal context allowed the pupils to do the tests in a short length of time, in absolute tranquillity and without any performance anxiety. The results were excellent, proving the efficacy of the teaching practices used during the experiential campus.

***Keywords:*** teaching practices, school improvement, learning, educational research

# The role of cognitive process in designing and constructing the national assessment of mathematics for primary school

**Francesca Ferrara – Stefania Pozio – Ketty Savioli**

Introduction. A recent research paper published in the journal Cognitive Development examines how page layout in children's math books may or may not promote engagement with mathematics and may be a relevant factor in children's engagement with it (Ernst et al., 2023). Specifically, the research considers counting books and the potential effect of their page layout on the counting processes developed by children, e.g. from the arrangement of illustrated sets of items on one rather than two pages. Other aspects that may influence counting development, such as the context of the story and the number of words on the page, are not studied here. In previous research, van den Heuvel-Panhuizen and Elia (2012) evaluated the effect of picture books for preschoolers on their mathematical learning. Several studies, focusing on the numerical domain, have investigated the mathematical content present–e.g., the range of numbers presented, the way number is represented–e.g., by item sets, numerals, or number words, features that characterize number sets–e.g., identical item or similar items that vary in color or size (Powell & Nurnberger-Haag, 2015; Splinter et al. 2022; Ward et al., 2017). Although the focus of these research studies is specifically on evaluating the media considered for learning in given contexts, they highlight that in the presentation of mathematical information, certain features or choices can influence its understanding. In research using eye-tracking methodologies to study the focus of attention in mathematical stimuli, moreover, it has been examined how task-relevant information is sought differently where multiple representations of the same object, e.g., a text, a formula, and a graph, are introduced (Andrà et al., 2009; Holmqvist et al., 2011; Andrà et al., 2015). The explanation, essentially, is that from the perspective of the comprehension process a text, a formula and a graph require different kinds of interactions, also involving the eye in different ways. After all, it is not news that register treatments and conversions among registers are essential for cognitive development in mathematics, as Duval (2006) has argued and as confirmed by the numerous studies on the multimodality of mathematical knowledge, which give new life to the interweaving of representational registers (e.g., Ferrara, 2014). Mathematical tasks then, especially when written, always require reading skills, as students need to read the task in order to solve it. Secondly, mathematical communication is considered one of the main aspects of school mathematics mastery, as reflected in national and international curricular documents (e.g., Niss & Højgaard, 2011). In other words, the ability to read (and write) mathematics is a fundamental part of mathematical proficiency, despite the fact that reading a mathematical text is not seen as a learning opportunity per se but rather as a type of skill needed to become active in situations where learning can occur, e.g. in solving given tasks (Österholm, 2006). Several studies indicate a strong association between reading and mathematics tests (Caponera et al., 2016; Ajello et al., 2018), but it remains important to avoid unnecessary reading demand in mathematical tasks (Bergqvist et al., 2018). Starting from this background of the literature, in this contribution we want to offer a reflection on how, in the construction of mathematics assessment tests (particularly for primary school, where we are personally involved), it is necessary to take into account the complexity of the thinking processes that accompany reading, understanding and solving a mathematical question. To this end, we will show examples of the changes undergone by a question in its transformation from conception to participation in a national test. This transformation is characterised by various modifications, which may be linked to difficulties and ambiguities that emerged, for example, during the pre-test, rather than to linguistic, representational or visual cleansing, in order to lighten the cognitive load required to solve the question. This study poses two interesting reflections, one on the depth implied by the creation of the assessment tests and the other on the didactic relevance these can assume in relation to the complex nature of mathematical learning.

Research object and hypothesis. The subject of our study is the complex and profound processes governing the transition from the conception of questions to their participation in the pre-test and eventually in the national assessment for primary school. Since the questions involve the use of different registers of representation, these being characteristic of mathematical thinking, as well as specific learning objectives (in line with the National Guidelines) and different cognitive dimensions (knowing, problem-solving, arguing), it is essential to have an eye on the cognitive load required by their resolution. This load can be interpreted in terms of the processes that each question mobilises. An open-ended question that asks for the result of a given operation expressed numerically, for example, is quite different from an argumentative question that asks to choose among several answer options the one that explains a given result in natural

language. Our hypothesis is that the construction of the mathematical questions for the national tests deserves and needs to address the many variables that affect the issue of readability, comprehensibility, accessibility with respect to the competences expected at a certain school level (and the learning objectives at that time). That is to say: How can a question be made readable, comprehensible, accessible? What are the characteristics and choices to be paid attention to?
Used data. The data we use for our study are primarily some questions from the national primary school assessment of mathematics (from grade 2 and grade 5), considered in their successive versions, from the original version of the question at the stage in which it was created to the final version in which it participated first in the pre-test, then in the national test. We will focus our attention on the changes undergone in the transition from the author's original version to the final version reworked by the group constructing the pre-test forms. In particular, we will reflect on the elements that from a cognitive point of view may affect the solving process and either compromise its success or facilitate the readability, comprehension and accessibility of the questions. Since the processes at work are fundamental to these reflections, we will finally show some protocols taken directly from the pre-test phase, which help strengthen our argument.
Method. For each of the questions used in this study, both open-ended and closed-ended, we take the original version and the final version. We therefore examine the changes which each question undergone in the transition from the first to the last version. At this stage, we will use at least two methodological tools for our interpretation: on the one hand, results from mathematics education research (such as the influence of certain features of the questions on the solving process); on the other hand, a software called READ-IT (made available by CNR researchers who study the dynamics of language: Dell'Orletta et al., 2011) which allows us to assess lexical, morpho-syntactic and syntactic characteristics of a text, taking into account both the projection of readability on the text and a global analysis of readability. The observations we will make in the light of our interpretation will be qualitative in both cases.
Results. Our study aims to reflect on the role of the cognitive process in the design and construction of primary school assessment of mathematics. Under the umbrella of cognitive process we can bring together all those elements that affect the construction of mathematical knowledge and meanings. Beyond the product, be it the expected outcome of a question or the question itself that partakes in the national test, there is in fact a process of growth and change that accompanies the transformations undergone by the question in its progress towards the test. This process of growth and change can be expressed through attention to the wording of the question, the use of different registers of representations and not only linguistic but also representational and visual cleansing. The conceptual and processing aspects, not just the executive ones, involved in solving the question are an integral part of the process and help to better understand the link between the complex nature of mathematics learning and the assessment of mathematics, allowing for a greater appreciation of the latter's value as a resource for teachers.

***Keywords***: mathematics learning, cognitive process, registers of representation, language, national assessment, conceptual

## An initial exploratory study of the perceived difficulty of students and teachers with respect to a mathematical task

**Bianca Nicchiotti – Camilla Spagnolo**

Introduction. Research in Didactics of Mathematics has focused on the aspect of difficulty in Mathematics, leaving in the background the perceived difficulty of a mathematical task, despite the fact that the perception of difficulty is a much-studied topic in metacognition research (e.g. see Efklides & Touroutoglou, 2010). Consequently, it has been possible to trace difficulty in mathematics back to a set of factors that cause and influence it, including text comprehension (Spagnolo et al., 2021), formulation of the problem (Bolondi et al., 2018), mathematical content (Radmehr & Drake, 2017) and affective factors (Zan, et al., 2006), whereas it has not been possible to give a similar characterisation for perceived difficulty. Within the scope of exploratory studies towards this direction, factors influencing students' perceived difficulty have been grouped into five categories: factors related to task resolution, factors related to ability and experience, emotional factors, factors related to task formulation and factors related to self-regard (Saccoletto &

Spagnolo, 2022; Spagnolo & Saccoletto, 2023). Analysing the perceived difficulty with respect to a mathematical task, the connection between teachers' and students' perceptions is particularly relevant; in fact, studies have shown that teachers are not always able to identify the reasons for students' errors (Faggiano, et al., 2023; Spagnolo et al., in press).

Research object and hypothesis. The research developed aims to carry out an initial qualitative exploration and comparison of the difficulty perceived by students and teachers with respect to the same mathematical task, using INVALSI data and questions. The decision to use INVALSI questions was dictated by the need to select tasks deemed appropriate for students at a certain point in their education. In addition, since 2018, Italian secondary school students have been taking INVALSI cbt, with the administration of tests containing questions that differ from student to student but are overall of the same difficulty. Starting from the same year, the percentages of correct and incorrect answers for each question are no longer disclosed, as they are not reliable, but rather each task is assigned a priori a "level of competence", on a scale ranging from 01 (the lowest) to 05 (the highest), which basically provides information on what should be the objective difficulty of the item.

Data used. The research involved 14 students of a second class of a scientific high school (liceo scientifico) and their mathematics teacher. A semi-structured online questionnaire was administered to all of them via Google Form in May 2023; the end of the school year was chosen to ensure that the teacher had had a chance to get to know the class. The questionnaire was divided into three distinct sections, containing two INVALSI questions taken from the 2018 test, accompanied by some questions strictly related to the assessment of the perceived difficulty for each of the tasks, the reasons for this assessment and some more general questions related to the perceived difficulty in general for mathematical questions. The version submitted to the teacher had the same characteristics, only some questions were modified with reference to the teacher's experience with her own students. The selected questions are both related to Space and Figures, and both refer to the Argumentation dimension; however, they are of a different type in that the first is a cloze while the second is an open-ended question. Both require knowledge of theorems and properties of plane geometric figures, but the levels of competence attributed are different: the first question is rated 03, while the second 05, since it is considered that to carry out the second question it is necessary to know theorems and geometric properties but also to be able to use this knowledge to create a demonstration.

The levels of competence assigned to each question allowed us to make a comparison between the difficulty of the questions and the perception of it by the students and their class teacher.

Method. The questionnaire was structured similarly to the one proposed in (Saccoletto & Spagnolo, 2022), since we used the categories determined in that study for the students as a starting point. The questions chosen, however, were different. From the methodological point of view, the research also refers to constructive grounded theory (Glaser & Strauss, 1967; Charmaz, 1990). In fact, we analysed the responses obtained using an inductive approach, starting from the general and then narrowing the vision to the particular and identifying the elements classifiable according to the predetermined categories, integrating these categories according to the responses obtained. We finally interpreted the results obtained by comparing the students' and teacher's answers.

Results. From the analysis of the answers obtained, we first found that the students and the teacher rated the perceived difficulty of the two tasks differently. The students found the two tasks to be, on average, of the same difficulty, to which they gave a medium-low rating; however, especially for the second task, the ratings given were very polarised at the extremes, with students who perceived the task as very easy or very difficult, almost totally missing the perception of medium difficulty. The teacher, on the other hand, perceived the first task as more difficult than the second. Despite this, the perceived difficulty rating given for the first task was in agreement with the level of competence attributed by INVALSI.

The motivations provided, by the teacher and the students respectively, also involve different elements. The teacher seems to assess the perceived difficulty mainly on the basis of the characteristics of the task (type of question, terms used, presence of figures) whereas for the students, elements linked to ability and experience, such as intuitiveness of the question and above all having already done similar problems in class, are more important. With regard to the latter element, it is worth mentioning that, although the teacher stated that the students had never tackled questions similar to those proposed, almost all the students stated that they had already tackled them. It might therefore be interesting to investigate the reason for this different perception.

***Keywords:*** perceived difficulty, mathematics education, teachers, students

# The space that teaches: *Laboratorium*

**Sebastiana Fisicaro**

This research investigates the physical space and methodological choices of some schools, which have received NRP funds. School space and methodological choices, which school construction conditions, constitute the construct related to what is normally understood by the term "learning environment." Learning environments are at the center of pedagogical and didactic reflection because they involve the school in a perspective view, affecting "a plurality of facets." The objective is to verify the level of awareness of schools in the use of spaces (classrooms, corridors, courtyards, etc.), linking to it the responsibility required by the regulations (Legislative Decree No. 81/2008) and the steps related to educational planning pertaining to the "School 4.0" Plan and investment line 3.2 "School 4.0," funded by the European Union - Next Generation EU - Action 1 - Next Generation Classrooms.

Physical space conditions the teaching actions of teachers, in fact, as Julia Atkin (Atkins et al., 2015) states, the effectiveness of the physical environment, for the purpose of improving learning, depends on the availability and choice that students and teachers have in the use of spaces and their degree of participation "in choosing" and arranging the spaces in which they work.

The design of learning spaces, in fact, stimulates a sense of belonging and enables their "identification with the new school environment" (Lippman, 2012b), so architectural choices need a didactic-pedagogical perspective (When Space Teaches, 2012- INDIRE). Student outcomes in standardized tests are determining, through the allocation of NRP funds, a widespread design of new learning environments for the implementation of laboratories, aimed at improving the levels of key competencies of students at risk of dropping out of school. The laboratories and, therefore, their design should intervene in a systemic way on school dropout in order to stem the phenomenon, which in pupils with level 3 is to be considered at high risk of implicit dropout.

Research object and hypothesis. The object of this research is, therefore, to investigate the role of awareness in the use of school space in the design of workshops for the improvement of levels at risk of dropout, and the contribution that an integrated systemic action has on the development of students' key competencies. The research, initiated in a qualitative-quantitative way, used the comparison of some documents published in the ministerial platform Scuola in chiaro, concerning a sample of schools in the province of Syracuse. The collection of data published by the schools makes it possible to cross-reference some data related to planning in the methodological field with data related to school construction in the schools under study. Initial results confirm that in schools with a higher number of pupils at risk of dropping out, some interventions that develop awareness in the use of space are considered strategic, although there continues to be a struggle in declining the combination of space (physical and digital) and learning (methodologies).

The problem of mismatch, highlighted repeatedly at the national level, as reflected in the FUTURA Guidelines (NRP 2022) "Designing, Building and Inhabiting Schools," requires schools to reflect on what is meant by functional and flexible learning environments. The problem is currently being partially addressed, from the sample of schools surveyed, due to self-justifying (pandemic) factors of learning loss and the steady increase in the lack of basic skills, 25 percent of pupils are below the minimum skill level (Save The Children).

The low levels of key skills of students at risk of dropping out Plan "School 4.0" and investment line 3.2 "School 4.0," - Next Generation EU - action 1 - Next Generation Classrooms- determined the allocation of funding of 10,004,315.01 euros for 70 schools between cycle I and II in the province of Syracuse. Most schools, which are receiving PNRR funds, have pupils with low levels of key competencies in standardized test results. Research conducted by the OECD PISA had already indicated how some of the difficulties that schools present must be attributed to the low percentage of teachers (only 50 percent) have technical and pedagogical skills necessary to integrate digital devices in teaching vs. the OECD average of 65 percent and, in addition, the inadequacy of time that teachers devote to preparing lessons with the integration of digital devices (60 percent vs. OECD average 68 percent).

On February 22, 2018, the document "National Indications and New Scenarios" had been presented at MIUR, in which students' digital competence (European Digital Competence Framework for Citizens -DigComp 2.0) is presented as the equity tool, which contributes to the reduction of the often unbridgeable gap in unfavorable socioeconomic context situations (INVALSI Report 2022).

Based on this premise, the survey involved:

- selection of a sample of 10 schools, secondary, in the province of Syracuse, recipients of PNRR funds for design action 1 - Next Generation Classrooms;
- documentary analysis of some of the most significant documents, published on the schools' website (PTOF, act of address of the DS, teachers' training plan) and on Scuola in Chiaro, especially, regarding the use of physical and digital spaces;
- documentary analysis of the proposed pathways for students to implement the PCTO;
- survey, through a questionnaire, of the most frequent relational dynamics and the level
- of perception, through focus groups (Weick, 1997), designed to detect the close relationship between cognition, emotion and students' school outcomes (James, 1980);
- survey on the sense of belonging and perceptions of the organization of space at school (Lewin, 1935);

The sample of schools identified, 10 upper secondary education institutions, was intended to construct, through mapping, a paradigmatic format, which monitors the correlation between the choice and dissemination of personalized methodologies and the data entered in the RAV by each school, the level of collegial perception about the preference of laboratory settings in promoting the improvement of their students' key competencies. The analysis of the learning environment provides each school with a toolbox to re-think the traditional model in favor of a laboratory school (M.D. 139/2007) as envisioned- Action 1- Next Generation Classrooms. The design of improvement actions, as required by the NRP funding, involves the responsibility of schools in increasingly laboratory-based teaching, helping to define the "agenda" of doing school in actions of functional methodologies and monitoring of learning outcomes in order to intervene in a more structured way on the current disconnect between internal and external assessments. Analyzing, comparing the data collected, intervenes in the reflection of key competencies, the school organizational model, and in light of the international debate on work tracks and the quality of learning processes.

Data used. The documentary survey, on a sample of 10 secondary schools in the province of Syracuse, found information on the use of physical spaces, the disconnect between internal and external evaluations, and the methodologies applied, regarding physical environments (school construction). The analysis found that interventions regarding the physical space involved, predominantly, the management and optimization of energy consumption about 13%, replacement of fixtures or redoing of works present in 25 %) of cases, while only 30% took care of the organization of spaces with practices for the usabilitỳ of spaces and/or optimization and rationalization works. The existing mismatch between the skills of internal resources regarding the organization of school buildings, understood as containers, and the skills related to the use of learning environments (methodologies) is 60% wide, especially in those addresses that entrust external personnel with the "management" of the physical space. The documentary survey of which partial results have been outlined will be supplemented, in September 2023, with the administration of questionnaires to sample schools for the collection of additional data useful for comparing the correspondence between the innovative processes initiated and the level of awareness in the educational spillover and use of spaces, especially, with reference to methodological choices in learning pathways.

Method. To conduct the survey, qualitative and quantitative, documentary analysis was done with comparison of data published by schools (Scuola in Chiaro and school websites), surveys, supplemented by focus groups, to support metacognition processes between assessment (SNV outcomes), self-evaluation (RAV) and learning environments, about processes and resources useful to define a taxonomy in terms of quality of physical environments and available resources (human and instrumental).

***Keywords:*** personalization, methodologies, competency-based assessment

# THEME 9. ITALIAN, MATHS AND ENGLISH: TEACHING PROCESSES THROUGH INVALSI DATA AND TOOLS

**ORGANIZER: INVALSI**

**COORDINATOR: VALERIA F. TORTORA**

**25TH NOVEMBER: 4.00 P.M. – 6.30 P.M. {ROOM 2 – TEACHING SESSION 4}**

---

## Connections And Inferences To "Unmask" The Text: A Vertical Route

**Anna Maria Moiso – Roberta Strocchio – Giuseppina Maria Sfasciamuro**

We got the idea for this project from the INVALSIopen videos curated by prof. Viale. These clearly illustrate that it is not enough to decode single words or sentences to understand the meaning of a text, but rather do a "textual immersion", where the text is understood through mental representations of the written text. "textual immersion" is a skill that can be used for all disciplines, because it implies knowledge of specific words and the intuition of what is omitted. Following prof. Viale's suggestions, we see modern communication as an ocean that can be approached in two different ways: as surfers who wait for the right wave, or as divers with heavy gear, going past the waves into the depths of the ocean to record, study and understand linguistic connections. Our path requires teachers to go beyond the concept of communication surfers The casa degli insegnanti has been working for years with different schools to build a vertical curriculum for language education and reading comprehension. In this project we tried to verify, with INVALSI tests, how the students in three school grades (elementray, middle and high school) develop competence in identifying two fundamental linguistic elements: connectors and references.

Object and hypothesis of the research. The educational path concerned text linguistics: analyzing the text as a product of communication, as a "structured intertwining of verbal and extra-verbal meanings" (Angela Ferrari, Linguistica del testo, Carocci editore, p. 13) helps understand the communicative product in depth, providing knowledge and analysis tools capable of rigorously describing the structure of the text.

The identification of connectives and the recognition of what is unsaid through inferences was developed with a spiral teaching, using very simple texts in the first two years of primary school, and then proposing increasingly complex texts in content and form, but respectful of the development students' cognition. A first theoretical part analyzed connectives in all their linguistic expressions, emphasizing that they are not just conjunctions and prepositions, but also verbs, interjections, entire sentences. The aim is to consolidate the identification of the function of connectives for students, because the lack of this skill compromises both comprehension and textual production. This is why we talked about the semantic-syntactic function, we reflected on the variety of connectives and their nuances of meaning and on the logical relationships that explain the development of a text. In the three-year upper secondary school year, we pointed out how connectives combine to create the style of a certain literary text. The second, indispensable element for understanding the text is the identification of what is left unsaid, of what is not explicitly expressed in the communication; in fact, much information is withheld or unexpressed because it is the reader who must discover it, making use of his encyclopedic knowledge and above all of the context. As Maryanne Wolf argues in Proust and the squid, we were not "born to read, but we are endowed with an extraordinarily plastic brain" which allows us to create "sophisticated connections between neural structures and circuits" and which help to form knowledge on non- repetitive, but creative and effective that always give rise to new thoughts and discoveries. Our culture is therefore the daughter of reading, but of an "understood" reading.

Data used. In the first phase of the project, the teachers selected a suitable number of INVALSI items pertaining in particular to the macro-aspects of reading comprehension and to the areas identified by the reference frameworks of the INVALSI tests in Italian:

A. MACRO-ASPECTS

1. locate and identify information within the text;
2. reconstruct the meaning of the text, locally or globally;
3. reflect on the content or form of the text, locally or globally, and evaluate them.

B. AREAS: Textuality and Lexicon and semantics

Method. The phases of the project were the following:

- for primary school: based on the research of Prof. Lo Duca we started, in the first two classes, with the analysis of the verbs, considered as the backbone of the text. We used nursery rhymes, fairy tales and very short texts to discuss the form and function of verbs
- in the second cycle of primary school we used short narrative and descriptive texts starting to identify the development of the text (coherence) and focusing on the most common connectives (and, but)
- in lower secondary school we administered INVALSI items relating to the areas at the basis of our research and we tried to broaden both the semantic-syntactic knowledge and the analysis of the function of the most common pronouns to understand their function within the text
- same path in secondary school, however choosing both more complex INVALSI items and therefore more suited to the students' abilities and evaluating the skills achieved with an ad hoc text that would allow the analysis of all the semantic, syntactic, inferential difficulties present in the text

In the last two years of secondary school, two literary texts were proposed, one poetic and the other in prose, to measure not only textual comprehension, but also the ability to identify the author's stylistic choices

In summary: in primary school we paid greater attention to respecting the childresn's linguistic knowledge, not to stuff them with abstract rules but trying to strengthen their ability to express themselves with suitable exercises

In the following school levels we stimulated reflection on language to develop the metalinguistic competence of the students.

In all grades of school we used teaching methods based on the development of expressive skills rather than memorization

Results. The experience was positive and motivated the students who worked both individually and in heterogeneous groups. The work has stimulated interesting discussions, that have allowed for changes of perspective and rethinking of the different ways of interpreting a text.

Surely, the language ability has improved and students were not bored in carrying out the various activities.

***Keywords:*** text comprehension, connectives, cohesion, coherence, vertical curriculum

## Beyond the sentence. The valential model, the complex phrase and INVALSI item

**Alessio Trevisan – Federica Ferrini**

Introduction. The complex phrase's context, in the Italian didactical tradition, is usually taken into account either during the last day of the Middle School or in the first two years of the High School, which leads to a superficial knowledge of the grammatical aspects that are built on a larger dimension than the one of a simple sentence. Furthermore, this didactical action affects both the comprehension and the written production, as the reciprocality between the reflection on language and and language usage. In the *Indicazioni Nazionali* of 2012 year of primary school, the student is able to recognize the nuclear phrase;

- at the end of the fifth year of primary school, the student is able to recognize form and usage of the most important conjunctions of the complex phrase;
- at the end of the third year of middle school, the student knows how to recognize both the logical and syntactical structure of the complex phrase, at least in a first grade subordinate clause, and the connectives, the punctuation marks in their specific function.

In the Italian traditional linguistic didactics, one can find a very rigid separation between a simple sentence and a more complex clause, as if their inner structure was not similar. The valential model, instead, leads to a different approach while teaching Italian grammar: it has to be more cyclic, circular and less sequential; the functions of the different syntagms can be taken also by the simple sentences. As a consequence, one can see the insufficiency of the traditional and rigid definitions which have always characterized the didactics of the Italian language at school: for instance, it is clear that the subject of a clause could be a sentence and not only a nominal syntagm. The vertical curriculum of the Scuola Madre Mazzarello, created thanks to the disciplinary epistemology and to the reflection on the so called Nuclei Fondanti, has allowed significant work on the analysis of complex sentences, especially in secondary school. In fact, the systematic use of the valential model while reflecting on the complex sentence has led the students to reflect on the graphic representation adopted by the INVALSI, which indeed shows the hierarchy of the period but does

not give the different features of the subordinate sentences on the basis of the function assumed in relation to the nuclear phrase. The following are the essential references for linguistic reflection and disciplinary epistemology:


- Berruto G., Cerruti M., *La linguistica. Un corso introduttivo*, UTET Università, Torino, 2011.
- Colombo A., Graffi G., *Capire la grammatica. Il contributo della linguistica*, Carocci, Roma, 2017.
- De Mauro T., *L educazione linguistica democratica*, a cura di Loiero S. e Marchese M. A., Laterza, Roma-Bari, 2018.
- Ferrari A., Zampese L., *Grammatica: parole, frasi, testi dell italiano*, Carocci, Roma, 2016.
- Pognante S., Poletti Riz J., *Educare alla lettura con il WRW – Writing and Reading Workshop. Metodo e strumenti per la scuola secondaria di primo grado*, Erickson, Trento, 2022
- Prandi M., *L analisi del periodo*, Carocci, Roma, 2014.
- Prandi M., De Santis C., *Manuale di linguistica e di grammatica italiana*, UTET Università, Torino, 2019.
- Sabatini F., Camodeca C., *Grammatica valenziale e tipi di testo*, Carocci, Roma, 2022.


Object and research hypothesis. The object of the present research is the use of data and items of the INVALSI to verify the epistemological theory that guide the didactic praxis of the first and second grade secondary school of the Scuola Madre Mazzarello of Turin. More in general, it aims to assess the validity of the didactic action of the group of research. This investigation dealt with grammatical questions related to complex sentences to show how the valential grammar not only confirms but also describes in detail the structure of the complex sentence more carefully than the block diagrams system, as the one proposed by INVALSI.

Data used and method. The data used in the experimentation revolve around the items below, compared with the results obtained by first and second year secondary school students. The items are created both by analyzing the questions related to the complex sentence which are available on the digital platforms and by writing questions based on the Frameworks of Reference (Quadri di Riferimento).

The proposed exercises are of the following types:

- recognition of the hierarchy of a complex sentence given its graphical representation in blocks;
- recognition of the function of argomento, circostante or espansione given by the subordinate clauses;
- recognition of the main logical and syntactic functions given by the circumstantial subordinates;
- production of complex sentences given their graphic representation in blocks or through a radial pattern according to the valential model.

Method. The survey method consists of three phases:

1. theoretical approach to the valential model for the description of the structure of the complex sentence: especially its validity and didactic implications;
2. analysis of the exercises proposed given the goals of the research and commentary;
3. presentation and commentary on the data obtained, with reference to epistemological assumptions and didactic praxis.

Results. The results that offer the possibility to identify:

- the high descriptive capacity offered by the valential model also for the analysis of the complex sentence;
- the need to rethink the didactic of the Italian language shaping it in a circular rather than sequential form;
- the contribution of INVALSI data and research as a constant contribution to test didactic experiments and praxis.


***Keywords*:** complex sentence, valencial grammar, linguistic reflection


## The floor to the students: in-depth interviews on reading comprehension tests during the field trial


**Antonella Vendramin – Francesca Resio – Antonella Mastrogiovanni**


Introduction. Mixed method surveys, and thus surveys in which qualitative research techniques are supported by quantitative researches and vice-versa, appear to be particularly useful, even in the field of educational research, mainly when the research object is as complex as reading comprehension.

In fact, to understand a text is necessary to process information on two levels. The analytical one, relating to the linguistic representation of the specific elements of the text, and the global one, which concerns the action of relating the elements of the text to other knowledge held by the individual. (Zanetti, 1999; Zanetti, Miazza, 2004).
Flanking qualitative analysis to support psychometric analysis for the construction of standardized tests of reading comprehension appears increasingly relevant and necessary to report such complexity.
Research object and hypothesis. The Covid-19 pandemic has brought out, even more strongly than in the past, the need to deepen also on a qualitative level the analysis of the tests produced by INVALSI for primary school. The students involved in the latest national surveys have in fact followed training paths less "linear" than in previous years. These students alternated periods of in-presence training with periods of distance learning, with longer or shorter periods of suspension of training activities based on health and territorial specificities. Wanting, therefore, to deepen the analysis of the reading comprehension tests carried out in the field trial, also on the basis of evidence derived from the experience gained so far and from the comparison with teachers who are experts in the field, the need has arisen to complement the usual quantitative analysis of the test, which must comply with all the psychometric parameters of standardized tests, with a qualitative analysis that would make it possible to highlight other aspects of the test itself and to read the quantitative results obtained from another perspective. In particular, two interview protocols have been prepared, one for each elementary school grade involved in the national surveys, which allowed the acquisition of valuable information in order to gain greater awareness with respect to the comprehension of the proposed texts and to the way primary school students relate to the text and the entire test. Through the two prepared protocols it was possible, in fact, to deepen the actual comprehension of specific passages of the text and of the text in general, a factor that allowed for a more precise focus on the answers given and the critical issues encountered by the students during the performance of the test itself.
Data used. This paper therefore examines the data arising from the qualitative in-depth analysis carried out in the field trial through interviews and the quantitative sample data found during the field trial and for the same test in the national survey phase (main study). In particular, the purpose is to see whether the outcomes of the qualitative analysis of the protocols produce evidence that constitutes added value in the test construction process.
Method. Interview protocols used in the classes involved in the in-depth qualitative study (7 for Grade 2 and 5 for Grade 5 classes, respectively, located in the 3 geographic macroareas: North, Central, South) are then explained. Specifically, regarding Grade 5, special attention is paid to illustrating what emerged regarding the narrative text contained in the test.
Results. The analysis of the interview protocols and data tests from the two school grades can be useful insights for INVALSI in the process of constructing reading comprehension tests, but these qualitative elements, which will be collected over the years through the administration of these protocols, may in the future feed into additional in-depth tools that INVALSI will make available to schools.

***Keywords*:** reading comprehension, educational research, qualitative analysis, interviews

## A trip beetween Italian vertical curriculum and INVALSI data.

**Alessia Ieva – Alessio Trevisan – Federica Ferrini – Giuseppina Maria Sfasciamuro**

Starting from the academic year 2021/2022, Madre Mazzarello School has started a process of vertically designed curriculum that involves the Primary, Middle and High School. This is a rare endeavor in the Italian educational landscape, which is characterized by a structural challenge in communication between the different school degrees. Some primary school teachers and a group of Italian teachers of Middle and High School have developed the design of vertical curriculum. The process of refining the vertical curriculum, in addition to the needs for internal training on the competency-based learning, required departments to engage in discussions and share epistemological orientations, teaching methodologies, designs, and practices that demonstrated a clear synergistic approach. INVALSI data, materials, and documents, available on the Institute's managed digital platforms, allowed the school to compare and evaluate its work using scientifically valid and validated documents.

To assess the validity of the vertical designed curriculum and the teaching choices derived from it, we examined the results of standardized tests taken by students who attended our institution (second and fifth grades of primary school) and who completed at least two educational levels within our school (e.g., middle school and high school, primary and middle school). Then, we compared the results to the national data.
We compared the national results to the results obtained in May and June 2022 by students in primary, middle and high school. The table below provides a concise and indicative overview of the item that served as reference points for the analysis conducted by the research group.

| ***Degree*** | ***Item – Reading Comprehension*** | ***Item - Language Reflection*** |
|---|---|---|
| 02 | Non fiction text "Se tu fossi un orso" – Prova 2019 | Similar items to the following:<br>*In ognuno dei gruppi di frasi indica la frase che non va bene (Prova 2017)* |
| 05 | Non fiction text "Scienziati in casa" – Prova 2019 | Similar items to the following:<br>*Qual è il soggetto della frase: «Manca ancora una settimana alla fine dell'anno scolastico»?*<br>(Prova 2009) |
| 08 | Non fiction text "Dalla filosofia naturale alla scienza" – Prova 2019 | Similar items to the following:<br>*Nelle frasi che seguono individua se il pronome che ha funzione di soggetto oppure di complemento oggetto (Prova 2016)* |
| 10 | Non fiction text "Alla scoperta dei documenti che non ti aspetti" – Prova 2019 | Similar items to the following:<br>*Leggi la seguente frase. "Maria ha ricevuto assieme a me il premio dell'amicizia". Solo una delle seguenti affermazioni è vera. Quale?* (Prova 2019) |
| 13 | Non fiction text "Uno che non dimenticava nulla" – Prova 2019 | Similar items to the following:<br>*Nella frase "permette dia vere su di lui un controllo relativamente facile" (evidenziata nel testo) l'avverbio "relativamente" è un modificatore* (Prova 2019) |

The research method unfolds in three phases:

1. Reconstruction of the didactic-pedagogical and epistemological foundations underlying the development of the vertically designed curriculum, with particular attention to grammar of dependency as a key factor for both language learning and textual comprehension and production.
2. Description of the working process that led to the construction of the Italian vertically designed curriculum and its overview.
3. Presentation and analysis of the emerged data, with reference to both the implied theoretical assumptions and the vertically designed curriculum.

In conclusion, a metacognitive emphasis will be placed, briefly reflecting on the inherent value of such experiments, which can be seen as small research-action experiences that are pivotal to a teacher's professional development.
The results obtained open the way for an ongoing evaluation of the vertically designed curriculum. Specifically, they allow to

- confirm the validity of the assumed epistemological orientations in the construction of the foundational core design;
- verify the effectiveness of the teaching experiences proposed to students for the learning of linguistic and metalinguistic competencies.

***Keywords*****:** vertical curriculum, valential grammar, competence-based design learning

# Implementation, improvement of INVALSI data, comparison with international OECD-PISA research for a European dimension of teaching

**Pompea Funiati – Lucia Funiati**

Introduction. The purpose of this contribution is to show how has developed in our school a learning measurement system in line with international best practice, in particular the Finnish, in all subjects covered by both national and international surveys.
Within the disciplinary department of Italian, the need was felt to work on an Improvement Plan (PdM) aimed at improving results in INVALSI tests in Italian grade 10 which have shown strong criticalities in the national surveys especially after Covid-19 pandemic.
An analysis of INVALSI results shows that linguistic difficulties, also due to educational poverty, and text comprehension difficulties affected answers given by students in the Italian test, compromising the results of main and basic skills to ensure active and responsible citizenship in an increasingly complex and globalised society.
The need to improve the teaching/learning processes and students' skill levels motivated the teachers of the Humanities Department to submit to the EU the project "ToM (towards mindfulness)" action k1 - Erasmus plus, aimed at carrying out an in-service training experience for the institute teachers involved in teaching basic Italian skills in a Finnish school (a country with good practices and many examples of a high-quality school system).
This decision stems from the need to get to know the Finnish school system as it records excellent student performance in OECD-PISA standardised tests in reading literacy for 15-year-olds students, in order to get to know its characteristics and the possibility of replicating some of them in the Italian school system and especially in our school.
Visits to several Finnish schools were also intended to explore another qualifying point of the Finnish school system: their self-evaluation process of individual schools and results, because the evaluation data are used for improvement plans.
To cope with this educational emergency, the Humanities Department worked on INVALSI test results improving and implemented the INVALSI methodology in curricular teaching practice, stimulating teachers' thinking on the possible relationship between INVALSI test experience and skill-based teaching, through the reading of INVALSI test results return.
In previous years in the institute there was widespread scepticism towards national tests, however with the project 'Improving the outcomes of INVALSI tests', the students' approach to external system assessment has completely changed.
Thanks to the project's spin-off, in our school all students made INVALSI tests. The project aimed to foster pupils' understanding of the purposes of standardised tests. The main objectives of the project were:

- Promoting pupils' understanding of the characteristics of the Italian standardised tests, the areas covered by items and the skills being assessed;
- Supporting pupils in developing test-solving strategies;
- Homogenise INVALSI tests results, in particular reducing the percentage of pupils in levels 1 and 2.

Subject and research hypothesis. The exchange of information and experiences on common problems in the education systems of EU Countries, particularly in Finland, has brought our educational institution closer to European standards and the aims to Erasmus+ programme.
Visiting some Finnish schools helped get to know and discover interventions on assessment strategies, methods and tools, designed to foster the professional development of teachers responsible for basic learning, in order to achieve improvements in outcomes and processes and reduce territorial disparities, through innovative teaching methods and metacognitive assessment activities.
In other words, it is essential to create a virtuous cycle starting from the use of the results of an external learning assessment to draw up a self-diagnosis and design, for both students and teachers, more targeted training actions based on innovative and inclusive methodologies.
Data analyzed:

- quantitative and qualitative analysis of INVALSI items;
- identification of critical areas and processes;
- didactic and methodological path of improvement to help students overcome cognitive obstacles in solving the items and verify the effectiveness of the actions;
- data from the initial and final parallel tests.

Method. Strategies to develop students' learning processes and reduce/decrease gender gaps by redesigning teaching thanks to INVALSI test results:

- initial training of Italian teachers in a two-year course on areas described in INVALSI reference framework;
- administration of an initial and final parallel test with a set of INVALSI items;
- based on the results of the initial survey, improvement actions are planned;
- post-school training modules for pupils belonging to different classes;
- use of teaching and methodological materials made available by INDIRE website;
- use of CBT tests to compare results using IT tools.

Goals attainment to strengthen skills was verified through the administration of initial and final tests and the evaluation of the gaps between entry and exit learning levels in order to assess pros of the pathways carried out and of the teaching/learning method adopted. The analysis of the results positively highlighted what has been done. Analysis considered a quantitative dimension (correct, acceptable, and unacceptable answers) with the help of calculation software. For the quantitative analysis, a worksheet was developed for grouping these areas: text, vocabulary and grammar.

Results:

- Developing and disseminating an evaluative culture by measuring learning thanks to national standardised tests, with the aim of achieving greater transparency of results and a greater correlation between results and investment;
- Spreading the culture of self-evaluation aimed at preparing plans for improving both didactics and the curriculum with the introduction of innovative learning method paths;
- Being aware of improvement actions, ensuring coherence and integration with curricular activities;
- Improving the correlation between skill certificate made by the school and the one certified by INVALSI.

***Keywords*:** teaching practices, school improvement, learning

## Students Preparing for an INVALSI Test: A Vertical Peer Education Experiment

**Roberta Strocchio – Aessia Ieva – Anna Maria Moiso**

Madre Mazzarello Institute includes primary school, secondary school and high school with three fields of study: languages, science and human studies. Starting from the academic year 2020/2021, the school has completely renewed its vertical curriculum. Together with the association "Casa degli Insegnanti", which has been working with the school for more than ten years in the field of updating teachers and in the design of educational courses, the vertical department of Italian Language has developed experimental paths aimed to improve verticality. The strength of the vertical curriculum consists in promoting competence over time and in designing teaching experiences in long term. In addition, we have wanted to combine the effectiveness of peer education, offering a didactic experience, designed together with students at the second year of high school and proposed by them to secondary school students and their fellow first-graders.

The research. The goal of our research was to strengthen the mastery of language of our students, by referring to some of the skills evaluated by INVALSI tests:

- reading competence, understood as understanding, interpretation, evaluation of the written text.
- grammar knowledge and skills.
- Semantic-lexical competence.

We have led second high school year students from all the three fields of study to build an INVALSI test for the students of the first high school year and for those of the third year of secondary school. The classes involved in the project were five and they have the same vertical curriculum, developed and shared by all the Italian teachers from different school grades.

Data used. In the first phase of the project, teachers selected an appropriate number of INVALSI items relating to the following areas, identified by the INVALSI Test Framework of Italian:

- lexical and semantic (meaning relationship between words; semantic fields and lexical families; polysemy; figurative uses and main rhetorical figures; idiomatic expressions; structure and use of the dictionary);

- textuality (signals of text organization and cohesion phenomena: anaphora, connective, punctuation, etc.; pragmatic aspects of language like speech phenomena, speech functions, etc. ).

Method. The project phases were:

- administration of INVALSI items;
- study by secondary school students of the Framework of Reference of Tests INVALSI of Italian;
- reading and analysis of the story by Jorge L. Borges, *La casa de Asterión*, included in the 1949 collection L'Aleph (translation by Francesco Tentori Montalto)
- construction of the text comprehension test INVALSI;
- administration of the test to the first classes of the high school and to the third classes of the secondary school of first degree;
- reflection and evaluation of experience

In September the project was presented to second year high school classes involved and they were subjected to a battery of INVALSI items related to the areas "Lexicon and semantics" and "textuality".
Later, thank to the speech made by Prof. Annamaria Moiso, from “Casa degli Insegnanti”, who in the past collaborated with the preparation of INVALSI tests, students were introduced to the logic behind the construction of the INVALSI tests.
Then, managed by the teachers and divided into groups, students were invited to build first-hand items like INVALSI on the story of Jorge L. Borges, previously read and analyzed in class.
The items prepared by the various groups (each of which focused on a specific aspect of the two selected areas) were discussed in class and, in some cases, modified or rethought.
Finally, the test was structured by choosing the best items for clarity and effectiveness; it was tested on the same second year high school classes. After that, students proposed it to the younger children.
The second-year students also took care of the correction, which was followed by a reflection on the results: most frequent typology of errors; their possible cause; possible defects in the formulation of the questions.
The material produced was then supplied to the teachers of the secondary school of first degree.
The results. Through experience, students' awareness has increased significantly, especially in linguistic and metalinguistic skills. The moment of peer education was also precious, because, as a student pointed out, "to explain things to others, you must have understood them well, but I mean, very well".

***Keywords*****:** peer education, verticality, reading compliance, lexical compliance, semantic compliance, mastery of language

# THEME 1. A DIFFICULT EDUCATION: CHALLENGES OF THE SCHOOL SYSTEM TO ENSURE SOCIAL INTEGRATION, SKILLS/KNOWLEDGE AND SOCIAL MOBILITY

**ORGANIZER: INVALSI – ESPANET**

**COORDINATOR: EMMANUELE PAVOLINI**

**25TH NOVEMBER: 4.00 P.M. – 6.30 P.M. {ROOM 3 – RESEARCH SESSION 16}**

---

## The top, the bottom, and the average. The trinity of school ability peer effects in Italy and Norway

**Emanuele Fedeli – Solveig Topstad Borgen – Moris Triventi**

Introduction. The study of social influence in adopting similar actions, behaviors, and choices is deeply rooted in the social sciences and cuts across any line of research. Contributions to the peer effect are numerous in education studies and usually analyze whether and how much classmates or schoolmates influence cognitive skills, social-emotional skills, and educational choices in the short or long term. (Angrist, 2014). What emerges is that peers play a very important role in influencing the life paths of students by influencing the vision of the world, the dynamics of friendship, or simply the study. To date, most of the literature has focused on the exposure of students to the ability of their peers very often not considering that some students are exposed to the presence of highly skilled or poorly skilled subjects (Lavy et al. 2020). Furthermore, these studies often look at a single country or region, neglecting the fact that educational systems are different and that the peer effect mechanisms, although universal, can vary between systems due to the organization of the school system, due to formal and informal mechanisms which prolong or less exposure to the same peers in subsequent years.

Goal of the research. The peer effect mainly consists of two branches of research. One focuses on how the composition of the classroom or school requires adaptation of teacher practices, both in terms of teaching and classroom management. Another one focuses instead on the interactive dynamics between classmates; (Epple and Romano 2011). According to the first strand, classroom management should synchronize with students' cognitive and non-cognitive needs and requirements based on their sociodemographic characteristics (Lazear 2001). In the "good or bad apples" model (Lazear 2001) the presence of low or very competent students leads to a change in the needs of the class. Teachers compensate by spending more time and effort with low proficiency ones, for example, thus reducing the time and attention devoted to other children. When held against these compensatory mechanisms, there is the actual peer effect due to mechanisms of social influence. Simply put, a student may be positively or negatively attracted to having very good or bad classmates. However, there are no contributions in the literature that analyze the effect of peers in a comparative perspective by evaluating or even breaking down this effect not only by looking at the average but at the top or bottom of the "ability" distribution.

In this work we fill these gaps by comparing two countries such as Italy and Norway which present quite different aspects and 3 institutional mechanisms. First, Italy is an example of an education system with early tracking although not binding for university access while Norway is an example of a late tracking system. Second, elementary school lasts 7 years instead of 5 as in Italy. Third, the transition between elementary and middle school is characterized by a de facto continuation with the same peers at school level, in contrast to what happens in Italy where there is a substantial and important reallocation. All these institutional differences make the comparison interesting beyond the simple interest of verifying whether this peer effect differs or not. Finally, we look at whether this effect is modified by the social origin of the students. The hypothesis is that the peer effect is less relevant for subjects with a high cultural capital but is important for those with a low level. In this perspective, a person with low cultural capital may benefit or be harmed by more than one peer effect depending on the perspective.

Data and method. We use Italian and Norwegian administrative data and perform 3 cohorts from elementary to high school in 3 academic years 2011/2012 2012/2013 2013/2014. Our dependent variables are the standardized tests in mathematics and Italian and the choice of school address. Control variables are individual-level sociodemographic baseline antecedent to treatment and we exploit variations over time in the same school for causal identification. As templates, we use OLS with school fixed effects.

Results. The results suggest that as middle class ability increases there is a negative effect on skills of 0.04 (sd) in both Italy and Norway. This result is in itself important because it suggests that the peer effect is

similar in two countries that are quite institutionally different. This is also confirmed when we look at extremely or low competent students. The effect is the same between the two countries, suggesting a similar pattern in terms of common peer exposure. The interaction with socioeconomic origin also appears to be of the same amount and of the same. Similar patterns are also found when we look at track choice even if the sign is opposite, i.e. having subjects with higher index competence to choose the academic address in high school. We are currently working on how some formal and informal institutional aspects such as exposure to the same peers over time may or may not have an effect. Preliminary results indicate that prolonged exposure to the same individuals accentuates the peer effect

***Keywords***: peer effect, ability, socioeconomic divide, Italy, Norway

## Assessing the educational inclusion of students with a migrant background. Italy and Spain

**Patrizia Rinaldi – Giovanna Filosa**

While the benefits of host-country education for the first generation have been the subject of much attention in the literature, yet more needs to be written about how education affects 'life opportunities' in the second generation. In fact, the current debate on education systems within the European Union focuses on evaluating inclusive schools in the conviction that policies must be evidence-based and ethically oriented for this strategic sector like education. The hypothesis is based on the current school structure incapable of reducing the inequalities between the native component and the component with a migrant background of the student population. These inequalities, even exacerbated with the pandemic, were already present before Covid.
The research question is: can the modernisation of teaching methodologically and infrastructurally overcome this gap?
This article aims to determine how cultural diversity is problematised in the school context and how differences are evaluated at school.
For this reason, the document examines the capabilities and limitations of standardised tests to assess school settings that include students with a migrant background. Specifically, from a comparative perspective, it looks at two education systems, the Italian, and the Spanish, to focus on the certification and methods of evaluating the educational background of first- and/or second-generation migrant minors.
To this end, in Italy, the results of the *INVALSI* tests of pupils with non-Italian citizenship are compared with those of their native peers. In particular, the differences between 1st and 2nd generation natives and foreigners will be examined in the standardised tests of Italian and Mathematics (G8) in the historical series from 2012/13 to 2021/22 school year. Examination of the historical series is critical because it allows us to distinguish, from a descriptive point of view, the effects of Covid-19 on the school inclusion of a particular student target from those due to structural limits of the Italian school system. The INVALSI national surveys now represent an essential component of the national evaluation system and an essential thermometer for evaluating the state of health of the Italian school. One of the indicators of this state of health is given by equity and the capacity for inclusion, which can be operationalised by measuring the differences in learning between different targets of the student population. In the Spanish case study, we took two types of evaluations: one organised at the national level by the Evaluation Institute (*Instituto Nacional de Evaluación Educativa - INEE*) that is carried out at the territorial level every three years by the individual Autonomous Community and a diagnostic evaluation at the territorial level delegated to the individual Autonomous Community and carried out annually, focusing the analysis on the comparison of native pupils with pupils of foreign origin. This agency combines two strategies: general system evaluation and diagnostic evaluation. Both aim to gather information on the education system by assessing the competencies students acquire about their socio-economic and family context. The methodology has been based on analysing the rules of the main Spanish Educational Laws from 2012 to post-pandemic time, together with participant observation, interviews or document analysis. The tests are administered randomly to pupils at the end of the fourth year of primary school (10 years of age) and at the end of the second year of lower secondary school (14 years). The subjects covered, in turn, the eight basic skills of the curriculum. For internal

consistency in this document, we will examine the equivalent subjects of the Italian case: Spanish and mathematics. To make a comparative analysis, the following aspects are analysed:

- Treatment of the evaluation
- Evaluation characteristics,
- Promotion and qualification,
- Other significant differences,
- Conclusions and food for thought.

A quality school is among the goals of the 2030 Agenda for Sustainable Development (and assessment systems can make a major contribution to this), and among these goals is the reduction of inequality and poverty in all its dimensions, including education.To close the gaps often highlighted by evaluation tests, it would be necessary not only to promote overall excellence but also to identify bad practices so that necessary corrective action can be taken, understand and apply cross-cultural education Considering that parents' socio-cultural resources are associated with their children's educational outcomes, the evaluation of schools should also consider how inclusive the school is, i.e., they can guarantee similar performance results for students with Italian and Spanish citizenship and potential "new citizens".

***Keywords*****:** migration, second generation, inclusive education, inequality, Covid-19

## Grading standards across provinces revisited

**Emanuele Fedeli – Moris Triventi**

Introduction. In recent decades, the economics and sociology of education has devoted increasing attention to the grading practices adopted by teachers in the classroom and to any discrimination implemented by them in the student assessment phase. In fact, the literature is well established in identifying the presence of different grading practices territorially (Argentin & Triventi 2015) but also in the same schools (Lievore, Fedeli, & Triventi 2023) and also that these may reflect discriminatory attitudes towards some social groups - demographics such as girls, immigrants or low income. However, these contributions show some scope for action to improve the identification of grading practices and in the future to better understand the quantum of discrimination with important implications in the public debate. Think, for example, of the annual discussion on how much or not some Italian territories favor their students. In this work we revitalize a body of knowledge to propose an analysis protocol to study vote differences with observational data.

Goal of the research. In fact, these two lines of analysis, the one on grading practices and the one on discrimination, show two different assumptions in establishing the association between grading and testing. In the first strand, the test is a function of the grade, while in the second, the grade is a function of the test. Although they may seem technical, this divergence betrays a different theoretical and empirical estimate. Alongside this important divergence, other aspects insist on these different lines of research. In fact, the design of the research is often not clear enough and it remains difficult to understand the limit within which grading practice is discussed or elements that lead to grading discrimination are discussed, due to characteristics due to observed or non-observed productivity or to true attitudes and their own discriminators. In addition, these studies overlook some relevant aspects. The first is that a single standardized test does not fully measure ability but rather is exposed to a measurement error (Jerrim and Vignoles, 2013; Bynner, 2015). Secondly, it is no less important to be careful when tests and grades are taken. If grade and test are taken in the same year there is a risk of bias as the student is influenced by the teacher and this can be reflected in the standardized tests creating an endogenous path with the result that we do not have a standardized test exogenous to the vote. Added to this is that all these studies do not conceptually and empirically take into account the incoming quality of students in the first educational cycle. The bias due to the mutual influence between grade and test and/or the lack of reflection on the quality of incoming students can generate incorrect estimates.

In the light of these two lines of research on grades and the different nuances, we propose a reformulation of the identification of assessment practices that comes closest to the data-generating process, i.e. that the grade depends on individual cognitive skills to which we add corrections for the bad day effect and for the average incoming quality control. We then apply this frame to the differences in practices between

provinces in Italy to evaluate whether and how much grading practices really differ between provinces. we choose the province because it is an intermediate institutional level characterized in most cases by municipalities with various similar characteristics.
Data. The INVALSI data represent a significant example since grades are collected after a certain period of time and not in the months immediately following the start of the study cycle, whether in elementary, middle or high school. This requires identification according to a new approach and correction using the previous standardized tests.
Method. Focusing on middle schools and using five cohorts of enrolled students, we estimate assessment practices through a simple linear regression model. In a second step, we adjust the results for prior and concurrent standardized tests using them as instrumental variables.
Results. What we observe in both middle and high schools is a marked presence of more generous assessment practices in some areas of Italy, especially in the south, but there is a significant variation within the macro-areas. In fact, the most generous practices are adopted in Sardinia, southern Sicily and in some areas of northern Calabria. However, some areas of central Italy have similar practices in contrast to the north east of the country characterized by stringent grading practices. It should be noted that thanks to our design we correct for any errors in the test measurements and in the second model we check directly for the previous test. In this way we keep the quality of the incoming student flow under control. From here there are some considerations to make. The first is that the practice of generous grading simply insists on those areas with a low level of skills entering middle school. The second is that there may be other mechanisms at play given the context. In fact, on the one hand it could be the quality of the teachers in elementary school that generates a stock of low-skill students but on the other there are obviously contextual factors that can affect the effectiveness of teaching.

***Keywords*:** grading standards, provinces, instrumental variables

# TEMA 10. SCHOOL 2.0

**ORGANIZER: INVALSI**

**COORDINATOR: CHIARA TAMANINI**

**26TH NOVEMBER: 9.30 A.M. – 11.30 A.M. {ROOM 1 – TEACHING SESSION 5}**

---

## INVALSI Tests And Gamification: Some Experimental Incidences

**Luigi Umberto Rossetti**

This paper aims to report the results obtained from the experimentation carried out in a second-grade school in the area of improving INVALSI results through "gamification".

Introduction. Gamification' is the use of typical game elements and game dynamics in non-game contexts, such as educational, work or social settings, in order to engage, motivate and stimulate the interaction of individuals (Marczewski, 2017). It is an approach that exploits typical game elements, such as points, rankings, badges, missions, challenges and rewards, to create an engaging and fun environment that encourages active participation and learning (Petruzzi, 2021). In an educational context, it can be used to make lessons more engaging, motivate students to achieve certain learning objectives and provide immediate feedback on performance. The effectiveness of gamification depends on the design of game dynamics, clarity of objectives and rules, and the ability to engage and motivate students. When applied correctly, it can foster learning, improve motivation, increase interaction and promote a sense of achievement and gratification (Newmann, 1992).

In the digital sphere, there are various platforms for the realisation of gamification activities, some of which are specifically structured for gaming, while others have game tools or dedicated applications within them. Few are those CBT (computer based tests) platforms that, although structured on gaming, allow for individual and group evaluative data returns that can be used as INVALSI exercises. It is precisely this latter aspect that laid the foundations for the realisation of this experiment reported in this paper.

Research object and hypothesis. Gamification does not originate as a learning theory, therefore its pedagogical assumptions are not systematised, however, having the purposè of motivating students to learn, it finds its foundation in psychological theories of motivation (Maestri, Polsinelli, Sassoon, 2018). The main objective of the research was to obtain a strategic information base relating to three specific variables measured before, during and after the experiment.

1. Context and environment of implementation;
2. Individual approach to the trials;
3. Final evaluation and individual feedback.

The research course was aimed at the second classes of a secondary school in the Campania region located in the capital city of Benevento (Ipsar Le Streghe - Professional Institute for Hotel and Catering Services). The innovative action entitled "Mi metto in gioco" - The game as a methodology and tool for engaging didactics, was carried out in April 2023, prot.llo no. 3369/E, as an exercise activity in view of the INVALSI tests of the second classes through gamification.

Research question:

Q1. The starting research question was an attempt to understand whether it is possible to use gamification in the INVALSI context and whether it can generate an improvement in the CBT results of the students themselves.

Q2. How does the student's approach change psychologically in terms of motivation, interaction and gratification.

Research methodology. The research methodology used was mixed using the descriptive method, which is the most common form of investigation, the quantitative nature of which allows the information gathered on the students to be statistically inferred. Whereas, in order to understand the phenomenon inductively, the disguised Observation technique was used, which enabled the acquisition of participants' reactions, and the Interview technique in order to capture behaviour and motivations for action through a focus on subjects and not on variables. The main idea on which the research was structured is that of better defining a situation, attitude or behaviour of a group of people on certain aspects. Particular attention was given to the choice of the administration platform, which had to have the following characteristics:

- User simply;
- With an internal gamification structure and with real-time visualisation of the challenge;

- Able to enter INVALSI questions as a competition;
- Able to generate individual and group reports at the end of the competition.

The structuring of the research took into account a series of investigations:

1. Student information context
- Behavioural analysis in the face of gamification communication;
- Analysis of attention when explaining the rules of the game.
2. Environmental/organisational context
- Analysis of the behavioural variation of students in the organisational layout;
3. Game execution
- Modalities and platform for gamification;
- Analysis of student motivation, interaction and gratification.
4. Final considerations
- Consideration of the experimentation and analysis of the students' reactions to the discovery of having carried out a competition on INVALSI tests.

Method and data used. The platform chosen for gamification was 'Socrative' via the 'Space Race' administration option. This tool, which is free of charge, is structured on the gamification of a space race in a fun way that engages students by allowing them to work together on a quiz (INVALSI test from previous years).

- Each student is assigned to a group (identified by a colour) anonymously and participates individually in the competition, which consists of answering a test correctly. Each partner is a potential competitor and with each correct answer the platform advances the team towards the finish line (there is no possibility of identifying the student/result association during execution);
- The platform provides the option of choice in relation to the number of groups to be made, choice of avatar, shuffling of test questions, shuffling of answers, option of feedback on the individual answer and final result showing the grade in tenths;
- The test is based on INVALSI questions from past years previously uploaded onto the platform;
- The game competition, being second classes, covered the disciplines of Italian and Mathematics;
- The team's progress was projected in graphical format in "Real Time" on the digital board in the computer lab;
- The results obtained, at the end of the action, were shared with the teachers of the relevant disciplines for appropriate focus;
- The prize-giving ceremony, with the issuing of badges, took place in the individual classes.

Results. The results obtained at the end of the experiment were:

- General behaviour of the students;
- Students' attitude before, during and after the competition;
- Students' behaviour after the award ceremony;
- Behavioural differences in the approach to the INVALSI exercises;
- Final class report in the following formats
- Excel sheet with group and individual results and class scoring;
- Pdf for each individual student with highlighting of result and correct or incorrect questions;
- Pdf summarising all students with corrector.

Perspectives. The research intervention was included among the evaluations of strategic choices for the purposes of a new INVALSI exercise mode for improving student results. The future objective is to extend the research to the fifth classes of the same institute and to initiate a path of comparison with other school institutes as well as the possibility of creating a competition not only in the class but between classes of the same institute and/or external institute, in a sort of championship covering the entire school year. It is also hoped that the research can be validated and made universal so as to make it applicable in all school institutes and other contexts.

***Keywords:*** gamification, environment, platform, test, competition, questionnaire, INVALSI, badge, assessment

# Teachers' beliefs concerning the role of AI in mathematics education

**Camilla Spagnolo – Giorgio Bolondi**

Since January, chatGPT has entered massively into the communication of major newspapers (Aydın & Karaarslan, 2022), even though the chatbot in question (in the first beta version) was released by OpenAI manufacturing company on November 30, 2022. After remarks by Microsoft CEO Nadella in Davos, it is known that chatGPT will be integrated into many Microsoft products primarily Bing the company's flagship browser. The impact undoubtedly of the new tool has been strong, the final product is awaited.
The impact will involve all sectors of our society, but for now the most direct seems to be the training and cheating effect.
ChatGPT, at the moment, is not certified and does not cite sources where it gets its information. A central and strategic question arises: what will be the impact on educational systems, specifically in the teaching of science disciplines and mathematics in particular? Can teaching not take into account the new perspectives and challenges that AI poses, is one prepared for this "challenge"?
The EU recently drafted an interesting document "Ethical guidelines for educators on the use of artificial intelligence (AI) and data in teaching and learning" (European Commission, 2022).
Artificial intelligence in education (AIEd) is an emerging interdisciplinary field that applies AI technologies in education to transform and promote the instructional and learning design, process and assessment (Chen et al., 2020; Holmes et al., 2019; Hwang et al., 2020; Xu & Ouyang, 2022).
There are currently few in-depth studies on the topic of artificial intelligence in mathematics education (Knill et al., 2004; Garrido, 2012).
Artificial intelligent projects like "Eliza" (Weizenbaum, 1966), "Student" or "GPS" (Russel, 2010) attempted to produce agents which can solve general mathematical problems. One soon realized that this goal was too ambitious. Not only because of technical hurdles, but also because one had to realize that many problems are computationally hard and can not be settled in reasonable time. While this slowed the momentum of AI developments in education, it did not stop it.
The purpose of this paper is to describe an experiment whose goal is to bring out teachers' beliefs, and attitudes in an assessment situation in which artificial intelligence (AI) is involved. On one hand, it is well known that assessment is an important element in defining teacher identity, and the role of AI in assessment processes is a central issue. On the other hand, comparing teachers' behaviors and attitudes in these assessment situations allows us to investigate their beliefs and attitudes toward AI. The study involved 153 primary and secondary school teachers and a task on prime numbers from the Italian national standardized assessments (INVALSI). Results at the national level were used as a benchmark of difficulty and significance at national level. The question required production of argumentation to support a statement or its denial. Teachers had to assess 5 answers to the task, taking in consideration accuracy of content, accuracy of language, coherence to the task, coherence to the claim, completeness in calculation, and completeness in argumentation. After their assessment, it was disclosed to them that one of the protocols had been produced by AI chatGPT in answer to the task. None of the protocols (not even the one produced by chatGPT) contained an answer that was considered correct according to standardized assessment standards. After that, teachers tried to identify the answer produced by chatGPT and to give explicit reasons for their choice. In a second phase, 72 teachers were involved in focus groups and individual interviews to compare criteria and ways in which they had assessed and addressed choices. Results highlighted differences between primary and secondary teachers in recognizing the chatGPT answer and the criteria used to identify it. Particularly in primary school teachers, more attention was paid to the language used by students. Focus groups also showed that teachers, after participating in this study, consider incorporating chatGPT within their classroom practice. In addition, beliefs emerged with respect to what AI is and its role in mathematics teaching and learning. These beliefs and attitudes need to be considered as an element of teacher identity.

***Keywords*:** artificial intelligence, AI, chatGPT, beliefs, perception, mathematics education

# PLAY-INVALSI. Gamification through INVALSI questions

**Silvia Servili – Chiara Saletti – Ivan Graziani**

In the National Indications for the first cycle (2012) it is written: "In mathematics, as in the other scientific disciplines, the laboratory is a fundamental element, understood both as a physical place and as a moment in which the pupil is active, formulates his own hypotheses and controls their consequences, plans and experiments, discusses and argues his choices, learns to collect data, negotiates and constructs meanings, brings to temporary conclusions and new openings the construction of personal and collective knowledge. In elementary school it will be possible to use play, which plays a crucial role in communication, in educating people to respect shared rules, and in developing strategies adapted to different contexts."

With this last sentence we do not really agree because play should remain for secondary school students and even university students as well, as it facilitates the acquisition of concepts and procedures in a more lasting way, promoting learning. In fact, doing mathematics through even playful activities for children and young people, who thus become protagonists of their own learning goes through two elements that support each other. The first from the methodological point of view is the mathematics laboratory, that is, teaching mathematics as a living science and not one that belongs to archaeology, with an experience that continues with students experiencing mathematics as if for the first time. In this way, the lab allows children and young people to be truly active. The other is about the content to be proposed, which will not be trivial, we can propose games as well, but which are actually math questions and problems.

As early as some 2,400 years ago, Plato argued, "No discipline imposed by force can remain lasting in the soul. Therefore, do not educate children in the various disciplines by resorting to force, but by play."

Gamification consists of the use of game mechanics in contexts, which, however, have nothing to do with play, to achieve a certain educational goal. Bret Terrill (2008) defined "gamification" as, "Taking the mechanics of games and applying them to other web properties to increase engagement."

In school, "Gamification" does not mean "making the lesson playful," rather it is applying elements peculiar to games and video games in education to stimulate learning in traditional subjects.

"Gamification is the process of game-thinking and game mechanics to engage students and solve problems." (Zichermann, 2011) If we then want to bring pupils closer to mathematics in its essence, we necessarily need to confront them with problems.

The psycho-pedagogical assumptions that we can find behind gamification recall constructivism; in fact, the student is placed at the center of his own formative process, he is the builder of his own learning, which he acquires through direct experiences; this allows him to amplify his motivation, enabling him to internalize and memorize information in a meaningful and more long-term way.

In addition to constructivism, cognitivism can also be traced in gamification; in fact, all information from the outside is stored in the memory which reorganizes and reprocesses it to make it reusable in other contexts just as this innovative teaching methodology intends to do.

There are two different sources of motivation: one based on mastery, the other on the product (and its comparison with the achievement of others).

Evidence shows that learning motivated by grades alone is effective only in the short term; in contrast, a willingness to learn broader skills ensures retention of those skills in the long term (Terrill, 2008). In addition, not to be underestimated, gamification allows the acquisition of knowledge to be segmented into consecutive levels thus enabling the ability to operate in each learner's proximal zone with respect to growth, activity and self-esteem needs, maintaining high interest and self-esteem. Progress is visible from the very beginning and error no longer takes on the discouraging connotation but on the contrary arouses in students the impulse to try again other strategies and solutions.

One of the biggest victories that we as teachers can bring home is when we are teaching and getting our students to work and, if we say that time is up and must be closed, they instead ask to wait, that they want to finish ("Wait! Let us finish").

If we get to the point where they say, "Wait a minute let me finish," that is definitely one of the good victories we can remember.

Lucio Lombardo Radice (1916-1962) wrote, "Why, in order to check what the pupils have learned, do you not have an hour of games in class (instead of questioning)? To play well is to have a taste for precision, a love of language, an ability to express oneself with nonverbal languages; it is to acquire together intuition and rationality, a habit of loyalty and cooperation." Playing math with students, and not only with younger students, has many advantages, because it succeeds in downplaying a subject that is often considered

hostile, engaging students through its natural competitive dynamics; it induces students to ask and ask questions; and it creates a valuable informal opportunity for self-assessment, allowing the development of peer interaction dynamics (peer education), creating and reinforcing useful automatisms and metacognitive pathways. It also makes students think that "school" is not being done, causing them to reflect on what they are doing, to develop new strategies, and fostering interaction with peers.
Our project proposal stems from some experiential workshop activities, in which the teaching methodologies most responsive to the educational needs of "digital native" students of the first cycle (classes IV and V Primary and I and II Secondary I) were researched and used; moreover, the most suitable tools were identified to foster the development of transversal skills for life, following an educational path that would not end only in the classroom but could continue in different modes and places. So we thought of having students play on some INVALSI items, selected on Gestinv 3.0, dividing them into groups of four with specific tasks and names:

- the mathematician: with the task of keeping the group on the items to be developed and the timeframe to be met;
- the researcher with the task of searching together with the mathematician for the various strategies to be followed also by resorting to internet surfing;
- the critical friend whose task is to check the feasibility of the various proposals found by the first two, giving reasons for his or her choices
- the blogger with the task of keeping a kind of logbook updated, keeping track of the various strategies brought forward, but also of those discarded.

We then assigned the work to the students and invited them to concretely examine and analyze and solve the proposed questions in small groups.
Our activity developed through a number of phases that were independent of each other but preparatory and consequential. After giving a brief introduction on gamification and the various technical and practical aspects, we analyzed the data collected as part of these experiences, going from time to time to emphasize the many aspects that the use of certain teaching methodologies allowed us to investigate.
After the work done by the students in the groups, we asked them to review their own work, thanks to the logbooks drafted by the bloggers, to analyze concretely the proposed questions and the students' feedback on the experiences they had had and how they had had an obvious impact on their relational and self-assessment dynamics.
In the final part, we experimented concretely with some digital platforms, convinced that only by putting ourselves out there and experimenting on ourselves can we truly appreciate and evaluate a product.
As for the online gamified part, the students used different platforms, pointing out their pros and cons, such as Kahoot, Panquiz, WordWall, Deck-Toys, Genially, Scratch, LearningApps, but also some non-digital, unplugged activities, but still based and set on the rules of the game.
The purpose of this research activity was to foster a new educational design, also based on innovative methodologies that offer multiple possibilities and can be adapted and repurposed at will by teachers to enable students to learn, consolidate, learn, evaluate and self-assess. All this, of course, always keeping in mind that the student must be the builder of his or her own learning.
Moreover, these methodologies incentivize a new model of schooling based on cooperation and sharing, which is also more in line with what is demanded by the world of work, to foster the creation of digital learning content and beyond.
Clearly through the use of the INVALSI tests in unusual contexts for them, this project has also enabled students to then take the tests more calmly and consciously.

***Keywords***: gamification, formative assessment, digital platforms, educational innovation

## INVALSI and PNRR VS learning fragility and implicit dispersion: the necessity of a virtuous school system

**Ileana Ogliari – Andrea Guarnacci – Mariarosaria Orefice**

Introduction. The interpretation of data provided by INVALSI to schools throughout the national territory increasingly constitutes a valuable key for reflecting on the processes that guide students' learning paths. The results in standardized tests—with their descriptive levels that provide detailed information on what

each student knows or should know—represent a mapping of strengths and weaknesses, including the curricular designs of teachers. Through the analysis of these test items, teachers can individually or collectively reason about interventions aimed at triggering improvements in existing organizational and teaching practices. The additional research directions presented by the National Institute for the Evaluation of the Education and Training System, which include definitions of learning fragility and implicit dispersion, pose a specific call to institutions that cannot be ignored. It is necessary to design intervention actions that combine the expertise of professionals and the funds from the National Recovery and Resilience Plan—a tool whose effectiveness will be primarily measured by the strategies that schools adopt to leverage it.
Subject and Research Hypothesis. The subject of the research focuses on the awareness of the specific context of the Manfredini Comprehensive Institute in Pontinia and the inevitable repercussions associated with it. The institute consists of six branches located in a rather extensive area that radiates from the city center towards the rural outskirts, with a heterogeneous student population in terms of socio-economic and cultural backgrounds, prominently featuring a significant number of immigrants, particularly of Indian origin. In a scenario like the one just described, with an extended-day Lower Secondary School, it is evident that finding an effective synthesis between the resources provided by INVALSI and those, both structural and non-structural, guaranteed by the National Recovery and Resilience Plan (PNRR) represents a valuable opportunity to design and stimulate activities deemed most functional in progressively raising the students' learning levels. This is the crux: describing and implementing the intervention strategies designed to achieve this goal.
Data. The data provided by INVALSI to schools, including the fragility index, constitutes the starting parameter. It allows for a realistic, synchronous, and diachronic overview of the students' profile. Within the institute, there is a designated contact person who reports on the outcomes of national assessments. Based on these results, they propose a work plan that involves a team of teachers dedicated to the detailed analysis of the evidence and the definition of common tests to be administered according to a shared schedule during the school year. This procedure serves a dual purpose: introducing students to the structure of INVALSI items (with a constant focus on the inference mechanism, which is peculiar) and then training them to tackle it. Over time, this organization has led to the creation of a database, significantly contributed to by platforms like "INVALSIopen" and "gestinv." It has also established an organizational structure based on leveled groups (recovery, consolidation, and enhancement) that is functional to achieving the objectives.
Method. School administrators, staff, system figures, and teachers: the availability of PNRR funds has made the need for broad-spectrum active participation even more pressing. The constant factor is designing based on data: to contain the risk of school dropout, to involve families in the proposed activities, and, above all, to motivate students to embrace the outlined initiatives. Whether it is mentoring or coaching, the objective remains the same: to enhance the abilities and talents of each student, which continues to be the task of a school. To achieve this, the intention is to provide continuity to the various actions already undertaken in recent years, especially those related to digital initiatives. Teachers are engaged in a serious and documented training process that has resulted in the experimentation of relevant languages and tools. Through these efforts, it has been possible to tangibly experience the enthusiasm of participation and the solidity of acquired skills, such as coding and problem-solving, among others, demonstrated by the students.
Results. Didactic and digital innovation, a playful yet functional approach to achieving interdisciplinary skills: the potential of the ‘school 2.0’ suggests significant results. To achieve them—beyond the creation of purpose-built learning environments—we start from a series of previously experienced and replicable experiences that guarantee success. Various responses have been received from students at the end of the pathways initiated at the institute: solid evidence of the achieved milestones can be found in the portfolio of produced works and in the journals documenting their experiences. The cooperative mode of shared work, guided web research, the design and creation of manual or digital artifacts, games understood as effective vehicles for rules and steps to follow, and even healthy and proactive competition that arises when adult professionals lead the processes (without diminishing the students' role as absolute protagonists) provide a reasonable certainty that the chosen path can be considered correct. Furthermore, through the received gratifications, it contributes to building a self-awareness of one's own potential.

***Keywords***: digital resources, teaching strategies, technologies and tools in education, innovation

# Learning Goals Related To The Use Of New Technologies Is Measured Also Through INVALSI Data

**Alice Severi – Alessandra Marrata**

Introduction. Nowadays the effective learning can be achieved through a positive relationship among various cognitive, motivational, and emotional variables of a student. In fact, there are no longer ideal strategies suitable for all educational contexts and learners, but flexible methods must be adopted that can produce the best impact on learning outcomes. The health emergency during the COVID-19 pandemic has made it necessary to use technology in the field of education, first for remote learning and then for integrated digital learning. Teachers have had to quickly modify their teaching strategies as the health emergency imposed physical distance from the classroom, and it was necessary to cater to students' diverse learning modalities based on their cognitive styles. This has accentuated the ongoing changes in education, initiated by the digitalization of schools and subsequently the introduction of Artificial Intelligence in education. The ongoing societal transformations and demands of the job market require an innovative educational approach to prepare students for rapid economic and social changes. After returning to in-person classroom teaching at the end of the pandemic, the different ways of learning have become even more evident, and it has been possible to observe the impact on learning outcomes measured by the levels of competence achieved through the use of innovative teaching strategies. By leveraging the skills acquired by teachers and students in recent years and through the sharing of best practices, we can have an overall understanding of what schools need today, namely flexible, personalized, and inclusive strategies necessary to address significant changes in students' learning styles. The levels of learning achieved, deduced from the results of INVALSI assessments, help us provide a comprehensive evaluation of the changes that schools are undergoing, to prepare students for the workforce and to cultivate aware citizens.
Object and hypothesis of the research. The presented educational research is based on studying how the use of new technologies in the institute's teaching methods has improved students' skills and, consequently, their performance in INVALSI tests. The hypothesis of this research aims to demonstrate that teaching strategies that employ new technologies are crucial for reorganizing post-pandemic education and bridging the gap resulting from physical distancing from the classroom. With technology and Artificial Intelligence, learning becomes even more student-centered, adapting to their pace and learning styles, allowing each student to achieve both disciplinary and cross-curricular competency objectives.
Data used. In this study, we used data from INVALSI for the upper secondary school ISIS Follonica, specifically focusing on the past 5 academic years. Simultaneously, data regarding the number and type of technological tools available at the school within the same investigated time frame are also analyzed. Furthermore, the school curriculum is examined, as it has been modified in recent years by individual departments of the institute. The analysis also encompasses the didactic design for Learning Units, which have enabled students to undertake authentic tasks, including the use of technology, aimed at assessing both disciplinary and cross-curricular competencies acquired. The analyses are conducted at the overall institutional level, as well as differentiated by specific tracks, as ISIS Follonica comprises a scientific high school, scientific high school with applied sciences option, human sciences high school, linguistic high school, technical-economic institute with tourism, administration, finance, and marketing orientation, business information systems, and professional institute for health and social care services.
Method. Starting from the analysis of the INVALSI data, in this work we made a comparison of students learning outcomes before and after the COVID-19, connecting the digital projects made at school. The methods of comparison among these data allow to make punctual analysis on the relationship between learning outcomes and tools used, and also between methods used and results obtained.
Results. From the INVALSI data analysis, our school showed in general that the results in Italian were at a higher level than the mean of Tuscany and center of Italy (with differences among the studies pathways) and mathematics were at a lower level than the corresponding means, both in the second and the fifth year of secondary school. Related to general results of our school (that we deepen in the differences found in the studies pathways) we can see that the levels in Italian were at a higher level than the mean of Tuscany and center of Italy after the COVID-19 world emergency. The ones in mathematics were at a lower level than the corresponding means, but it has shown an improvement from 2021 to 2022. The improvement of technological equipment in different school buildings and the integrated digital education used during the COVID period, caused the raise of the use of the digital tools in the ordinary didactics. Therefore, we can see teachers keen on the specific training more than in the past, and students use new technologies in more

quantity and awareness than before. As a result, students have had an improvement of the skill levels, specially in STEM subjects (we can see it in mathematics level from INVALSI data). New projects about digitalization, that will be done by the Piano Scuola 4.0, action 1 (Next Generation Classrom) and action 2 (Next Generation Labs), have the purpose to improve learning also in Italian subject by increasing the use of digital tools in non-STEM subjects. Another aim is to raise the mean mathematics level reached by students, as expected from RAV 2022-25 in our school ("reduce the percentage of students that reach a level under 3 in the INVALSI test).

***Keywords***: digital, innovation, skills

## Digital skills assessed with ICILS 2018

**Paola Giangiacomo – Valeria F. Tortora – Monica Papini**

Introduction. Digital competences are skills to use information and communication technologies (ICT) to search, manage, evaluate and share information in different contexts and purposes. These skills are essential to participate in contemporary society and education, where texts and information are often presented in digital format and require specific reading, comprehension and interpretation skills. The international survey IEA ICILS (International Computer and Information Literacy Study ) evaluates the level of preparation of eighth grade students in this area and is carried out every five years. The second edition took place in 2018 and involved around 46,000 students from 14 countries, including Italy. The results of ICILS 2018 can provide participating countries with useful indications for improving the teaching and learning of digital skills in schools. ICILS aims to answer an important question: how prepared are students for study, work and life in the digital age? To do this, ICILS measures digital competence: that is, the ability to use the computer to research, create and communicate in different contexts. These competencies combine computer literacy, critical thinking, technical skills and communication skills applied to a range of contexts and purposes. ICILS also studies teachers' experiences, attitudes and familiarity with the use of computers for teaching, thus comparing different points of view on issues of crucial importance for those who work in the school world. School principals also answer the questions of the questionnaire addressed to them to collect crucial information and data to evaluate the school context of the students. ICILS is based on a computerized test administered to students and on questionnaires administered to students, parents, school administrators and teachers.
Object and research hypothesis. The aim of the study is to analyze the level of digital competence of Italian students in the third year of lower secondary school, compared with that of students from other countries participating in the IEA ICILS survey. This aspect was assessed through a computerized test administered to students, which required them to carry out tasks related to information literacy, communication and collaboration, digital content creation and problem solving . solving in the academic, personal and social fields. The article also aims to analyze the factors influencing the development of students' digital competence, such as socio-economic background, gender, language proficiency, access and use of ICT at home and at school, self-perception of own skills, attitudes towards ICT and support received from teachers. The research hypothesis of this study is the significant evidence of differences between the countries participating in the survey, between the Italian macro-areas , between the socio-economic levels of the families and between the genders of the students. In particular, it is assumed that the countries with greater economic and technological development have obtained higher scores than the countries with less development; that the regions of Northern Italy obtained higher scores than the regions of Southern Italy; that students from affluent families scored higher than students from disadvantaged families; and that girls scored higher than boys. These hypotheses were formulated on the basis of the results of the ICILS survey, carried out in 2018, and of the evidence from other international research on students' digital skills.
Our hypotheses are based on some theoretical and empirical considerations:

- Digital competence is influenced by the economic, social and cultural context in which students live and learn. It is therefore assumed that countries with greater economic and technological development offer more opportunities and resources for the development of students' digital competence, compared to countries with less development.

- Digital competence is also influenced by the regional context in which students attend school. It is therefore assumed that the regions of Northern Italy, characterized by greater economic and technological development, obtained higher scores than the regions of Southern Italy, characterized by less development. This hypothesis is supported by the results of ICILS 2018, which showed a strong disparity between the Italian macro-areas .
- Digital competence is influenced by the family context in which students grow up and train. It is therefore assumed that students from wealthy families, with a high level of parental education and employment, scored higher than students from disadvantaged families, with low level of parental education and employment. This hypothesis is supported by the results of ICILS 2018, which showed a positive correlation between students' score in the CIL and the socio-economic status index (SES), calculated on the basis of information provided by parents on years of education, profession and the possession of cultural and technological assets.
- Digital competence is influenced by student gender. It is therefore assumed that girls scored higher than boys. This hypothesis is supported by the results of ICILS 2018, which showed a gender difference in favor of girls in the CIL. This difference could be explained by the fact that girls tend to have better language and communication skills than boys.

Research method. In Italy, the sample involved 3,500 students, 3,500 parents and 1,400 teachers in 175 schools and their school principal. From the standpoint of data analysis, the battery of questions on ICT use at home and at school, self-perception of one's abilities and attitudes toward ICT proved particularly interesting. The use of the linear regression method made it possible to investigate how much the context data of the students, the possession of cultural and technological assets in the family, information on the level of education and employment of the parents affect the acquisition of digital skills by the students. Furthermore, the analysis of the data collected through the teachers' questionnaire, in particular the use of ICT in teaching and the difficulties encountered in school application, has allowed us to highlight how much digital skills are linked to teaching practices.

Results. The international average score is 516 points, with a variation between countries from a minimum of 429 points (Chile) to a maximum of 600 points (Republic of Korea). Italy is in ninth place with 487 points, below the international average and below the European average (513 points). 18% of Italian students did not reach the minimum level, 25% were placed at level 1, 36% at level 2, 19% at level 3 and only 2% at level 4. Compared to ICILS 2013, the Italian average score increased by 13 points, but the percentage of students at the minimum level remained unchanged. Furthermore, the results confirm, as for the other IEA surveys and the national INVALSI survey, a strong disparity between the macro-areas: the North East obtained the highest score (508 points), followed by the North West (479 points), the Center (471 points), from the South (435 points). Finally, the results show a positive correlation between the socio-economic level of families and the score obtained in the tests: students from wealthy families obtained an average of 51 points more than students from disadvantaged families. Factors influencing students' digital competence include gender (girls scored higher than boys), socio-economic status (students from wealthy families scored higher than students from disadvantaged families), fluency in the test language (students who speak the same language at home and at school scored higher than students who speak a different language), access to ICT (students who have access to a computer at home and at school scored higher than students with no access) and ICT use (students who frequently use ICT for educational and personal purposes scored higher than students who rarely use ICT). Students who received more support from teachers in using ICT achieved higher scores than students who, conversely, received less support. From the analysis of the psychometric aspects it can be seen how much self-perception affects the results, the students who have greater confidence in their IT and media skills obtained higher scores than the students who have less confidence; moreover, attitudes affect the acquisition of digital skills, students who have a higher motivation and interest in the use of ICT have achieved higher scores than students who have a lower motivation and interest.

***Keywords*****:** digital skills, attitudes towards ICT, use of ICT, digital teaching

# THEME 9. ITALIAN, MATHS AND ENGLISH: TEACHING PROCESSES THROUGH INVALSI DATA AND TOOLS

**ORGANIZER: INVALSI**

**COORDINATOR: PAOLA GIANGIACOMO**

**26TH NOVEMBER: 4.00 P.M. – 6.30 P.M. {ROOM 1 – TEACHING SESSION 6}**

## From concept to question: the revision of INVALSI Teacher Questionnaire

**Giuseppina Le Rose – Doriana delli Carri – Francesca Leggi**

Since the school year 2011/12, INVALSI has been administering a Teacher Questionnaire to Italian and Maths teachers of all sampled classes participating in National Surveys. Over the years, tests have been modified and updated according to the different research purposes, while keeping constant the objective of exploring teachers' attitudes towards the National Surveys and identifying strengths and weaknesses related to the school context that affect students' performance.

There have been three main reviews of these Questionnaires: the first one was after the introduction of computer-based tests; the second one after aimed at adding a series of specific questions on Integrated Digital Teaching in the pandemic period; the last one, which is the topic of this paper, both to adapt the section about didactics for grade 13 of the secondary school, to build new questions in the section about English teaching (grades 5 and 8) and also to add a new thematic section about school drop-out.

This paper aims to describe all necessary steps in order to introduce these new these important changes: from the involvement of Italian, Mathematics and English teachers in a three-day presence seminar in March 2023 to the main results that emerged in brainstorming and group work activities and, thus, the updating of the theoretical reference framework.

During the seminar days there were plenary sessions with introductory speeches and parallel sessions with brainstorming or group work sessions. From a methodological point of view, it was decided to use evaluative brainstorming rather than making indicators in a rationally-deductively way, because the difference between the two approaches is not in the formal representation of the indicators but in those connotative and emotional aspects that belong to the strings. What emerged, however, was validated in terms of the study and in-depth study of the literature when updating the theoretical reference framework.

Taking into account budget constraints and representativeness objectives, a sample of 36 units was identified; non-probabilistic sampling was used on the basis of reasoned choice and supplemented by elements of random selection, bearing in mind the following characteristics: geographical origin (North, Centre and South of Italy), Professional role (Head Teachers, INVALSI Referees and Teachers), school order and grade (Primary, Lower Secondary School, Upper Secondary School).

Regarding data analysis, the main advantage of brainstorming over other techniques is that there is not information, in real, to be analysed but a set of indicators available for later use. In this context, they were used to build up a kind of conceptual map that was used in the next phase of group work. On the material that emerged in the group work without the presence of moderators, a content analysis was carried out about: 1) data preparation; 2) the definition of the text units to be examined; 3) coding system definition; 4) the verification of the coding system on a small portion of the text; 5) the coding of all the text material; 6) the study of the validity of the coding system; 7) the organisation of what emerged from the text; 8) analysis procedure explanation.

The organisation of the database had been made by comparing the crucial emerging aspects during two different moments of group work: proposals for revision of items made in each group and annotation of the salient points highlighted during the plenary discussion. On the basis of these proposals, a number of macro-categories were identified, which were initially applied to a portion of the text and then extended to the entire material. What emerged was used by the research team to adapt the existing questions but especially to make new items.

***Keywords:*** questionnaire, Italian, Mathematics, revision tools, evaluative brainstorming

# Outcomes and INVALSI questions, "strategic allies" for our school improvement

**Giuseppina Maria Grazia Cardillo**

During the academic year 2019/2020, the comprehensive school with 1,138 students, has started an improvement process on the basis of the critical issues that emerged from the RAV and declined in the PdM, regarding two disciplines: Mat and Ita

To make this action effective, vertical dips within the disciplines were introduced right from the first year in order to give a sense of uniformity to the work that was starting up.

The undersigned, as a teacher of the Instrumental Function for the coordination of the Ptof, has assumed the role of coordinator and mediator of the actions proposed in both departments.

The starting point was, at the start of the school year, the return of the results obtained in the previous year as a starting point for the identification of the main critical issues in both disciplines and in both school orders. This phase which is repeated every year, at the start of the school year, is fundamental, as it is up to the undersigned to facilitate the process of exploration and reflection among colleagues, as an important and precious moment for elaborating conjectures in order to identify the possible causes that have determined these outcomes. During the first year, a collaboration was established with the teachers of the Institute, which gradually became active and constructive over time and at the same time a path of self-training and systematization of good practices was developed, also involving the teachers in the year test.

Introduction During the 2022/2023 school year, starting from the critical issues that emerged after the analysis of the data on academic achievement and INVALSI results, it promoted a self-training activity for teachers, with the involvement of new recruits and their respective tutors. The actions put in place had as a focus the revision of the working methodology and in particular proposing in one's own classes some activities that had a strong engaging and motivating character towards mathematics and the elaboration of educational activities aimed at improving this perception. The strategic orientation of our improvement plan therefore started from the reflection on organizational and teaching practices starting from the logic of assessment for learning. The enhancement of resources has set self-training among teachers as the main objective, as an important aspect in their professional exercise.

In particular, the most important resources available on the net have been used for primary: To Count; and for the secondary M@tabel and INVALSI open, in order to be able to "draw" on them in an individualized and autonomous way. Furthermore, both during the PTOF approval committee and in the identification of the process objectives connected to the RAV priorities, the teachers were invited to deepen the INVALSI framework and the GESTINV platform for a reflection on the learning objectives and goals of skills according to a perspective of vertical development, such as to monitor progress along the work path.

Object and research hypothesis: This year the internal NIV evaluation unit, following the new SNV 2022/2025 cycle, played an important role:

1. revisited and redefined the strategic priorities - starting from the return of the INVALSI results, for the formulation of the improvement actions in the new plan, intended for publication in Scuola in Chiaro in a dimension of transparency, sharing with the community of belonging and promotion of improvement.
2. has established tasks, roles, timing for improvement actions 1st year - 22-23, to really trigger continuous improvement processes.
3. use of INVALSI data and items, to design and implement didactic activities focused on developing logical-mathematical skills for analyzing their thought processes and for intervening on the most recurring criticalities/misconceptions.

This growing awareness of the usefulness of INVALSI tests, not only for the purpose of establishing and comparing student outcomes, but also as a professional reflection on the daily work carried out, has undoubtedly contributed to improving our teaching practice. Thanks to the data provided by INVALSI and the support in interpreting these data, more and more colleagues choose to reflect on their work, to consolidate the most effective teaching practices and investigate the most critical aspects, in order to stimulate teaching innovation and improvement evaluation.

During the school year that has just ended, I have noticed more and more interest among colleagues in the comparison and in the possible repercussions, in a "positive" key, of the data released by INVALSI as a tangible tool for making comparisons and reflections, even in a diachronic key, on the you hesitate.

Method. The working method chosen was to plan the improvement interventions within the departments. vertically of both basic disciplines: Italian and Mathematics, subject to school improvement. This work will describe only the actions concerning the mathematical field.

Starting from the identification of the main criticalities encountered in their own context, also from the INVALSI results and making use of activities specifically selected on the curricular objectives of the discipline, a planning and experimentation path of teaching practices was proposed and agreed within the departments aimed at a reflection on the curriculum and the learning dynamics of the pupils.
Specifically, this work describes the actions and tools used in order to propose mathematics in an engaging and "concrete" way, experimenting with activities that facilitate the understanding of the relationship between theoretical abstraction and events of everyday life, thanks to the collaboration between teachers of school and "expert figures".
Phase 1. Established data, knowledge and skills
Identification of standardized questions with the highest percentage of error, appropriately selected so as to range over three of the four content areas (Numbers, Relations and functions, Space and figures), for administration as entry tests in the classes of both orders: primary and secondary. Thanks to the Gestinv.it platform, I traced each question to the Achievements of competence and to the Learning Objectives. Therefore, the focus of the first intervention, to help teachers, was based on these aspects/tools, useful for starting departmental work:

- Construction of a grid/scheme for the collection of information/observations deriving from the return of the test to the student;
- Focus attention on incorrect answers;
- Investigate the cognitive processes that the student has activated to answer the question
- Investigate the difficulties encountered by the student.
- What distance is there between the teaching practiced and the national skills assessment tests?
- Detection of learning through verification tests (e.g. by administering established items, etc.)

Phase 2. Reporting outcomes at the first vertical meeting: in order to highlight the major criticalities and carry out context analyses. Reflection and discussion within the "vertical" department began with some didactic questions to guide us in correctly identifying the improvement proposals When and why do kids make mistakes? investigate/how to act on the most recurring misconceptions and obstacles? to which the first considerations arrived: the lack of the production of conjectures by the students, therefore in daily practice the students must be stimulated in their argumentative competence, and the mathematical discussion within the work in small groups, as well as the use of examples and counterexamples by the teacher must always be present. Therefore, from these assumptions, it was essential to design a mini-didactic path, consistent with the reference programmatic documents, with the aim of eliminating the main criticalities and their recurring difficulties. In the INVALSI mathematics tests, teachers can find examples of problems to reflect on and to confront, thinking about the goals for the development of skills and the objectives set at ministerial level
Phase 3: identification and planning of "vertical" educational activities, with reference to the Goals for the development of skills and the Learning Objectives for the activities, to which all subject teachers have adhered, starting from the analysis of the needs of the students and effective teaching practices: preconditions for students' educational success.
In summary the actions wereManipolazione di artefatti-strumenti Prime ipotesi di soluzioni a criticità che si possono incontrare nella pratica didattica, sollecitando lo scambio fra pari delle esperienze;
Phase 4. Evaluation of the learning outcomes achieved and reflections on the experience in progress among the teachers of both levels in relation to the achievement of pre-set objectives.
Data used. The coplanning phase took place in two close temporally face-to-face meetings, in which I presented the scheme - see worksheet with the founding nuclei and related skills for the vertical development of the discipline. This phase was of fundamental importance for guiding the action and the didactic choices and the focus of the investigation. In fact, starting from the INVALSI refunds of May 2022 for the II and V Primary and the entrance test for the secondary, the percentage of correct answers is lower than in the other reference categories (macro-region) especially for the secondary and for some classes of the primary.
On average, the percentage of correct answers in the various categories

- Numbers 39.20 % correct answers
- Data 43.94% correct answers
- Figures 40% correct answers
- Functions 29.30 % correct answers

Furthermore, before planning the activities, an examination of the main misconceptions for the various content areas was carried out, drawing inspiration from the collection of evidence available on the GESTINV platform, focusing on the following dimensions:

- Argue
- Problem solving (knowing how to identify the appropriate strategies for solving problems.

Starting from these considerations, during the second meeting all the disciplinary teachers shared their proposals for organizing a teaching activity in the various fields of the discipline.

In the following two months, the colleagues planned, started and carried out the classroom activity to be completed by April, in order to create space for discussion and collective discussion regarding the evaluated results obtained. An important methodological aspect that all the teachers had to respect was the use of some established items, resorting once again to the Gestinv platform for the opportunity to select the items in a specific way, relating to the area chosen for the didactic intervention in classroom, in order to verify the possible persistence of errors and analyze them in class as a valuable training moment.

Initial favorable feedback from this classroom action.

- Changes in educational objectives: few, well-defined and achievable during the school year.;
- Changes in teaching deliveries and methodologies: the laboratory and playful approach is extremely attractive for pupils, especially for primary school
- Changes in the planning of educational activities: proposing to the pupils problematic issues, understood as authentic and significant issues, related to daily life, and not just exercises of a repetitive nature and stimulated by the teacher's guidance and by discussion with peers, the pupils identify possible solution strategies, leading to significant learning.

Results. On the teacher front, important changes have been achieved in teachers' attitudes with reference to their professionalism, greater awareness of the aims of ministerial documents, and of their teaching practices. The activities promoted and launched as part of the school improvement plan showed an improvement in particular of the students' argumentative skills, easily observable from the solutions to the questions in the final assessment phase. A further sign of the greater habit and greater care in arguing one's reasoning. Finally, this working approach has contributed not only to the strengthening of the disciplinary contents, but has also shown that stimulating mathematical discussion in the classroom favors a process of development of argumentative skills; in this way the vision of mathematics as a set of procedures is abandoned and the importance of pursuing activities related to problem solving and argumentation in the classroom is affirmed.

In summary, it can be stated that

- there was an average improvement in the correct answer, in the classes whose teachers reproposed the items;
- investigate the nature of the errors, as a permanent way of working in the classroom
- knowing how to identify the educational objectives that can be evaluated using the questions proposed by INVALSI.

The different study methods can explain the better academic performance obtained by one group of pupils, compared to another

Certainly, another aspect that can make the difference is a partial or hasty reading of the text which leads the pupils to concentrate their attention only on a datum to the detriment of other information useful for solving the question, while only a small group carry out a more careful reading aimed at the search for precise and targeted information.

Strengths (Positivity)

- The activities carried out in laboratory form, also resorting to the use of INVALSI items, as a useful teaching tool, have (re) opened up new teaching perspectives.
- Collaborative study spaces with the aim of supporting reflective processes (for example, through peer to peer, virtual learning environments) have determined a good didactic-disciplinary setting;
- sharing and exchanging experiences among peers;

Challenges and Purposes. Perfecting the management and documentation tools related to improvement actions Obviously it is crucial in a mathematics teaching that gives importance to thought processes and not only to products. To ensure that no student is left out and that everyone can achieve adequate levels of learning - education that is equitable, inclusive and of good quality. Create a school service capable of keeping merit and inclusion together, valuing excellence and paying attention not to leave anyone behind, cultivating talents and helping those in difficulty

***Keywords:*** Teaching practices, scholastic improvement, strategic planning, learning weaknesses, INVALSI tests, competences

# INVALSI Data To Build A Effective Didactic Approach And To Realize A Professional Growth Of The Teachers

**Alessandra De Angelis – Maria Chirico**

The tests INVALSI and the analysis of the data obtained make it possible to monitor students' learning at different stages of their school career in order to identify their level of competence on a national scale. This value of the tests is now recognised by all teachers and most students also because, often, the data obtained by the individual school is highlighted with the regional and national data in the representation of the results.

The data obtained from the tests INVALSI, however, has another value, which we can define as more active because it does not merely take a snapshot of an existing situation, but identifies and evaluates the most effecive strategies for improving the situation represented.

It is therefore a question of using data as a tool to assess and at the same time to improve students' transversal competences, which are that set of skills that are fundamental for getting around in today's society, in which innovation is continuous and 'obliges' constant and lifelong training ("I dati INVALSI come strumento per migliorare e valutare le competenze trasversali", P. Falzetti). Those precious data, collected with effort, with intelligence, with economic commitment shoud be used for their own purpose of improving teaching and training, and for research purpose, for which they offer indispensable tools ("("Che cosa fare - e che cosa non fare - dei dati INVALSI" A. Sobrero).

In the literature, many researches have highlighted how the data obtained in the tests can be used in a didactic approach; Castoldi in his research work ("Dalle prove INVALSI alle pratiche didattiche", in Falzetti 2017: 19-32) set himself the objective of verifying to what extend the INVALSI tests results returned to schools could be useful for initiating improvement actions with more effective didactic approaches; particularly in Italian and Mathematics, referring to the experience conducted in a school in Biella. In the same work, Castoldi emphasises how the use of INVALSI data on a strictly didactic level, and thus at classroom level, represents not only an opportunity for improving students' skills but also an opportunity for professional growth for the teachers themselves. Indeed, it should not be forgotten that the main actors in this process must be the teachers themselves and its is they who, firts and foremost, must be convinved of the didactic value of INVALSI tests.

"Evaluation can truly be an agent of change, " argues Prof. Tronchero of the University of Turin, "provided that: i) the data is given the correct meaning; ii) the school is able to understand the potential suggestions that evolution can make and is open to positive change. Evaluation applied to a 'learning school' can really help it to realise its full potential." (Il Servizion Nazionale di Valutazione e le prove INVALSI. State of the art and proposal for assesment as an agent of change", R. Tronchero)

In order to counter 'oppositional or uncooperative mental attitudes' on the part of the teaching staff, there should be greater involvment in the school evaluation process through targeted refresher courses, offered exclusively by INVALSI and not by training agencies, covering the whole country.

Teachers could thus become an active part of school system improvement pathway through the acceptance of the importance of evaluation aimed at identifying inherent problems of funcioning and consequent ways of solving them.

Research Hypothesis. Based on the above, the aim of the present work was to investigate the perception of the value attributed to the INVALSI tests by both students and teachers through the administration of two questionnaires; in particular, the investigation concerned two Higher Education Institutions, an Artistic High School located in Udine (Sello) and a Hotelier Institute in Brindisi (Pertini). This also gave us the opportunity to compare the data taking into account the geographical difference and type of school.

Data. In order to refute the research hypothesis, we relied on data collected from questionnaires administered to teachers and students; in particular, the questions were addressed to all teachers in the school and to all second-year students after they had taken the INVALSI tests; in addition, it was planned to also address the questionnaire in early September (2023-2024) to first-year students, with reference to the

INVALSI test they took in their final year of secondary school, and to students who had just finished their studies at the secondary school.
Work Method. The work method used to develope the was that of interviews, administrated through Google forms.
Results. It turned out that about a ¾ share of students did not consider the INVALSI tests useful as a time of self-assessment; more than half of the students felt the tests were difficult, both because they did not understand the requests of the question and because the subject of the question had not been expained in class. The emotional states with which the trials were faced were mostly of indifference and boredom. More than ⅓ of the students admitted that they had used, or at least attempted, to use suggestions from fellow students or other media. More than half of the students stated that their families do not know the INVALSI tests and their contents. At least half of the respondents did not practice before the INVALSI tests, either drawing on the questions of previous years published on the INVALSI website, or preparing themselves on special texts. Anyway, the exercises didn't happen in class. As regards the analysis of the results of the questionnaires administered to the teachers, a need emerged on the part of the majority about the usefulness of linking the results of the INVALSI tests to the certification of competences, both at the end of the second year and at the end of the course. This need appeared following the observation by the teachers of a lack of commitment by the students in the performance of the tests, confirmed in several cases also by the extremely reduced time that the students used to finish the test. The use of the multiple choice test typology, still for a good part of the teachers, is considered not very effective, as well as the usefulness of measuring the learning of some fundamental skills, indispensable for school learning of other disciplines, as well as in life, for citizenship or at work. With regard to possible changes to the structure of the tests, INVALSI circa 3of the respondents proposed the inclusion of open-ended questions to facilitate the assessment of the skills acquired by the students. Analyzing the data of the teachers, it emerged the need for training and/ or updating on the issues of evaluation and self-assessment of the Institute, reading and interpretation of INVALSI data, innovative teaching methodologies, class management and communication, learn to learn.
The above data on the results of the proposed questionnaires are fairly homogeneous from North to South, with modest and detailed variations being reported in the final publication.

***Keywords:*** INVALSI, indicator, school Improvement, high School, INVALSI form, geographical comparison

## Teaching practices and improvement of learning

**Lisia Piovano – Ester Valloreo – Valentina Mandruzzato – Manuela Carchesio**

The proposal is part of the self-assessment process, a.s. 2022/2023, of the Omni-Comprehensive Institute of Città Sant'Angelo, specifically in the analysis of INVALSI data for students of the fifth classes (three classes of 65 pupils) of the Primary School in order to understand the results and use them . An aware and shared use of the teachers of the Primary School and of the Lower Secondary School of the Institute who work within the same building as the Primary. The study, starting from the INVALSI data relating to the 2020/2021 formative testing, pays attention to the results of the students, when they attended the third primary classes; the reading of the items by the teachers; to the answers of the students themselves at the end of the data entry of the paper tests after the administration in May 2023. The study aims to illustrate how the teachers read the data collected to better understand the critical elements with respect to the actions implemented as teaching practices. To this end, the school career of these students was monitored and analysed, from October 2022 to May 2023, through the observation of Primary School teachers (24 teachers) who used part of the programming hours to observe each other as suggested by the research by the Agnelli Foundation and INVALSI FABBIANI2023@. https://www.omnicomprensivocsangelo.edu.it/nuove-indicazioni-i-ciclo-valorizzare-le-due-ore-di-programmazione-progetto-a-s-2022-2023/
These observation interventions were also oriented with a view to continuity in the lower secondary school classes in order to understand which and how many requests are made to lower secondary pupils, how much they are in line with previous teaching practices, with the current profile leaving Primary School and

which cognitive, logical and systemic processes are needed to make INVALSI data tools for guiding teachers and guaranteeing the educational success of students.
Research object and hypothesis. The object of the research is represented by the process of identifying and sharing the educational and evaluation choices in order to guarantee coherence and continuity in the most delicate moment, that of transition to the Lower Secondary School. The INVALSI data of the School, from 2018 to 2023, are interpreted by a mixed group, Primary and Secondary, for a comparison with the priorities and goals of the RAV and the Improvement Plan. Since the current year, the INVALSI dossiers of the fifth PRIMARY have been read through the organization of an Excel folder in order to understand the logic that underlies both the dossier and the distribution of questions. The analysis of the items and the related cognitive processes guides the teachers to identify the necessary interventions for the 2023/2024 academic year, to support the growth in terms of skills and competences of students in the lower secondary school. The research hypothesis consists in verifying whether targeted interventions implemented by the primary school, in synergy with the lower secondary teachers, can contribute to building a shared vision on assessment and on how to outline a student profile that is guided to to evolve and to be an ally, together with the family, of the teachers. In particular, the mutual observation of the teachers, the strategic orientation between primary and lower secondary schools, constitute a research-training path that leads the first cycle to an evaluation that gives continuity to the formative evaluation of the primary school.
Method. The methodological approach used in this work was that of sharing a tool that allows us to understand the data used and to build a device to record the pupils' answers and to use it, subsequently, as a model for the construction of parallel tests and ride. We intend to share and discuss how we have built grids on Excel files in which to tabulate pupils' answers, how to re-modulate didactic planning so that it is more effective and more responsive to the training needs of all students, to the vision of the School, to the expectations of primary and lower secondary teachers. In any case, the improvement of the levels in the national tests, even the raising of a single point, entails a longer time, a depth of learning, a synergy with the families and with the teachers of the Institute to promote significant attitudes and behaviours. Furthermore, with our work we try to answer a question that we perceive as a priority: "Why, despite our efforts, the levels that emerge in the tests are not satisfactory compared to our expectations?"
And again: "Among the factors that come together to determine the skills measured by INVALSI, which factor within our Institute represents the weak link in the chain of actions, reflections and studies that INVALSI provides for the benefit of schools?"
We have tried to answer these questions by assuming an objective position, that of the reflective professional, to identify the factor/weak link in the process. We began to deconstruct the tests carried out by the pupils by following new paths of reading and interpretation. This complex approach has led us to recognize teacher training as a fundamental lever. In the current year, 10 primary school teachers and 7 lower secondary school teachers followed the PON Value courses.
Results. We used the results of the INVALSI tests in Italian, English and Mathematics carried out by fifth grade students. From the analysis of the data we have understood, first of all, how important it is to adopt aware attitudes with respect to the training function that the restitution data have towards the teachers. We have understood that, for an effective use of the data made available to schools, it is necessary to develop an awareness of their educational function, a function aimed at everyone, parents and pupils, but above all aimed at teachers. The analysis processes relating to what emerges from the tests represent the framework within which to act to raise the levels of competence of the students of our Institute.
In the study conducted, at the basis of this work, we realized that the re-modulation of teaching planning in itself is not enough, it is not sufficient if it is not accompanied by an attitude of training and Shared Leadership among teachers. We have begun to carry out a process of correcting the tests to better understand, data in hand, what happens between these 4 factors: didactics, learning, teaching, student performance. We are perfectly aware of the fact that this experience that we present, despite having its own logic, does not claim a statistical value, but a comparison to frame the critical elements.
Furthermore, what we consider significant is the effort we have voluntarily made to build functional tools to be able to read the data returned by INVALSI more effectively. This is our horizon of meaning towards which we are reaching out in the logic of improving all those variables on which we can act, understanding their nature and impact on our Institute.
In conclusion, we are convinced that working on our training and using the reading of INVALSI tests as tools can guide us in learning to be competent in interpreting and teaching, giving meaning and value to what we do in a perspective of continuity within the eight years that make up the First Cycle.

***Keywords:*** teaching practices, school improvement, learning, educational research

## The class and understanding of the result of the INVALSI test

**Marina Paola Mariano**

Introduction. At the beginning of each new school year, the data of the INVALSI tests are returned to the schools that take action to bring them to the attention of all stakeholders. In this work we will deal with the students, precisely those who have carried out the tests and who are the first interested in knowing the results, together with their teachers. The students who worked in this research are part of those who, at the end of the tests in spring 2022, had been administered a satisfaction questionnaire on the work of the SNV just concluded and of which we reported last year in the VII Seminar.
Object and hypothesis of research. It is part of the teacher's teaching practice, once the keys to read the results of their class have been received, to review the results, whether positive or not, especially on which items the strengths and weaknesses have been concentrated, and in the case of the latter, on which it is necessary to reflect and intervene. The teacher knows that the design of any learning unit provides a moment of review in order to be more effective in what is the *mission* of the school, that is quality teaching. To the question of the students who ask how their INVALSI tests went, the most appropriate answer that can be given is <<let's see it together>>. The restructive phase of learning, that of metacognition, is the most effective and is the one where the teacher guides the students and places the teaching of doing at the center of attention, making the students in fact the main observers of their work, forced to reflect and argue the choices made regarding their answers.
Data used. Review in class the tests carried out in Italian and Mathematics, distributing the files kept and guiding a work of analysis of the individual items, dwelling right where any wrong answers were the most recurrent trying, through the guided conversation, to understand and make people understand what was the path that led to the answer, on which topic it is necessary to return to deepen. It is also interesting to compare the results of the tests carried out in spring 2022, had by INVALSI in the following autumn, with the personal perception that the class declared at the end of the tests, through the answers provided with the satisfaction questionnaire administered about ten days after having done them. This too becomes a moment of growth through reflection and metacognition on what has been perceived and the objective reality that stands out from the correction of the questions.
Method. Sharing, dialoguing with one's classmates in a discussion where the awareness of one's own mistake is no longer something to be ashamed of, but rather, becomes a moment of growth that allows students to implement social learning, to get to reflective learning, a fundamental element that helps today's students to overcome the difficulty in managing the frustration that derives from a negative performance. Immersed in an overprotective family environment that manifests itself in an all-out defense of their work, students risk losing control of their actions and choices, seeing that there is always a justification to explain their mistakes. The teacher-led discussion in class stimulates peer education where students exchange their skills and compare hypotheses about the possible answers given, working together.
Results. Building an educational relationship in the school, based on recovering the centrality of the student who is the first actor of his own learning path and in the first person is aware of all the steps that his own growth entails. Become able to rework the tests carried out, together with the classmates, giving space to the observations made by the teacher that become a stimulus for reflection for the students who through this process learn to evaluate their work and also to help the classmates who struggle in this perspective, with a view to inclusion and sharing. The new evaluation of primary school pushes teachers to move from a summative evaluation (assessment of learning) that is done at the end of a path and that goes to measure what the class has learned, to a formative assessment, which is parallel to learning (assessment as learning), indeed it is part of it, as it accompanies the learning process itself allowing continuous correction to reach the final goal, full mastery and achievement of the expected competence. Self-evaluation becomes school practice and also passes through the INVALSI tests that do not remain an isolated episode, almost alien and detached from the everyday life of school life, but an integral part of the learning process.
Working on the data returned by INVALSI, as well as being fertile ground on which to work with the students, also becomes a moment of re-elaboration and comparison with the colleagues with whom the teacher has designed at the beginning of the school year the activities that guide towards the learning

objectives. The learning objectives that make up the curriculum of a given school, are declined in educational activities that, initially designed together, the interclass teachers, gradually compare during the year to establish alternative or corrective paths depending on the feedback of the students. The comparison of the results of the INVALSI tests by class, far from being a moment of pride or personal shame, instead becomes a working ground precisely on those issues where at the level of parallel classes a similarity of answers has manifested itself. The teaching practice makes use of a valuable tool that in addition to helping students to reflect on what was declared about the conduct of the tests and what it really was, helps the teacher to avoid stereotypy in their judgments, that is, to be affected by an a priori thought, to avoid overestimation or underestimation compared to previous tests; not to be influenced by previous judgments, but rather to grasp through the passages of the conversation in class, those reflections, those motivations that the students externalize in arguing their answers, and that offer the teacher an additional element in the practice of direct observation; something recommended by the New Evaluation that favors the use of different evaluation tools to be able to formulate the descriptive judgment.

***Keywords:*** teaching practices, learning, educational research

# THEME 12. FROM SELF-EVALUATION REPORT TO SOCIAL ACCOUNTABILITY: THE VALUE OF INVALSI DATA

**ORGANIZER: INVALSI**

**COORDINATOR: MICHELA FREDDANO**

**26TH NOVEMBER: 4.00 P.M. – 6.30 P.M. {ROOM 2 – TEACHING SESSION 7}**

## The Use INVALSI Tests Data For Training, Informing And Improving. Teachers' experiences, attitudes, opinions, and perceptions

**Paolo Barabanti**

Introduction. INVALSI has different institutional duties. Some of them are the ones about the management of the National Evaluation System (SNV) which require periodic and systematic tests on students' knowledge, abilities and skills and on the overall quality of the educational and training offer, with the aim of promoting the full implementation of the autonomy of educational institutions, avoiding the risk of self-referentiality, and, most importantly, promoting the educational success of the students.

INVALSI, by means of its tests, tries to fulfill this task by providing a rich and useful information gathering both for schools and for policy makers. However, the delay of the Italian education system in introducing the school and the learning evaluation, in addition to a scarce culture of accountability by public institutions compared to other European Countries, meant that the start of mandatory INVALSI tests, after a few years of experimentation, arrived chronologically before the cultural consensus around the need to make these measurements. Moreover, the fluctuating regulatory uncertainty has not been able to legitimize, from the very beginning, this task given to INVALSI. All this has contributed to raise and power a certain cultural resistance which has been manifested in different forms, including abstention from tests, their boycotting, the persistence of phenomena such as cheating and teaching to the test. The oppositions and conflicts from teachers, school principals and sector trade unions, boosted or even exacerbated by the press (Corsini and Losito, 2013), about these tests and deriving from a different/wrong understanding of their functions and purposes, have however helped INVALSI work, over the years, to make changes in order to improve, for example, some methodological, communicative and informative, training and content aspects (Ajello, 2014; Trinchero, 2014). Actually, INVALSI has done a lot in order to improve its partnership and its communication with schools, leading to a significant improvement in the sentiment about these tests, as well as to a more and more growing awareness of the importance and the usefulness of this kind of data to direct both didactic and organizational actions to activate improvement paths, to support the macro and micro decision-making processes towards more informed choices and more responsive to students' needs (Brookhart, 2011; De Luca, 2012; Calvani, 2013; Darling-Hammond e Adamson, 2014).

Research object. In the light of all these considerations, this research aims to investigate both the processes that the different school stakeholders experience with regard to INVALSI test and the deriving dynamics in terms of organization, relationships, reflections and didactic. More specifically, teachers' perceptions, opinions and representations are investigated, also as spokesperson for other school stakeholders, about the content and operational system adopted by INVALSI and by the schools themselves – i.e. the organizational, communication, training and information processes – in the previous period, during the administration of the tests and at their conclusion. Gathering this information, in a logic of bottom-up involvement, can also help INVALSI improve some dimensions and aspects about its national tests in order to offer a better quality service, to corroborate the partnership with the educational institutions and to try to overcome some resistances that today – even if weaker than in the past – still remain.

Data and method. For these purposes, data collected by over 150 primary, lower secondary and lower secondary school teachers who took part in the online training course "INVALSI data for an informative, training and the improvement" will be used (training course provided by INVALSI by means of SOFIA platform between May and July 2023). At the end of the package of hours of live lessons, participants were asked to carry out some final activities in an asynchronous mode, including answering some open questions, collected by Google Forms and then analyzed with NVivo, a qualitative analysis program. The research has adopted a qualitative approach, allowing to shed light on experiences, perceptions, strengths and weaknesses, doubts and improvement proposals of all actors involved, by collecting rich and significant productions on some thematic cores: INVALSI tests content; operational and organizational methods in the

various pre - ongoing - post administration phases; methods of communication within schools, between teachers, to families and students; teachers' training needs.
Main results. The research has highlighted a plurality, not always linear and rarely uniform, of opinions and perceptions of the usefulness of INVALSI tests but also their weaknesses. Strengths are about macro-system dimensions but also meso-institution and micro-class; the critical points concern several topics, including the risk of competition between students or between schools and of judgment for teachers, the difficult organization of test days, the persistence of a certain distance from the teaching, the complexity of reading and correctly interpret the data, the difficulty in motivating students to carry out these tests in a serious and responsible way. Among the proposals for improvement, the need for a closer communication with schools and a constant training often emerges, in order to help stakeholders understand the meaning of these tests and to support them in making the best use of the data returned. Furthermore, teachers involved in the research talk about multiple ways to organize the administration phase as well as to inform students and families about the days of test. Schools are aware that this is a process that has improved over the years, which has been modified in order to improve internal communication, motivation and work tools so that mistakes could be avoided for an effective carrying of the tests.

***Keywords:*** INVALSI tests data, data literacy, self-evaluation and improvement, school community, teacher training

## The Results Of The INVALSI Tests As A Tool For Strategic Planning

**Marianna Rasetta – Graziella Marrone – Marta Feliciani**

Introduction. From the a.s. 2013-2014 the comprehensive institute of Loreto Aprutino started a reflection, a study and analysis of the results of the INVALSI tests. Each year it focused to an aspect highlighted by the INVALSI contact person, by the Internal Evaluation Unit. In ten years of activity, the widespread mistrust regarding the National Tests has been overcome, most of the teachers have understood its importance and value. However, it is difficult to use fully and consciously the data returned by INVALSI to schools. The objective of the work concerns the opportunity to review one's own educational, teaching and evaluation practices backwards in an improving perspective through a possible self-training path using the INVALSIOpen platform. The school's focus is on achieving the goals set in the short and long term in the Self-Assessment Report through actions planned and included in the Improvement Plan. The contribution aims to illustrate the path and the improvement actions that the comprehensive institute of Loreto Aprutino has planned for the three-year period 2022/2025 on the basis of the INVALSI results analyzed at the beginning of the 2022/2023 school year.
Context. The new school context, following the sizing and inclusion of the Civitella Casanova complexes, the presence of a new Headmaster have required even more a sharing of processes and procedures concerning the INVALSI National Surveys starting from the return of the INVALSI results in the first Board of 6 September 2022, thanks to the publication of those on 29 August 2022. The return of results, the analysis and interpretation of the data, provided at the beginning of the school year, represent the last ring of reflection for the elaboration of the Social Report for the three-year period 2019/2022 but also the starting point for designing and planning the new three-year period 2022/2025.
In the light of the return of results, the reading and interpretation of the INVALSI data, the priorities and goals to be achieved in the short and long term, the process objectives and improvement actions were identified. The Self-Assessment Report, drawn up in October 2022 and published in Scuola in Chiaro in January 2023, highlights the difference between the scholastic results, in which the school self-evaluates itself with a score of 4 and the results in the standardized tests in which the school score assigned is 2.
Subject. The school aims at the educational success of the students but at the same time fails to achieve significantly higher levels in all the disciplines being evaluated. This discrepancy between school outcomes and results in standardized tests has led the Headmaster, the INVALSI referents, the Internal Evaluation Unit, the instrumental functions to examine in depth the INVALSI results of the second and fifth primary classes 2021/2022 to design the new three years.
Data used. The INVALSI data analyzed at Institute level show this starting situation: in Italian fifth primary school, the school effect is equal to the regional average and the southern macro-area with acceptable

results but with results to be improved in comparison with the national data; in fifth grade mathematics, the school effect is below the regional average, in the southern macro-area and in Italy, the contribution of the school is not adequate and the results need to be improved; in fifth primary English the results are positive, by crossing the data of the two tests it can be seen that all pupils have achieved A1 learning levels in reading, 8.5% have a pre-A1 level and 91.5% an A1 level in listening. The system figures, coordinated by the Headmaster, have identified and shared in the departments all the documentation produced, focusing in particular on two goals to be achieved at the end of the three-year period 2022/2025, linked to the results in the standardized tests and to the remote results: in Italian and mathematics to obtain a school effect on the results of third-year secondary school pupils at least equal to the regional average; in the 2024/2025 INVALSI tests, obtain remote results of students who attended the second and fifth primary years in 2021/2022, equal to the regional average in Italian, mathematics and English. The process objectives connected to the identified priorities and goals concern actions related to the curriculum, planning and evaluation and to the development and valorisation of human resources: analyzing the planning and evaluation tools to make any adjustments aimed at the success of the pupils in the INVALSI tests; implement training and/or self-training courses for teaching staff, also through the sharing of good practices, in order to improve teaching/learning processes.
Method. The self-training path entitled "Knowing to improve" was launched in February 2023 with a circular from the Headmaster to implement a first phase called "MAT-ITA SUPER", with the aim of knowing and understanding in depth by the teachers whose students took the INVALSI tests in 2021/22, information on the results in their classes, with particular reference to the learning processes that highlight any critical issues. This action involved in a first phase the INVALSI representatives with the task of supporting the councils of the classes interested in researching, reading and interpreting the graphs and tables available in the INVALSI Cineca area. This first action allowed the class councils to draw up a final written report, to enter into the details of the single classes, to analyze not only the general tables but above all the graphs and tables relating to the details of the single tests. The strengths and critical points that emerged in the learning processes investigated were highlighted and one or more hypotheses were formulated on the reasons that led to the not entirely satisfactory INVALSI outcomes. The analysis activity was extended both to lower secondary school teachers who have students in their classes and in the school year 2021/22 took the fifth grade test and primary school teachers whose pupils, in the 2022/2023 school year, would have taken the established tests. At the same time, the second self-education and training action was planned to be promoted in September 2023 aimed at the entire teaching staff. The trained or qualified internal staff will promote, starting from insights made available by INVALSIOpen for Italian, mathematics and English, with moments of critical reflection, planning and implementation of training modules and subsequent verification and sharing of results. With INVALSIOpen, as underlined by the President Roberto Ricci, "INVALSI wants to make the data of the national Surveys and the information that emerge from them available to the Schools. By reflecting on the results of the Tests, it is in fact possible to intervene both on the critical issues that have emerged and to work on strengthening the strengths present." INVALSI makes a range of training and information resources available to teachers. It also offers training and in-depth courses for teachers on some central themes of Mathematics, Italian and English. The use of the materials is optional and the teachers can independently establish the methods and timing of their use. INVALSI training materials are in fact designed as teaching support tools, useful for providing suggestions, ideas and operational materials with the aim of helping the school to improve the students' learning outcomes.
Results. From this path, the acquisition of greater skills by the teaching staff is expected, as well as the sharing of good practices, the development of planning and evaluation tools to carry out activities aimed at developing the pupils' learning processes, have a greater impact on the critical issues that have emerged in relation to the different learning processes that the INVALSI tests involve and solicit to improve the results in the standardized tests.

***Keywords:*** INVALSI Data, self-assessment report, strategic planning

# INVALSI Data: External And Internal Evaluation

**Maria Carbone**

The evaluation of learning is a demanding activity for schools: it requires a strong assumption of responsibility in the choice of coherent theoretical models, rigorous methodological arrangements, valid and reliable tools for the collection of knowledge and skills. What needs to be avoided is self-referentiality and this is helped by the national surveys carried out by INVALSI.
Knowledge is the main lever of the socio-cultural development and economic-financial competitiveness of a State so it is necessary to rethink the role of school evaluation, increasingly connoting it as a tool "at the service of learning"such as to implement individual skills and enhance the quality of education. It is important to use the conscious use of multiple devices, quantitative and qualitative, related to testing assessment and authentic evaluation depending on the processes and products to be evaluated. The evaluative action, has countless fields of action and must ensure throughout the national territory the achievement of learning targets, which allow each citizen to contribute in a conscious way to the individual growth and consequently the state. The DPR 80 of 2013 defines the Regulation of the national system of evaluation in education and training and establishes the SNV that evaluates the efficiency and effectiveness of the education and training system in accordance with Article 1 of Legislative Decree 19 November 2004, n. 286. It consists of the INVALSI, which assumes the functional coordination, the Indire and the inspection contingent. The evaluation process of the schools is developed in four phases that are realized in continuous cyclical succession: self-assessment of the educational institutions, external evaluation, improvement actions, social reporting of the educational institutions. The evaluation process, defined by the SNV, begins with self-assessment and the tool that accompanies and documents this process is the Self-assessment Report (EWRS) whose format is made available at national level, open, however, to the integrations of schools to grasp the specificity of each reality without reductions or oversimplifications. The report provides a representation of the school through an analysis of its functioning and forms the basis for identifying development priorities towards which to direct the improvement plan. A specific section of the RAV is dedicated to the results of standardized tests and the school, in which I teach, in this section has identified its first criticality and has clearly and numerically defined the goals to be achieved; in the PDM has detailed the actions to be carried out to achieve the goals. INVALSI data in its many facets is a very useful indicator to monitor results over time and avoid self-referentiality. These evaluations are not in contrast to the daily formative and summative evaluation carried out within the school, but despite the limitations, they alone can guarantee the comparability of the results achieved by pupils and schools over the years. The Institute contributes on the one hand to making transparent and accessible summary information on the most relevant aspects of the education system, especially useful for policy makers in making choices on the education and training system; on the other hand it provides information that can be used by individual schools to enrich the self-assessment processes aimed at improving their educational processes. Precisely because of the objectivity and standardisation of the tests, they are useful in the RAV and in the subsequent social reporting, at which time the school must give account of the results achieved or of its deviation ,implementing the PDM. To make the most of the data returned to my school, some important actions were carried out: stabilisation of the INVALSI contact person and meetings dedicated to sharing and critical reflection on the data received. The percentages in level 5 of the INVALSI certificate and in the advanced level certified by the school were compared and a substantial difference was recorded in favour of internal certifications, in addition to the criticality in the general scores. The evaluator and the manager agreed to activate specific meetings in the disciplinary groups to undertake actions to improve the overall results and especially the evaluation action to reduce this deviation. Evaluation models with agreed indicators and descriptors are used in the school, but given the results, the need to review them in the light of the INVALSI evaluations has been realised. Evaluation is a process that not only affects the cognitive and meta-cognitive sphere of the students but strongly involves the affectivity of the students and on this involvement depends, to a large extent, their attitude towards the school the objective tests of profit , are an indispensable support to teachers for the initial setting of their teaching activity, for the control and targeted programming during the school year and for the final verification of what they learned. Over time, the INVALSI contact person has followed specific training courses activated by the networks or chosen independently, has participated in regional meetings with the head of the school and has achieved a good knowledge of the roles of the SNV and the INVALSI tests. This acquired knowledge is sterile if it remains personal and of no use for improvement. Achieving cascade training is very important because improvement is possible if the School

has the ability to invest in collaborative activities and horizontal reflection and create communities of practice. The topics covered in these meetings were:

- The SNV :structure and functions
- INVALSI
- National indications and INVALSI tests
- Level descriptors in INVALSI results
- Interpretation of the data
- Definition of INVALSI results improvement actions

The dissemination among colleagues of the SNV's actions and the importance of standardised evaluation as a support for personal evaluation has increased the understanding of the data and their possible use. The data has been used in an interpretative way and this has led to a greater awareness of the paths outlined in the PDM that are not only a task of the NIV but affect the entire school community. Starting from the data and using the tools provided by INVALSI, In the sections and in the magazine INVALSI open, the processes underlying the questions were analysed in a shared way to agree on possible changes to the teaching methodologies. Particular attention was paid to the description of the levels of competence and the type of questions related to them . The interpretive reading of the data was a particularly challenging moment because it is from the comparison of the data of the School with the regional and national data that can result in improvement actions to ensure effective training for all and contribute to the reduction of territorial disparities and early abandonment of education and training. La riflessione su quanto esposto determina la scelta consapevole di azioni di miglioramento per garantire a tutti il raggiungimento di competenze spendibili.

***Keywords:*** data, evaluation, horizontal training

## School inclusion: special educational needs and students' sense of belonging through IEA TIMSS data and the RAV

**Francesco Annunziata – Laura Palmerio**

Introduction This work concerns school wellbeing and sense of belonging, with the aim of promoting students' wellbeing and inclusion. Several studies have shown that the sense of belonging in school plays a crucial role. Students who feel accepted, included and supported in their classroom are more motivated in learning, actively participate in school activities, develop positive social relationships and perform better academically (Dimitrellou & Hurry, 2019; Osterman, 2000). Schools play an important role in ensuring the integration of students with disabilities, both in terms of relationships and learning. The development of supportive relationships can be a valuable resource in the inclusion process, and it is therefore desirable that all educational activities of students with disabilities are carried out in the classroom together with their peers. In schools in the first cycle (grades 1 to 8), students with disabilities who have limitations in their autonomy spend most of their time in the classroom (27.6 hours per week) and carry out educational activities outside the classroom only for a limited number of hours (2.6 hours per week). However, if the student's limitations are more severe, the time spent outside the classroom increases considerably (6.4 hours per week). According to the European Agency for Special Needs and Inclusive Education, (2021), countries have implemented various actions to try to overcome these difficulties, among which providing an environment in which children feel accepted is key to fostering the development of a sense of belonging to the school environment and promote inclusion. In particular, for children with special educational needs, feeling respected and treated similarly to their peers without special needs could foster the development of a strong sense of belonging to the group.
Data from the IEA - International Association for the Evaluation of Educational Achievement international surveys. Data from a subsample of 196 fourth-grade students participating in TIMSS (Trend in International Mathematics and Science Study) 2019, collected through the student questionnaire, were used in this study. In TIMSS, all students in the selected classes regularly participate in the test. However, students with special education needs (SEN) are divided into two different groups according to their disabilities:

1. Students with certified functional, cognitive, behavioral or emotional disabilities (PCS). To promote inclusion, they can take the test if the teacher deems it appropriate. However, their results are not included in national and international reports.
2. Students with Specific Learning Disorders (SLD). They participate in the test using the same aids they normally use during school activities. Their data are included in national and international reports.

This study focuses on the data of students excluded from the international database due to their disabilities, but who participated in the survey for inclusion reasons. Students who participated in the survey using special aids (SLD) were also considered. Furthermore, in order to avoid distortions due to comparisons between students from different schools/classes, a random subsample of students who took the test regularly and attended the same class as the students with severe disabilities and SLD was selected.

Method. To achieve the clearest and most truthful objectives and conclusions possible, a mix method approach between qualitative and quantitative research was used.

Measures. In the quantitative analysis, four variables related to the sense of belonging present in the International Student Questionnaire were considered: "I like being in school"; "I feel safe when I am at school"; "I feel like I belong at this school"; "I am proud to go to this school". Students had to express their agreement or disagreement with these statements on a 5-point Likert scale, ranging from "Disagree a lot" to "Agree a lot ". The mean of these responses was zero, with a standard deviation of 1. An exploratory factor analysis was conducted on the entire sample: the data showed a single-factor structure explaining 53% of the variance and with an adequate Cronbach's alpha (.74). Regarding the qualitative study, the inclusion of school-internal intervention strategies was explored by examining the Self-Assessment Reports (RAV) of the schools involved in the study. The RAVs represent a fundamental part of the evaluation process of the National Evaluation System and are compiled by the school headmaster and the internal Evaluation Board, which is composed of a selected group of teachers. The reading and analysis of the RAVs provide further elements for comparison and understanding of the inclusion strategies adopted within the schools involved in the study. These qualitative data allow us to deepen our knowledge of inclusion practices and policies within the schools, and to assess possible links with the results that emerged from the quantitative analyses. In the context of this study, we focused on the section of the RAV on Inclusion and differentiation. This section provides information on the actions implemented to promote inclusion, the tools used, the strengths and weaknesses found in relation to inclusion and differentiation. The self-assessment rubrics in this section have also been considered.

Expected results. The combined analysis of quantitative data (from questionnaires and rating scales) and qualitative data (from the reading of the RAVs) makes it possible to obtain an in-depth view of the phenomenon of school inclusion, identifying effective practices and key factors that can influence the well-being and sense of belonging of students with disabilities.

Specifically: an analysis of variance was conducted, with student category (three levels: PCS; SLD; regular) as the intermediate factor and student sense of belonging as the dependent variable. The results of the ANOVA show a difference in the sense of belonging between the students of the different categories ($F[217; 2] = 3.961$; $p<0.02$). In particular, the post-hoc analysis performed with the Bonferroni method shows that PCS students have a statistically lower sense of belonging than SLD students (-.24 vs. .35). Reading the RAVs shows that the most common tools for inclusion include the use of shared criteria for assessment, measures to ensure the accessibility of resources and the adoption of a welcoming protocol for students with disabilities or special needs. Other less common tools are compensatory software and measures to make spaces and facilities accessible. However, only slightly more than half of schools use accessible versions of textbooks for students with sensory disabilities. Strengths in differentiating inclusion include teamwork between teachers, such as collaboration on adaptations and systematic observations, and synergy between curricular and support teachers in the use of inclusive methodologies. Classroom work' is another key point, involving group activities, personalisation of objectives and valorisation of students' different learning styles. Weaknesses include the need for qualified staff, the lack of specialised and tenured support teachers, and educational discontinuity caused by fixed-term contracts. Family involvement is critical, as is participation in decisions concerning educational activities. Finally, the Evaluation Rubric of the sampled schools emphasises the importance of differentiating educational pathways and evaluating achievement. However, according to this study, less than half of the surveyed schools regularly adopt differentiated intervention activities based on student performance. The results of this research could provide valuable insights and information for improving the well-being of disadvantaged students in schools. By understanding the factors that influence the sense of belonging and school well-being, it will be possible to

identify targeted strategies and interventions to promote an inclusive school environment and offer appropriate support to disadvantaged students in their educational journey.

***Keywords:*** sense of belonging, disability, educational inclusion, TIMSS, RAV

## Data and Tools: Sharing Best Practices

**Maria Carbone**

The National Evaluation System (SNV) is a strategic resource to direct school and training policies towards a cultural, social and economic growth of the country and to encourage the full implementation of school autonomy. Improvement is the common mission of the Schools, of INVALSI and of the SNV and each one makes its contribution. The SNV operates: the evaluation of educational institutions, the evaluation of management and the enhancement of the professional merit of teachers. ( D. Lgs. 286/2004); it provides a common platform on which to compile the strategic documents that interface with each other: Social reporting, RAV, PTOF with a special section dedicated to monitoring and PDM. INVALSI provides reliable and comparable data on students' skills in specific segments of the school system: grade 2,5,8,10,13; and has the data implemented with the School Effect, therefore represent data on which to be able to report. The experience I share concerns the use of the data returned by 'INVALSI within the educational institution where I teach. In implementation of the Regulation on the National Education and Training Evaluation System (DPR 28 March 2013, n. 80)The School draws up documents which, although separate, are interconnected and the point of intersection can be represented by the results of standardized tests. The documents are : the social reporting, the school questionnaire, the RAV, the PTOF, the PDM and among them in the RAV in the section Outcomes a subsection is devoted to the results in standardized tests . INVALSI data, do not have an evaluative nature of either the teachers or the School but, are a reference point. They provide levels of expertise in the areas investigated by using homogeneous indicators at national level ; are comparable and return the added value compared to the benchmark; avoid self-referentiality and photograph the reality of the school from different angles. INVALSI data have been used for years as a tool , in the school institution, where I am a teacher, to make known to the school and the territory where it is located and what it wants to reach , throughout the organization and have guided the methodological and organizational choices over the years; INVALSI's service to schools is very important because it provides comparable data at different levels and of different kinds, so the School, in its autonomy ,has selected step by step the results he wanted to achieve and addressed in a clear and targeted didactic-organizational actions in the PDM The general objective is to stimulate those self-assessment processes in order to improve the teaching and organizational processes ,which form the basis of the National Assessment System (SNV) of schools. The purpose of the work presented is to share the consolidated experience in my school on the use of the results of standardized tests and the dissemination of this protocol in two schools of the territory : a comprehensive Institute and a High School-scientific-classical. The actions implemented regularly in temporal order are: 1)Stabilization of the contact person INVALSI 2)collection and return of data of standardized tests in the teachers' college, by the contact person INVALSI . 3) illustration in the School Council of those data that involve the whole School Organization as the general data, the trend over the years, the heterogeneity between classes, the added value as they are decisive data for the accontability, 4) meetings in the departments that are fundamental ,for the formulation of the curriculum of the Institute and the choice of shared didactic-methodological actions operating as a learning organization. The choice not to limit the meetings to the teachers of the disciplines involved was determined by the conviction that some mental processes in the questions of mathematics are mobilized by all the technical-scientific disciplines and in the language by all the other teachers . 5)communication to parents of the importance of testing as a formative moment for children.The NIV, composed of instrumental functions and other teachers, coordinated by the head of the school Ing.Michele Antonio Iovine compiles the report document and the various sections, of the RAV and before publication it shall be made available to the teaching staff to share their choices and priorities. The sharing of the document is important because it is not a bureaucratic fulfillment, but represents the starting point for improvement. The first priority identified in the EWS was the improvement of standardised tests, not as a single datum but in three sections: reduction of variability between classes, reduction of concentration of pupils in the first and

second levels, decrease in the difference in the levels of assessment attributed to pupils by INVALSI and the teachers, attributing a percentage of improvement in the targets ,in order to be able at the end of the three-year period to identify the achievement or deviation from what is assumed As an INVALSI contact as an instrumental function Area 1 I worked with the NIV to define the actions to be implemented and to be included in the different sections of the improvement plan aimed at improving the results INVALSI: development and enhancement of human resources, curriculum design and evaluation, integration with the territory. The PTOF represents the project and the PDM the actions . However, the INVALSI tests cannot evaluate a pupil who lives in a motivational emotional relational dimension, so other activities that make the school a well-rounded educational agency are included in the section of the expansion of the PTOF's educational offer. Separate meetings are held with the teachers of Italian Mathematics and English to reflect together on the results in the standardized tests and on the connection of the INVALSI Frameworks with the National Indications 2012; In addition, the material available on the site in the section tests and insights of INVALSI, the types of questions that affect the different levels of evaluation. In conducting these reflections, the ideas and materials provided by the magazine INVALSI Open, especially the training videos on the different themes, were very useful. Moreover in the departments I have illustrated the platform Gestinv whose fruition is deemed very easy to use in curricular times and it releases the teachers from those practical almost of training that do not help the improvement of the learnings.. This organizational model has become a constant in my institute and has been proposed to the managers of two schools in the area with a view to sharing Best Practice.
Particular interest has been aroused by the viewing of information videos accessible by INVALSI open that have offered food for thought and discussion on central issues of the teaching process/learning in the perspective of the verticality of the curriculum both in Italian and in math. The participants expressed their willingness to use the proposed material to address some critical issues that emerged from the analysis of the results received and reflect on the structuring of a possible vertical path to be included in their PDM around core foundations starting from from primary school to secondary school. The second Institute involved in sharing Best Practice was the scientific-classical high school "E.Torricelli". The manager Prof.ssa Anna Giugliano organized a training/sharing path on the proposed INVALSI data usage model. The course involved several teachers of mathematics, English, Italian, INVALSI contact person and department head. In this course it was underlined why the INVALSI data for their objectivity and comparability represent a certain indicator to measure the improvement of an institution and the legitimacy of the didactic-methodological choices used. The methods of reading and selecting the many data provided according to what you want to pay more attention to have been illustrated. The teachers came together in disciplinary groups and in peer-to-peer training they focused on the processes underlying specific questions, connecting them to the thematic groups addressed in their lessons, identifying the most marked shortcomings . Also in this school they confirmed the validity of the proposed model and the tools made available by INVALSI that have been illustrated . In both schools I found a lot of interest and we worked in a climate of serenity and with the will to provide pupils with skills suitable to operate as conscious citizens.

***Keywords:*** RAV, PDM, model, tools, data, best pratictice

# AUTHORS

***Giovanni Abbiati*** is currently research fellow at the department of economics and management of the University of Brescia. His main research interests are related to methodology of the social sciences, educational inequalities education and evaluation of public policies.

***Tommaso Agasisti*** is a Professor at the School of Management, Politecnico di Milano. He also serves as the Associate Dean for International Relations and Quality at PoliMi GSOM - Graduate School of Management. His research interests lie in the field of Public Economics and Finance, Public Management and Policy, and Public Administration, with a particular focus on schools, universities, and local governments.

***Carmen Aina*** is a Professor at the School of Management, Politecnico di Milano. He also serves as the Associate Dean for International Relations and Quality at PoliMi GSOM - Graduate School of Management. His research interests lie in the field of Public Economics and Finance, Public Management and Policy, and Public Administration, with a particular focus on schools, universities, and local governments.

***Anna Maria Ajello***, full professor of Developmental and Educational Psychology at the Sapienza University of Rome, retired. He has carried out research about learning processes, skills and assessment. President of INVALSI for two terms and, in the autonomous province of Trento, President of the Evaluation Unit of the European Social Fund.

***Koray Aktas*** is a junior assistant professor of economics at Ca' Foscari University of Venice. Previously, he worked as a postdoctoral researcher at the University of Milan-Bicocca and Catholic University of Milan. He obtained his PhD in Economics at the Catholic University of Milan in 2017 and his research interest mainly lies in the field of empirical labor economics with a particular focus on the income dynamics and inequality; and economics of education.

***Filippo Albertini***, teacher of Computer Science and Technology Laboratory at the "Belluzzi-Da Vinci" Technical Industrial Institute in Rimini since 2002, also engages in freelance work providing IT consulting and analysis and development of software solutions. He is a trainer in courses for teachers and school staff on tools and data analysis for evaluation.

***Lorenzo Alderighi*** is a PhD student at the European University Institute. In the 2022 he won a scholarship for a 6 months apprenticeship at the Regional Economic Analysis Division of the Bank of Italy in Florence.

***Maria Francesca Ambrogio*** is a Primary School teacher at the I.C. of Santena (TO), she is the Instrumental Function for the Institute Evaluation and Self-evaluation. She is included in the national list of expert trainers of O.M. 172/20. In her Institute she holds the role of contact person for Training, INVALSI, DDI, Innovative Teaching Methodologies, Stem and Classrooms 4.0. She is a Welcoming Senior Tutor and she is part of the Institute working group for the PNRR.

***Josep Amer-Mestre*** is a researcher at the Competence Centre on Microeconomic Evaluation (CC-ME) of the Joint Research Centre (JRC). Prior to joining CC-ME, he obtained a PhD in Economics at the European University Institute and worked for the Research Department of the Bank of Spain. He has worked on a range of research projects, mainly in the areas of education and employment.

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

***Giovanni Antonini*** is a master student at the Catholic University of the Sacred Heart in Milan in Public Policy. He holds a three-year degree in Economics, Markets and Institutions from the University of Bologna. He did a curricular internship at the office of Associate Professor Gianluca Argentin where he did a thesis on teacher precariousness and inequalities in student learning. (New evidence from INVALSI data, 2017-2022.

***Graziella Arazzi***, Professor of Philosophy and History at the Liceo Human Sciences and Art of Imperia Collaborates with INVALSI (Area 5 Innovation and development). Latest publications: School and territories: plots of coevolution to reinvent the world, in ""IUL Research"", vol. 3, no. 5 (2022), p. 230-239.

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

***Massimo Armenise*** is a researcher at the National Institute of Statistics (Istat). He works on the evaluation of public policies and on territorial indicators for development policies related to the economy and to research and innovation. He graduated with honours from the University of Bari. He holds an M.Sc. in International Economics from the University of Rome Tor Vergata and a PhD in Theories and Methods of Individual and Collective Choice from the University of Bari. He has worked as a civil servant at the Ministry of Justice and as a researcher at the National Institute for Foreign Trade (ICE) and the Manlio Masi Foundation.

***Davide Azzolini*** is a research fellow at the Research Institute for the Evaluation of Public Policies of the Bruno Kessler Foundation (Italy) and an affiliated scholar at the Urban Institute (Washington DC, US). He has carried our research in the areas of student achievement, migrant education, financial aid and education technology.

***Stefano Babini*** teaches mathematics and physics. He is passionate about problem solving, didactic communication and new technologies applied to teaching. He deals with learning processes and assessment in various educational and system contexts. He is part of the research group in mathematics didactics 'Divertical-Math'. He has been collaborating for years with INVALSI. He collaborates with the University of Parma.

***Barbara Baldazzi*,** Researcher of Socio-Demographic Statistics Area at ISTAT since 1997. Researcher in SDGs Project: "Sustainable Development Goals". The United Nations Statistics Division entrusted Istat with the task of coordinating the production of indicators for measuring sustainable development and monitoring its objectives. In particular, in this project, I deal with analyzing, proposing, improving and monitoring statistical measures on health and weelbeing (Goal 3) and on education (Goal 4). Researcher in BES Project: "Measuring Equitable and Sustainable Well-being in Italy". In particular, I have the coordination and organization of activities of the thematic work groups of Education and Training. Project Manager of "Adult Education Survey" (2017 and 2012), in two waves of survey on households for study the participation of the adults at education and training during the life. Lecturer at Tor Vergata University since 2008.

***Rosario M. Ballatore*** holds a degree in Economics at the University of Bologna. Since 2011 he is an economist at the Regional Economic Analysis Division of the Bank of Italy in Cagliari. Starting from 2022 he is on charge of the Division.

***Paolo Barabanti***, PhD in Evaluation of Educational Processes and Systems. He has been a teacher at primary school and lecturer in Sociology of Education; now he is a researcher at INVALSI. His main studies are about learning outcomes through standardized tests, high achieving students, multicultural schools and the use of INVALSI tests data returned to schools.

***Gian Paolo Barbetta*** is associate professor of Economic Policy at the Catholic University of Milano.

***Martina Bazzoli*** is a research fellow at FBK-IRVAPP. In 2008 she obtained the Master degree in "Demographic Social Statistics" at Faculty of Statistics (University of Padua), with a dissertation titled "Measure and perception of disability among elderlies: an European comparative analysis". Before coming to IRVAPP, Martina worked as a researcher at the Social and Economic Permanent Observatory (OPES).

***Piera Bello*** is an Assistant Professor of Economics at the University of Bergamo. She holds a PhD in Economics from the Università della Svizzera italiana (USI). Before joining the University of Bergamo, she worked as a postdoctoral fellow at the University College London and at the University of Zurich. Her research has focused on gender and education economics, and household finance. Her current research agenda aims at understanding the role of teachers' stereotypes in affecting the performance and education choices of students, and at investigating the determinants of female under-representation in academia.

***Andrea Bendinelli*** got master degree in Statistics and works at INVALSI's statistical service. He carries out statistical analysis activities on large databases and conducts research activities in the assessment of students learning.

***Pierina Giovanna Bertoglio*** is internship supervisor at the Degree Course in Primary Education Sciences in Reggio Emilia for which she coordinates the projects Learning to learn, Playing with juxtapositions and overlaps to learn to argue. She is concluding the PhD Reggio Childhood Studies with a research project on cultural heritage and citizenship.

***Alice Bertoletti*** is a researcher at the Joint Research Center of European Commission in Seville (Spain). She holds a PhD in Management Engineering from Politecnico di Milano (Italy), with a dissertation on the evaluation of the impact of higher education systems on regional economic development. Her research interests involve the field of economics of education and the use of quantitative methods to evaluate educational policies.

***Francesco Bogazzi*** is a student in Statistics at the University of Florence. In the summer of 2023 he carried out an internship at Irpet (Istituto Regionale per la Programmazione Economica della Toscana).

***Giorgio Bolondi***, currently hold academic tenure as a Full Professor in Mathematics at the LUB, a position which he formerly held at the Alma Mater Studiorum-Università di Bologna and the Politecnico di Milano. His scientific research began in the field of mathematics, specifically complex geometry. His PhD thesis at the University of Nizza led him to the study of algebraic geometry, particularly the classification of projective curves. The last part of this research involved a study of larger geometric objects (surfaces of four-dimensional spaces and two-co-dimensional projective sub-schemes). In the same period, he began to look at historical and epistemological issues, particularly concerning didactic problems. Since 2000, his research has been centred upon Mathematics Education and the training of teachers.

***Alessandra Boscolo*** is a researcher in Mathematics Education at the Department of Mathematics of the University of Genoa, and a CRESPI collaborator. After graduation in Mathematics, she did an international double-degree PhD at the LUMSA University in partnership with the Australian Catholic University, with a dissertation in Mathematics Education.

***Hugo Armando Brito Rivera***, PhD in Psychology of Interaction, Socialization and Communication (University of Rome La Sapienza). Researcher in the project "Learning to learn in Italy, Europe and Latin America, coordinated by INVALSI. Member of the Mexican Network of Researchers in Organizational Studies (REMINEO) and of the Inter-American Society of Psychology (SIP). Independent Researcher, Research Professor at UNIPUEBLA and International Consultant at UNICEF.

***Enzo Brox*** is a postdoctoral researcher at the Swiss Institute for Empirical Economic Research (SEW) of the University of St. Gallen. He received his PhD in Economics from the Graduate School of Decision Sciences of the University of Konstanz in 2021. His research focus is on education - labor - and sport economics.

***Sergiu C. Burlacu*** is a research fellow at the Research Institute for the Evaluation of Public Policies of the Bruno Kessler Foundation. He has obtained a PhD in Development Economics at the Universities of Trento and Florence, Italy, and has a MA in Economics from Central European University. His main research interests are social inclusion, human capital formation, the psychology of poverty, family economics,

violence and decision making, and behavioral public policy. His research is based on impact evaluation studies, field experiments and lab-in-the-field experiments.

***Antonietta Cammarota***, Director of CEA (Centre for Environmental and Sustainability Education) in the Picentini area, expert facilitator for outdoor education and member of WEEC network (World Enviromental Education Congress – Networking Italia). She is involved in several activities on environmental and sustainability education for CEA and Asvis and in PON courses. She is an expert in statistical analysis.

***Ornella Campo***, Technical Director of USR Sicily. Self-assessment and school system valuation expert. She collaborate with the "INVALSI" Institution as member of "NEV" member ("Nuclei Esterni di Valutazione" or "external valuation nuclei ). Experienced trainee of learnings' valuation and member of the Regional support team for the implementaion of National guidelines, inclusion and introdution of innovative processes for schools.

***Emiliano Campodifiori***, graduated in Statistics and Economics at the University of Rome "La Sapienza". Currently he works in the Statistical Service of INVALSI, he performs statistical analysis of the National Assessment data.

***Lorella Camporesi*** holds a degree in Foreign Languages and Literature and a degree in Pedagogy from the University of Urbino, was a literature teacher until 2012 and is currently the school principal at the "A. Bertola" lower secondary school in Rimini. As a member of the Provincial Evaluation Unit, she is involved as a course director and trainer in courses on evaluation and self-assessment for Educational Institutions in the Rimini province.

***Marta Cannistrà*** is a PhD student in Data Analytics and Decision Sciences at the Politecnico di Milano - Department of Management Engineering. Among her main research interests, there are the use of statistical methods to predict the students' performance and the evaluation of educational interventions aimed at improving the educational careers of the students most at risk.

***Stefano Cantalini*** is Assistant Professor at the Department of Social and Political Sciences at the University of Milan. With a PhD in Economic Sociology and Labour Studies, he conducted empirical research particularly in the area of social stratification and demographic change. His main research interests include family formation and fertility, especially in their relations with occupational careers and social inequality. He is also interested in topics related to migration and gender inequalities.

***Roberto Capone*** is a chemical engineer, mathematician and psychologist. He is currently a researcher at the Department of Mathematics at the University of Bari Aldo Moro, where he teaches Methodologies and Technologies for Mathematics Education and Calculus. His main research interests concern interdisciplinarity in mathematics education, the use of technologies for the teachers' professional development, competency-based education, and the psychological aspects of mathematics Teaching-Learning.

***Maria Carbone***, Degree in Biological Sciences , with grade 110/110. Lecturer in mathematics and science in middle school from 1995 to the present.Instrumental function of area 1 and Invalsi reference for 13 years, NIV component for 5 years . I coordinate study groups and national, regional and PON projects. ICILS 2023 coordinator and test giver.

***Manuela Carchesio*** is a Mathematics and Science teacher for the lower secondary school. She is passionate about research and innovation in mathematics education. From 2006 to 2015 he carried out research activities in the scientific field, in Italian and foreign academic contexts.

***Roberta Cardarello***, Senior Professor of Methodologies of teaching and special education at the University of Modena and Reggio Emilia, where she currently directs the Research Centre "Teachers and Teaching Innovation" (CERIID). Many of her studies have focused on the promotion of reading and text comprehension, and teacher training.

***Giuseppina Maria Cardillo***, Teacher of Mathematics and Sciences, she is coordinator of the NIV Internal Nucleus of Evaluation and instrumental function of the PtOF. Trainer for newly hired teachers "School evaluation", tutor of the M@t.abel, PQM, VSQ plans and PdM aware observer for Emilia-Romagna I and II year Project P66 "Improvement desk". Collaborates with INDIRE and Giunti.

***Michele Cardone*** holds a degree in "Statistics for demographic and social sciences" and a Master (I level) in 'Statistics for the management of information systems' (Università di Roma "La Sapienza"). Working for INVALSI since 2004, member of the Statistical Service since 2010, mainly involved in the analysis of the school tests data and in the management of the annual data return to schools.

***Diego Carrasco*** is a full time academic researcher at the Centro de Medición MIDE UC, at Pontificia Universidad Católica de Chile. His research is devoted to the study of contextual effects and learning environments comparisons, using international large scale assessment studies (ILSA), specially for non-cognitive outcomes in civic education. He has served as a methodological expert for instrument design, scoring and/or results production using context background questionnaires in ILSA studies for UNESCO-OREALC, UNESCO-UIS, UNICEF-MENA, and the OECD. He received the IEA Richard M. Wolf Memorial Award in 2020, for its journal article "Civic knowledge and open classroom discussion: explaining tolerance of corruption among 8th-grade students in Latin America". He often participates as instructor of workshop on statistical modelling and results production with secondary data of ILSA, and as a lecturer of quantitative research methods.

***Giorgia Casalone*** is an Associate Professor of Public Finance at Università del Piemonte Orientale, non-resident fellow at Dondena Centre for Research on Social Dynamics and Public Policy of Bocconi University, member of the multidisciplinary Ph.D Program in Healthy Ageing Studies. She received a PhD in Economics from the University of Bologna and a Diplôme d'Etudes Approfondies from the University of Paris Nanterre. She participated as a researcher or as unit responsible in research projects funded by the European Commission, the Italian Ministry of University (PRIN), Fondazione Cariplo, Fondazione CRT. She worked as a consultant for the Sevilla Joint Research Center of the European Commission. Her main research interests are in the fields of the public economics, labour economics and the economics of education.

***Clelia Cascella***, PhD in Research Methodology and PhD in Economics, is a researcher at INVALSI, Department of Test Development. Her main research interests, both at INVALSI and abroad where she served as a Marie Curie Fellow and as a Lecturer in Social Statistics and Research Methodology, are mainly about the study of gender differences in students' academic performance, and the intersectionality between gender and other contextual/environmental (e.g., attitudes) and personal (e.g., socioeconomic status) characteristics, from both a theoretical and a methodological perspective.

***Laura Castellana***, graduated in mathematics, has been Headmaster since 2019. She carried out research at the CNR and taught in the first and second grade school. She trainer M @ tabel, PQM, Didatec, Vales and PNSD consultant. She held the role of coordinator tutor of the TFA and adjunct professor at the University of Bari and Foggia.

***Juan Carlos Castillo***, Associate professor in the Department of Sociology of the University of Chile and principal investigator of the Centro de Estudios de Conflicto y Cohesión Social COES and of the Nucleo Milenio de Oportuniodades y Desgualdes Digitales - NUDOS. Doctor in Sociology of the University of Humboldt, Berlin and Magíster in Public Management of the University of Potsdam. Currently, he is Investigador Principal de Proyecto Fondecyt Regular Nº1160921 "Economía moral de la meritocracia y preferencias distributivas", and is in charge of the development and implementation of the Observatorio de Cohesión Social de COES. His areas of investigation include social deprivation, distributive justice, public opinion and civic education.

***Chiara Cataldo*** teaches English and English literature. She got her C2 level in march 2022. She taught in many schools of Piacenza, Verona and Parma. She is passionate about collaborative and cooperative learning. She has collaborated for years with UNIBO and UniSalento. She took part in the SIMELP (Simpósio Mundial de estudos de língua portuguesa) and she participated in an Audiovisual Translation course (2015).

She presented her study on “Scientific discourse through TV series: Dr. House case" in 2019. She is a legal translator, too.

***Giorgio Cavadi***, former Technical Director, is expert trainer in evaluation processes and school organisation.

***Sara Caviola***, Researcher at the University of Padua. Previously Marie-Curie research fellow at the University of Cambridge (UK) and lecturer at the University of Leeds (UK). She published several papers in important peer review journals on the topic of mathematical learning and mathematical anxiety.

***Arianna Cecchini***, President of the Valdera Area Educational Conference since 2019, Mayor delegate for educational policies for the Valdera Union since 2019, Mayor of Capannoli since 2014, former President of the Valdera Union 2019/22. Promoter of the ValVal project. Political experience in governance and educational policies.

***Rosalba Ceravolo***, PhD in Prosociality, Innovation, and Collective Efficacy in the educational and organizational contexts. She works at INVALSI research dept. – methodological and psychometric unit. Her research interests are protective factors during development from a relational and educational perspective.

***Zhijun Chen*** is a PhD candidate at the Department of Education at the University of Bath, UK. She also works with Dr. James KO at the Education University of Hong Kong, focusing on teaching assessment in multiple cultures. Her research interests include educational effectiveness, education inequality, and leadership.

***Maria Chirico*** teaches scientific support teacher in a Hotel Vocational School of Brindisi. She is a staff member of School Leadership, in particular for educational guidance, combating and preventing early school leaving. Invalsi observer. Training designer, coordinator and assessor of vocational training courses.

***Zaneta Chonteva***, PhD, works at the Bureau of Education Development, Ministry of Education and Science, Republic of North Macedonia, as a advisor for research and research projects.Also, Chonteva is engaged as an external associate at the American University College-Skopje for the courses: Introduction to Statistics for Psychologists and Educational Psychology.

***Alessia Cividin***, middle school teacher and assessment coordinator, holds a PhD in Territorial Planning and Public Policies. Her research interests lie in intercultural contexts and teaching methodologies for learning in technological environments, as well as the analysis of complex socio-cultural settings.

***Ellen Claes*** is an associate professor at the faculty of Social Sciences of KU Leuven, Belgium. In her work, Ellen takes a didactic perspective on political science exploring the roles secondary schools have in shaping democratic knowledge, skills and attitudes of young people. Recent studies focus on the civic and intercultural competencies of (student) teachers.

***Francesca Coppa*** teaches at Plinio Seniore High School in Rome, PhD student in Mathematics at Sapienza University of Rome.

***Cristina Cosci***, Former school principal. As an expert in organizational, individual evaluation and management of the main quality systems, she collaborates with the Department of Public Function for training, research and evaluation activities in public administration. She participated in the European team for the revision of the CAF Education Model 2021 edition. She has worked with INVALSI on school evaluation since the first Vales and VM experiments up to SNV. She has organized training activities for school leaders and teachers. Those activities concerned issues of evaluation and management and she carried them out collaborating with USR. On behalf of Piemonte USR, she has coordinated teams focused on the evaluation of school principals. She has published several articles for INDIRE ForDirigenti website and number of texts documenting the evaluation work of the Valdera territorial school system.

***Annalisa Cristini*** is Full Professor of Economics at the University of Bergamo where she is Deputy Rector for Welfare and Sustainable Development. She holds M.Phil. and D.Phil in Economics at the University of Oxford. Her research interests range from macroeconomics to labour market issues. Recently she has

addressed gender inequality in relation to implicit stereotypes in primary schools and university performance during COVID. Among the institutional activities, from 2015 to 2021 she directed CESC, the University center for socio-economic dynamics and cooperation; during her career she held the positions of Department Director, President of Master Course, member of the Academic Senate and of the Board of Directors of the University of Bergamo; for the three-year period 2018-21 she was appointed by the Bank of Italy as a member of the Commission for the assignment of Bonaldo Stringher scholarships.

***Franca Da Re*** has a degree in Work and Organizational Psychology at the University of Padua. Throughout her career, she has been a primary school teacher; in courses for foreign adults; educational psychologist. Subsequently, she was Headmaster and held positions at the Veneto Regional School Office. From 2014 to 2022 she was Technical Director (Inspector) of the Ministry of Education at the Veneto School Office. You have carried out a lot of teacher training activities on topics concerning didactic organization, methodologies, learning and system assessment. She has been a member of the National Scientific Committee for Kindergarten and First Cycle Education Indications at the Ministry of Education and other national work commissions. Since 2022, September, she has been retired. She is the author of numerous publications on topics concerning didactics, evaluation, teaching of civic education.

***Ettore D'Agostino*** teaches Maths and Physics. He is a member of the school Digital Innovation Team and contributes to computerize school processes, collect and elaborate data for school statistical surveys.

***Maddalena Davoli*** is a Postdoctoral Researcher at the University of Zurich (Department of Business Administration and Swiss Leading House on the Economics of Education, Firm Behavior and Training Policies). She received her PhD in Economics from the Goethe University Frankfurt in 2022. Her work is in the field of applied microeconomics, with a focus on economics of education. Most of her current research relates to inequalities and determinants of financial literacy. She has won several awards and grants for her work on the determinants of gender gaps in financial literacy.

***Paolo Davoli***, Graduated in Physics, PhD, former Technical Manager and coordinator of the Regional Inspection Service at the USR-ER, administrative manager of a territorial Office, school principal, teacher at high schools and Universities. He works on various areas of the Italian National Evaluation System (SNV), on school systems, technologies,with training and support for schools and networks.

***Alessandra De Angelis*** teaches mathematics and physics in secondary school. She is a PhD in energy. She is the Invalsi test contact person at the school she belongs to. She is teacher since 2005 at the Polytechnic Department of Engineering and Architecture (University of Udine) and she is part of the research group in physics education. He holds the role of regional secretary for Friuli Venezia Giulia of the AIF.

***Kristof De Witte*** is full professor at KU Leuven and professor by special appointment at Maastricht University. He is program director of the Educational Master in Economics, and director of the faculty research group 'Leuven Economics of Education Research' (LEER). The research interests of Kristof De Witte comprise education economics, performance evaluation, and political economy.

***Doriana delli Carri***, graduated in Architecture at the University of Rome "La Sapienza", she currently works in the Statistical Service of INVALSI as technical collaborator research institutes.

***Marta Desimoni***, PhD, is a researcher at the INVALSI. She is the scientific coordinator of the psychometrics and methodological unit of the INVALSI National Assessment. Her research interests are in the fields of educational and psychological measurement and latent variables modeling.

***Teresa Di Tullio***, graduated in Natural Science, is a primary school teacher at I° Circolo Didattico "E. De Amicis" in Modugno (BA) for 16 years. She was vice Principal from 2014 to 2017; from 2008/09 she is INVALSI Referent and a member of the evaluation team. She is Digital Animator and external observer for INVALSI.

***Anna Dipace*** is full professor in Experimental Pedagogy at the Department of Medical Sciences and Surgery of the University of Foggia. She is head of the university E-Learning Center. Since 2022 she is dean of the

IUL Telematic University. Secretary and member of the executive board of SIREM – the Italian Society for Research on Media Education. Her research topics include technology-enhanced learning and assessment, digital and blended learning environments.

***Silvia Duranti*** is a researcher at the Istituto Regionale per la Programmazione Economica della Toscana which supports the region in developing and evaluating public policies. Her main areas of interests are the analysis of the education system and the evaluation of the effectiveness of labour and education policies.

***Nurullah Eryilmaz*** is pursuing his PhD in Education at the University of Bath, UK and also, he is a Research Analyst at the IEA, Hamburg. He holds an MA degree from the University of East Anglia. His doctoral thesis focuses on studying educational inequality and its effects on student outcomes using International Large-Scale Assessments (ILSAs). His research interests focus on the domains of comparative education, sociological theory and psychometric analysis..

***Eleonora Faggiano*** is currently an Associate Professor in Mathematics Education at the Department of Mathematics of the University of Bari Aldo Moro. She obtained her master's degree in Mathematics, the Qualification for High School Teaching in Mathematics and Physics and her PhD in Computer Science, discussing a thesis on e-learning and educational technologies with application in mathematics education. Her main research interests: use of digital technologies and their synergy with manipulatives, to construct shared mathematical meanings in the classroom; methodological and technological resources to foster teachers' collaboration and professional development concerning mathematics and STEAM; connection between National INVALSI standardized mathematics tests and teaching practices.

***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.

***Emanuele Fedeli*** is a Postdoc research fellow at University of Trento within the INEQUALITREES project funded by Fondazione della Compagnia di San Paolo. He defended the thesis at University of Trento. Previously, He studied at LUISS University and Collegio Carlo Alberto. His research agenda relies on two pillars. In the first, he focuses on education & health. I disentangle what "happens" in classroom, the role played by hierarchies, networks, and peers in shaping students' socioemotional skills, academic achievement, educational choices, and healthy habits. In the second, he studies the levels, drivers, and spacial distribution of unfair socio-economic inequalities and how it affects students' outcomes. He has recently won the INPS scholarship for young researchers.

***Marta Feliciani***, Primary school teacher at I.C. of Loreto Aprutino; graduated in Foreign Languages and Literatures. Member of the Internal Evaluation Unit participates actively the school life of our country. In the s.y. 2022/23 is engaged in the analysis of the discrepancy between school outcomes and results in standardized tests.

***Tommaso Feraco*** INVALSI Postdoc researcher at the University of Padova where he obtained his PhD in 2022. His main focus of research concerns learning, intelligence, and personality traits and skills in students. Published 19 papers in international peer-reviewed journals and presented his work in numerous national and international conferences.

***Francesca Ferrara*** is Associate Professor at Dipartimento di Matematica "G. Peano" of the University of Torino and researcher in mathematics education. She is author of several articles on national and international volumes and journals and a member of the Scientific Committee of the European Society for Research in Mathematics Education.

***Federica Ferretti***, PhD in Mathematics, researcher in Mathematics Education at the Department of Mathematics and Computer Science at the University of Ferrara. Her research concerns the Mathematics teacher professional development, didactic contract at all school levels, formative assessment in

Mathematics and the formative use of standardized assessment. For years she has been involved in Mathematics teacher's professional development.

***Federica Ferrini***, Professor of humanities in the first and second grade secondary school since 2011. Graduated in philology and literature at the University of Turin with a research thesis Il liber Psalmorum tra i Rerum Vulgarium Fragmenta: l'amore umano verso l'eterno. The most recent studies, which take as their starting point the results of INVALSI tests, concentrate on the teaching of Italian by competences, with particular attention to the valenziale grammar didactics method.

***Giovanna Filosa*** is a social psychologist and developmental psychotherapist with a psychodynamic focus. PhD in Psychology of communication, socialisation and interaction, as a technologist, she carries out research activities at the Social Economy and Migration Processes Section of Inapp in Rome.

***Brunella Fiore*** is Researcher at the University of Milan Bicocca, where she specializes in Sociology of the Family and Sociology of Education. Her main research interests focus on evaluating family dynamics and educational dynamics, teaching and the use of technology, with specific attention to learning.

***Sebastiana Fisicaro***, Latin/Greek teacher with long experience in training and networks, positions of responsibility (D.T. expert in planning. Coordinator NEV, Evaluation DS, SNV and OECD PISA. INDIRE trainer (PNSD, PQM, MULTILINGUAL, DIDATEC, Language in a European dimension), CLIL-TFA teacher -Catania University. APPLE Professional Development.

***Sara Flisi*** is a researcher at the Competence Centre on Microeconomic Evaluation (CC-ME) of the Joint Research Centre (JRC). Prior to joining CC-ME, she worked for the Centre for Research on Education and Lifelong Learning (CRELL) at JRC. She has worked on a range of research projects, mainly in the areas of education and employment..

***Arianna Fontanot***, Italian, Latin, History and Geography teacher since 2020 in Turin. Degree in Linguistic Sciences. Since 2020 she deals with Italian teaching: valential model and connections between sociolinguistic and linguistic learning.

***Roberta Franchi*** is a Language and Literature teacher. She has been Vice Principal for three years. During her university studies, she worked as a free-lance journalist for Il Messaggero and at the moment she is the Editorial Director of the School Magazine "L'Angolino". She has a large experience in National Operational Programmes on Education.

***Beatrice Franzolini***, PhD in Statistics, is currently a researcher and lecturer at Bocconi University in Milan. Her main research activity focuses on methodological statistics. Since 2020, she has been collaborating with RiminiInRete for the statistical analysis of school evaluations in the schools of the Rimini province.

***Michela Freddano*** is head of the Research Area School Evaluation at INVALSI where she has been researcher since 2013. PhD in Evaluation of Educational Processes and Systems, She is Adjunct Professor of Methodology of Action Research. Clinical interview and interview in organizational contexts at IUL Italian University Line.

***Pompea Funiati***, Literature teacher at IISS "E. Ferdinando" Mesagne (BR) Tutor of teacher training courses Expert in Europlanning FS area 2 "" Support for teachers' work and teaching coordination".

***Lucia Funiati***, Literature teacher at IISS "E. Ferdinando"- Mesagne (BR).

***Carmela Gabola***, Primary school teacher since 1996, referent teacher for the environment and sustainability, Senza Zaino, has been an expert and tutor in PON courses for creative writing, school counsellor and in charge of the new teachers' onboarding.

***Letizia Gambi*** is a doctoral researcher at the research group 'Leuven Economics of Education Research' (LEER) at KU Leuven. Her research focuses on the organization of schools, students' performance evaluation and educational policy evaluation.

***Alessandro Gambini*** is Associate Professor in Math Education at Sapienza University of Rome. He earned his PhD in Mathematics with a thesis in Analytic Number Theory. He has been involved in many research projects in the math education field, and has many years' experience as a teacher, conducting professional development, and developing mathematics dissemination activities.

***Annamaria Gentile*** is a PhD fellow in Reggio Childhood Studies at the Department of Education and Humanities of the University of Modena and Reggio Emilia and the Reggio Children Foundation. She is a member of the "Learning to Learn: L2L Project" research group and a junior member of CERIID, a research centre on teachers and teaching innovation.

***Andrea Giacomantonio*** "is an associate professor at the Pegaso Telematic University, where he teaches Experimental Pedagogy and E-learning and Media Education. From 2007 to 2021 he was researcher in General Didactics at the University of Parma, where he taught Competence based teaching and learning and Pedagogy of Marginality and Deviance. His current research topics are twofold. On the one hand, he is studying the soft skills and key competences assessment. On the other, he studies the grip that biopower has on competence-based teaching and learning, and assessment. Since 2000, he has been supervising and conducting teacher training courses about learning and school level assessment. He recently published on the topic of strategic competences: Bonazza, V., Giacomantonio, A. (2023). Strategic Competences and Academic Success. Linearity and profiles. In V. Bonazza (Ed.), The Hephaestus' Mark. Hypotheses, Models and Evidence About Competence-Based Teaching and Learning (pp. 61-93). Lecce: Pensa Multimedia; Bonazza, V., Giacomantonio, A. (2022). Disposizioni interiori, dispositivi valutativi e azione educativa. In A. La Marca, A. Marzano (Eds.), Ricerca didattica e formazione insegnanti per lo sviluppo delle Soft Skills. Atti del convegno Nazionale SIRD Palermo, 30 giugno, 1 e 2 luglio 2022 (pp. 774-788). Lecce: Pensa Multimedia; Giacomantonio, A. (2019). Valutare le soft skills: la struttura fattoriale del QSA-R. QTimes – Journal of Education, Technology and Social Studies, 9(2), 56-70.

***Letizia Giampietro***, Researcher at the INVALSI Innovation and Development Research Area and contact person for the Valu.E for Schools supporting the improvement of self-evaluation skills of schools. Her research focuses include educational policies, evaluation uses, school development and improvement, and educational and social inclusion processes.

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

***Patrizia Giannantoni***, PhD in Statistics and Demography from a multinational program with University of Rome and Lund and Max Plank Institute in Rostock. She has worked on psychometric evaluation of developmental tests in collaboration with CNR and University of Parma and participated in research projects on migration as research fellow at University of Naples. Since 2017 she has joined the Statistics Office of INVALSI, keeping her research interests on migrant integration, and educational inequalities.

***Sara Giannone***, graduated in Sociology at "La Sapienza" University of Rome. She is a research assistant at INVALSI and deals with the development and review of the English Listening and Reading standardised tests within the context of the INVALSI large-scale National Assessment.

***Marco Giganti*** is completing a PhD in Personal and Educational Sciences (Educational Research curriculum) and is a member of the research group of the Center for Studies and Research on Education Policies (CeRiForm) institutes at the Catholic University of the Sacred Heart. For several years he has been involved in empirical research methodology applied to the issues of formative evaluation and in general docimologie, pedagogical design, evaluation in personal services and training design for Vocational

Education and Training Institutes (IeFP). He is a member of several national and international scientific societies: the Italian Society of Educational Research (SIRD), the Italian Society of Pedagogy (SIPED), the Italian Evaluation Association (AIV), and the Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE).

***David Giofrè***, Currently associate professor at the University of Genoa. PhD. from the University of Padua, and lectured at the Liverpool John Moores University. Published over 70 papers in important peer review journals. Awarded with several grants, including Fulbright, JCPS, and PRIN. Keynote and chair for several international conferences and events.

***Fabrizio Giovannini*** is researcher at ISFOL - Institute for the Development of Vocational Training of Workers since 1996. He has been committed to the themes of methodological and socio-institutional strategies aimed at defining models, methodologies, and tools for the treatment of competences (training standards, transparency, and certification). Since 2022 he has been in charge of the Research Group Key competences for employability in the IVET, within the Training Systems Department of INAPP – National Institute for the Public Policies Analysis (formerly ISFOL), focusing his interest on the issues of training and assessment of the strategic competences necessary to face the challenges of life and work, conducting analysis and action-oriented research paths on policies and training practices in the VET context. On these topics he has published several articles, studies and research reports.

***Ivan Graziani*** teaches mathematics and science. Trainer in mathematics didactics. Expert in educational design and assessment. Passionate about problem solving and didactic communication. He is a member of the "Gruppo di Ricerca e Sperimentazione in Didattica della Matematica - Pisa" (GRSDM) and the "Divertical-Math" research group. He has collaborated for years with UNIBO, INDIRE, INVALSI and Mondadori-Rizzoli educational. He is a member of the Territorial Training Teams for the dissemination of the PNSD.

***Andrea Guarnacci***, received the degree in Literature from 'Universita' degli Studi RomaTre'. Teacher at the high school 'Manfredini' in Pontinia, collaborator of the school principal, INVALSI coordinator, coordinator of the Committee for the Self-Evaluation of the school between 2012-2020, member of the NIV and coordinator of the INVALSI Commission.

***Stefania Gubbiotti*** is associate professor in Statistics (SECS-S/01) at the Department of Statistics of Sapienza University of Rome, since 2018. She is involved in the activities of PLS (Statistics). Her research mainly focuses on Bayesian inference in the context of clinical trials.

***Alessia Ieva***, Teacher of humanities at Madre Mazzarello High school. In recent years she has explored the study of Italian language and the study of items from INVALSI in collaboration with the Casa degli Insegnanti. With the colleagues of the humanistic department, she focuses on the valenziale grammar method.

***Simona Incerto***, graduated in Philosophy at "Aldo Moro" University of Bari. She is a research assistant at INVALSI and deals with the development and review of the English Listening and Reading standardised tests within the context of the INVALSI large-scale National Assessment.

***Maria Magdalena Isac*** is a researcher at the Centre for Political Science Research at KU Leuven, Belgium. Magda's research is focused in the area of comparative evaluation of educational systems, with special emphasis on understanding how different formal and informal educational approaches contribute to young people's citizenship learning.

***Beti Lameva***, PhD is head of the Sector for Exams, IT and Research, National Examination Centre, Republic of North Macedonia. She is responsible for organizing state matura and national assessments, also she is the National Project Manager for PISA, TIMSS and TALIS studies. She has more than 20 years of professional experience in research in educational.

***Cristina Lasorsa*** works at INVALSI in the Research Area - Methodology and Psychometrics team. Her main research interests are in the field of the educational research, and particularly the proficiency evaluation in Maths.

***Carla Lavista*** is a teacher of Human Sciences. She worked for many years as a primary school teacher and since the 2021/22 school year she has been teaching Human Sciences at the Liceo of Città Sant'Angelo where she spent her training year as a newly entered teacher.

***Giuseppina Le Rose***, psychologist, psychotherapist and expert in psychological evaluation and counseling, currently works at INVALSI. She has performed numerous educational and vocational interventions and collaborated in the preparation of psycho-aptitudinal, cognitive and personality tests.

***Francesca Leggi*** graduated in Sociology, specializing in Economics, Labour and Organizations at the University of Rome "La Sapienza". Currently, she works at the Statistics Office of INVALSI, focusing on the statistical analysis on large databases.

***Melisa Diaz Lema*** is a Junior Assistant Professor at Politecnico di Milano, Department of Management Engineering. Her research falls within the intersection of public management, economics of education, and digital innovation. Her main research interest focuses on the usage of data analytics for decision making in the public sector.

***Yuan-Ling Liaw***, IEA Research Analyst, PhD in Educational Measurement & Statistics (2015, University of Washington). Research in psychometrics, test fairness, measurement invariance, ELLs. Expertise in item calibration, equating, bias detection, using simulations and real data to enhance test validity.

***Bigozzi Lucia***, Full Professor in Developmental and Educational Psychology, holds various institutional positions at the University of Florence. Many years of scientific and teaching activity on topics such as specific learning disorders and the identification of predictors of developmental dyslexia.

***Francesco Mammarella*** is an English teacher. He is currently performing his duties as Vice Principal. During his university studies, he worked as a free-lance journalist for Il Messaggero. He has a large experience in International School Projects.

***Simone Mancini***, External consultant for data analysis of school indicators. He began his work experience in the provincial school Observatory in Pisa (2001). Referent of the MISI project Models for Innovating Education Services promoted by UPI (2008/10) in the Provinces of Rome, Naples, Viterbo, Rieti, Pesaro. At the moment consultant for the regional school Observatory in Tuscany. He followed local projects of local authrities and school networks. He collaborates with the Emilia Romagna, Puglia, Lazio, Friuli Venezia Giulia and Umbria Regions on the topics of education and data analysis.

***Valentina Mandruzzato***, teacher of Mathematics and Sciences at the First Grade Secondary School at the Omni-comprehensive Institute of Città Sant'Angelo since the school year 2011/12. Special project teacher Area 1 s.ys. 2019/20, 2020/2, 2021/22, Member of the Evaluation Group Area 1 a.s.22/23, Head of Mathematics and Technology Department.

***Sara Manganelli***, PhD, Researcher at INVALSI. She has worked on the IEA PIRLS, TIMSS and ICCS international comparative surveys and on national school evaluation projects. She has conducted numerous studies and published in international journals on the development of psychometric tools and the analysis of large databases using multilevel techniques and structural equation models.

***Elisa Manzella*** is a research fellow at University of Brescia and a PhD candidate in Sociology at Catholic University of Milan with a thesis on the role of teachers in the reproduction of educational inequalities. His research interests include inequalities in education and school guidance, experimental evaluation of educational interventions.

***Lorenzo Maraviglia***, PhD in Sociology (title obtained at the University of Pisa). In the last twelve years I have directed the statistical office of the province of Lucca and I have participated in various research projects in the social and statistical fields promoted by ISTAT and various research bodies. Actually I work as a statistical researcher at INVALSI.

***Emanuele Marcora*** has been Headmaster of the Istituto Omnicomprensivo Europeo of Arconate and Buscate (MI) since the 2019/2020 school year. With a degree in History, he has taught since the year 2001/2002 and is passionate about educational innovation in all its forms, particularly the construction of digital learning environments. He was invited as Keynote Speaker to the official closing event of the open public consultation for the New Digital Education Action Plan promoted by the European Commission - Directorate General for Education, Youth, Sport and Culture.

***Miriam Mariani***, research fellow at the Research Area School Evaluation at INVALSI. PhD in Design at Sapienza University of Rome, where she is Adjunct Professor of Information Design and Visual Communication. Her research area focuses on Information Design, Data Literacy and complex processes.

***Marina Paola Mariano***, Teacher of Primary School and FS Evaluation of the Institute; QUADIS project referent of her institute, TIMSS and ICILS referent of her institute. Invalsi external servator. Teacher trainer INDIRE New Evaluation of Primary School. Registered ANPE Pedagogists. Graduated in Pedagogy at the Catholic University of the Sacred Heart and postgraduate specialized in learning and cognitive-behavioral disorders at the same University.

***Alessandra Marrata***, degree in Environmental Science and Agronomy, teaches science at ISIS Follonica (Gr). She is Instrumental Function for curriculum, didactics and assessment, coordinator of Invalsi, STEM projects and ICDL. She has published articles about didactics of STEM subjects and teaches a course on IBSE method, digital learning and inclusion.

***Graziella Marrone***, Primary school teacher, collaborator of the headteacher at the I.C. of Loreto Aprutino and psychologist. Member of the Internal Evaluation Unit, promotes methodological reflection, study and actions aimed at raising the quality of teaching and good practices. He has carried out projects on inclusion, learning difficulties and disorders, school well-being.

***Michele Marsili***, graduated in Statistics at Sapienza University of Rome. He worked in Business Intelligence consulting, providing software development solutions for analysis and support for company's decision making in insurance and pharmaceutical industries. Since January 2018 he has been working in the Statistical Service of INVALSI.

***Sonia Marzadro*** is a research fellow at the Research Institute for the Evaluation of Public Policies (IRVAPP). Her research interests concern social inequality, with a focus on educational opportunities, and poverty phenomena. She is involved in the evaluation of programs and public policies related to the use of digital technologies in educational processes.

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

***Chiara Masci*** is Junior Assistant Professor at Politecnico di Milano, Department of Mathematics, in the statistical branch. Her research interests are in the development of innovative statistical methods in the area of mixed-effects regression and classification models, both parametric and nonparametric, and in their application in the educational field.

***Antonella Mastrogiovanni***, graduated in Psychology, at "Sapienza University of Rome". expert in developmental age and systemic relational theories. Senior researcher INVALSI, Head of Italian tests, deals with the construction of standardized tests, in the context of reading comprehension..

***Sabine Meinck*** is the co-head of the Research and Analysis Unit (RandA) and head of the Sampling Unit. She coordinates IEA research activities and supervises sampling-related tasks, supports the development of survey designs and the analysis of large-scale assessment data, and conducts related workshops.

***Michela Milioni***, PhD in Prosociality, Innovation and Collective Efficacy in the educational and organizational contexts. She works at the INVALSI research dept. – methodological and psychometric unit. Her research interests are protective and risk factors in psychosocial adjustment and maladjustment during development from an educational perspective.

***Daniel Miranda*** is a full-time research assistant professor at the Centro de Medición MIDE UC at the Pontificia Universidad Católica, research associate at the Center for Social Conflicts and Cohesion Studies - COES and adjunct researcher at the Millennium Nucleus of Opportunities and Digital Inequalities - NUDOS, Chile. His research focuses on citizenship education, political inequalities and their intergenerational transmission.

***Catalina Miranda***, Political Scientist and Master in Sociology from the Pontificia Universidad Católica de Chile (PUC). Student of the Doctorate Program in Education at the PUC and member of the Red de Politólogas. Her research topics focus on gender equity, political participation, intergroup attitudes, educational policies and social research methodologies.

***Anna Maria Moiso***, Italian and literature High school teacher. Supervisor for SIS Piemonte (languages). Trainer. Textbook author. Coordinator of teaching research groups. Autor of Invalsi tests. Supervisor of Stanza di Italiano for Casa degli insegnanti, Torino.

***Sara Mori*** has a PhD in Evaluation of Educational Systems and Processes; she is a psychologist and a cognitive-behavioural psychotherapist, and has been a lecturer in Test Theory and Techniques. She has worked at INDIRE since 2011 and has been an adjunct lecturer at IUL in Developmental and Educational Psychology since 2013. Her main research interests are the assessment and development of transversal skills and the development of students' motivation and well-being, with a focus on methodologies for the personalisation of teaching..

***Francesca Morselli*** is an Associate Professor at the Department of Mathematics of the University of Genoa, where she involved in pre-service and in-service teacher education programs. Her main research interests include: first approach to mathematical argumentation and proof, formative assessment, and the role of affective factors in the teaching and learning of mathematics.

***Jessica Niewint*** is the Head of Research on Innovative Laboratory Teaching Methods and New Technologies at INDIRE and a principal investigator in international and national research projects based on STEM education and laboratory teaching. Her main interest is in exploring the potential of technology to support and enhance the teaching and learning process. Her particular focus is on the relationship between teacher and learner in a student-centred learning scenario to promote the personalisation of learning in the classroom.

***Bianca Nocchiotti*** is a PhD student in Learning Sciences and Digital Technologies at the Free University of Bozen-Bolzano. Her main research interests include the didactics of mathematics, the use of technology and difficulties in mathematics.

***Ileana Ogliari*** received the degree in Literature from 'Universita' degli Studi La Sapienza' in Rome. Teacher at the high school 'Manfredini' in Pontinia, coordinator of the POF between 2012-2014, collaborator of the school principal since 2013, NIV member, expert in student orientation and coordinator of the Digital Team.

***Mariarosaria Orefice*** received the degree in Mathematics from 'Universita' degli Studi di Napoli'. Teacher at the high school 'Manfredini' in Pontinia, where she is a deputy in the logic-mathematic area and Invalsi Commission. She is also a member of NIV and of the Committee for the Valuation and school Self-Evaluation.

***Federica Origo*** is Full Professor of Economic Policy at the Department of Economics of the University of Bergamo, where she is also Delegate by the Rector for relations with schools and students guidance. She is

also Research Fellow at the Institute of Labour Economics (Bonn). After graduating in Economics and Business at the Catholic University in Milan, she earned an MSc in Economics from the University of Texas at Austin and a PhD in Economics from the Catholic University of the Sacred Heart in Milan. Her research interests focus on labour economics, economics of education, policy evaluation and gender economics.

***Francesco Orlando***, Graduated in Agricultural Sciences, former Technical Manager and coordinator of the Regional Inspection Service at the USR-ER, administrative manager of a territorial Office, school principal, teacher.He deals with technical and professional education, evaluation of schools and managers, training, state exams, support for schools and networks.

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

***Nazareno Panichella*** is an Associate Professor of Economic Sociology at the University of Milan. He serves as the Principal Investigator for the MEI project (Multidimensional Analysis of Migrants' Penalty on Occupational Achievement and Health in Italy) and coordinates the interdisciplinary research group on Social and Geographic Inequalities in Italy (GESI). He is Associate Editor of the "European Sociological Review (ESR)" and a member of the editorial board of "Polis - Research and Studies on Society and Politics". Additionally, he is the Scientific Coordinator of the Spring School NASP-ECSR and a board member of the ECSR (European Consortium of Sociological Research). His research interests include geographic inequalities, labor market dynamics, educational opportunities, and the sociology of the family.

***Ornella Papa*** holds a PhD in Methodologies in Social Science, and she is a researcher at INVALSI. Her work is in the Methodology and Psychometrics team. In her research, she focuses on the techniques and methods of social research as well as the evaluation process and machine learning model.

***Donatella Papa*** holds a Ph.D. in Methodologies in Social Science, and she is a researcher at INVALSI. Her work is in the Methodology and Psychometrics team. In her research, she focuses on the techniques and methods of social research as well as the evaluation process and machine learning model.

***Monica Papini***, graduated in Statistics, at the University of Rome “La Sapienza”, she obtained the II Level Master in Data Intelligence and Decision making strategies. She currently works in the Statistical Service of INVALSI where she performs support activities for statistical analysis on large data bases of national and international surveys on learning. National Data Manager for ICILS 2018.

***Davide Passaro***, teacher at B. Russell High School in Rome, Ph.D Student in School of Statistical Sciences, Sapienza University of Rome.

***Serafina Pastore***, PhD in ‘Instructional Design and Educational Assessment’ and Fulbright Research Fellow, is associate professor in the Research and Humanities Innovation Department at the University of Bari (Italy). Her research examines the complex intersections of assessment practice, teacher education and educational policy as operating within the context of school and university innovations.

***Salvatore Patera***, sociologist and PhD in Education. Currently I am Associate Professor (M-PED/03) at the Faculty of Interpreting and Translation – Università degli Studi Internazionali di Roma (UNINT). Since 2008 University teaching assignments (M-PED/03 and M-PED/04) in several universities in Italy and abroad. From 2017 to 2020, I was a post-doc research fellow at INVALSI on the topic: "Research in the field of learning-to-learn skill, educational innovation, also from an international perspective".

***Emmanuele Pavolini*** is Full Professor in Economic Sociology and Social Policy at the University of Macerata (Italy). As an economic sociologist he has focused his research interests in two fields: a) welfare state studies from a comparative perspective, with specific attention to the Italian and Southern European welfare states, inequalities in the access to welfare state provision; b) labour market research and economic development, with a focus on the changing roles of professionals, networks among enterprises, strategies for competitiveness, and the role of migrants in the labour market. He is a National Expert for Italy of the

European Commission Expert Network ESPN (European Social Policy Network). He is currently co-editor of the Journal of European Social Policy. Since 2020 he is also Affiliated Professor at the Scuola Superiore Sant'Anna. Among his more recent publications there are: The Dark Side of the Moon: il ruolo del welfare fiscale nel sistema di protezione sociale italiano; Policy Change and Partisan Politics: Understanding Family Policy Differentiation in Two Similar Countries; Comparing occupational welfare in Europe: The case of occupational pensions; The Matthew effect in childcare use: a matter of policies or preferences?.

***Andrea Pietrolucci*** is Post-Doctoral researcher within the GESI project (Geography and Social Inequality in Italy). He attained his PhD at the Doctoral School of Social Sciences of the University of Trento. His main research interests focus on wealth inequality, social mobility, and inequalities in educational opportunities.

***Giuseppe C. Pillera***, PhD, Researcher at INVALSI. He has been a research fellow at the University of Catania, lecturer on SSD M-PED/04 for the universities of Catania and Messina and visiting researcher at the Universidad de Sevilla. He is focused on evaluative issues and methodologies principally related to e-learning, teacher training, intercultural education.

***Carmina Laura Giovanna Pinto***, Teacher of Mathematics and Physics, seconded to USR MARCHE since September 2015, from 21/07/2022 she is PhD in Computer Science and Mathematics. She participated with an article in the volume "I dati INVALSI come strumento per l'innovazione e il miglioramento scolastico. IV Seminario" and is author of the text L'EQUITÀ DEL SISTEMA SCOLASTICO Currenti Calamo Ed. She is involved in the teaching of Mathematics and is a trainer.

***Giuliana Pinto***, Full Professor in Retired Developmental and Educational Psychology, holds numerous institutional positions at the University of Florence. Many years of scientific and teaching activity on topics such as the development of symbolic systems, literacy and school learning.

***Lisia Piovano*** is a Primary school teacher. She is a member of the Digital team and Digital Supporter of the Institute since s.y. 2021/2022. Special project teacher for Area 1 since 2019.

***Antonio Piscopo*** is the deputy CEO of Teach For Italy. His main focus is the organisation's impact, from the supervision of the whole impact chain to its monitoring and evaluation as Learning Organisation. He is also responsible for Research and Policy.

***Donatella Poliandri***, PhD and Senior researcher, she heads the Innovation and Development Research Area at INVALSI and the PON Valu.E Project. She is the PI of a Project of relevant national interest on the evaluation of School to work programs policies. Her research focuses on social research methodology, evaluation of educational policies and evaluation capacity building.

***Stefania Pozio***, master degree in Geology and PhD in Experimental Pedagogy, is responsible for the national assessment tests of mathematics. She is chief researcher at INVALSI and her research mainly focuses on the study of the errors by students answering the mathematical questions of the national and international assessment surveys.

***Simone Quartara*** is a Mathematics teacher at I.I.S. Calvino in Genoa. Graduated in Mathematics at the University of Genoa. Member of the DIVA (Didactics Inclusion Formative Assessment Argumentation) research and action group of the Department of Mathematics of Genoa.

***Simonetta Radi***, Manager of the Unione Valdera, Coordinator ValVal Working Group. Administrative experience in educational and school services management and school governance systems.

***Marianna Rasetta***, Primary school teacher at I. C. of Loreto Aprutino; graduated in Primary Education Sciences. As INVALSI referent and Member of the Internal Evaluation Unit is involved in data analysis and interpretation to build a vertical curriculum with the aim of improving students' learning levels.

***Francesca Resio***, graduated in Political Sciences at "Sapienza University of Rome". She is a research assistant at Invalsi and is responsible for supporting the construction and revision of standardized tests and research in the context of the National Italian Test.

***Enrico Rettore*** is Professor of Econometrics at the Department of Economics and Management of the University of Padova. Previously, he has been Professor of Econometrics at the Department of Sociology and Social Research of the University of Trento. He received a PhD in Statistics from the University of Padova. His research interests focus on programme evaluation within the broader field of drawing inferences from non-randomly selected samples. On these topics he has published in various journals, including the Review of Economics and Statistics, the American Economic Review, the Journal of Econometrics and The Journal of the Royal Statistical Society (A). He has been the project leader of research projects supported by the Italian Ministry of University and Research, the Italian Ministry of Economy and Public Finance, the Italian Ministry of Welfare and the European Commission.

***Patrizia Rinaldi*** is a post-doctoral researcher at the Institute for Migration Research (University of Granada, Spain), where she received her PhD (Cum Laude) in Migration Studies in 2021. Previously, she earned a Master's degree in International Migration and Development at the Pontifical University Comillas in Madrid, Spain.

***Ugo Rizzo*** is an associate professor of Economic Policy at the Department of Mathematics and Computer Science of the University of Ferrara. He is a member of SEEDS (Sustainability, Environmental Economics and Dynamics Studies), an interuniversity research center in the field of environmental economics, and part of the team of the UNESCO Chair on Education, Growth, and Equality at the University of Ferrara. His research interests include the economics of innovation, with a particular focus on entrepreneurship, the economics of science, and environmental innovation.

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

***Elisabetta Robotti*** is an Associate Professor at the Department of Mathematics of the University of Genoa. She is the scientific director of the Mathematics Education Research Laboratory devoted to issues concerning inclusive mathematics education and outdoor education in primary school.

***Luigi Umbero Rossetti*** is professor of Business Economics in high schools and lecturer at the University of Sannio. He holds a PhD in Management and Local Development. Registration ORCID 0000-0002-9922-1166 and ResearcherID ACA-0187-2022. Chartered accountant, auditor, expert trainer. Author of several scientific contributions. Digital animator and member of the USR Campania Territorial Training Team.

***Daniela Ruffolo*** School Headmistress since 2010, she is graduated in Russian language and literature. A secondary school English teacher from 1994 to 2010, she is a teacher trainer and Marchio Saperi auditor. She is the principal of Don Milani-Linguiti Comprehensive School of Giffoni Valle Piana, Senza Zaino School HUB for community pacts.

***Arduino Salatin*** is a PhD in Education Sciences. He has been Dean of the Salesian University Institute of Venice (IUSVE), aggregated to the Faculty of Educational Sciences of the Salesian Pontifical University of Rome. Since 2020, he is emeritus professor at the IUSVE. From 2007 to 2011 he directed the Provincial Institute for Inservice Training and Educational Research and Experimentation (Iprase) of the Autonomous Province of Trento. From 2007 to the present, he has been a consultant of MIUR on the State vocational upper secondary education and School–Work Alternation. From 2016 to 2021, he was a board member of INVALSI. He carried out several comparative researches on education and training systems, with a focus on VET.

***Chiara Saletti***, primary school teacher, graduate in Literary Subjects, Tutor Coordinator UNIFI, author of school texts and articles teaching mathematics. She collaborates with Giunti as a consultant on assessment. Expert trainer OM 172/2020 and NEV SNV. She deals with assessment with training acquired at MIUR, INDIRE, INVALSI and POLIMI.

***Citlalli Sanchez-Alvarez***, bachelor degree in psychology from the Autonomous University of Baja California, Mexico; Master's and PhD in Educational Sciences from the Institute for Educational Research and

Development (IIDE-Mexico) with a focus on large-scale educational/psychological constructs, language proficiency/certification assessment instruments. Teacher in educational and social sciences undergraduate and postgraduate programs since 1997. Instructor in professional teacher development programs since 1999 aimed at primary, middle school, and university level teachers for their skill development in: • the design of classroom and formative assessments, • teaching practices and strategies, and • student learning skills development. Research interests: • Development and validity of large-scale psychological constructs, and student achievement assessment instruments, using AIG and traditional methodologies. • Development and validity of student context questionnaires, and surveys that focus on professional development and the teaching profession. • Analysis of personal, school, and teacher factors associated with student achievement in national and international assessments, in subject matters such as Mathematics, Language, Science, and Civic education.

***Andres Sandoval*** is currently the Head of the Department of Education at the University of Bath, UK. Prior to that, he worked as Head of the Research and Analysis Unit at the International Association for the Evaluation of Educational Achievement (IEA) in Germany..

***Marta Santanicchia*** is researcher at INAPP (formerly ISFOL), she has been dealing with research projects on training-assessment and recognition of key competences within the research group *Key competences for employability* (Training Systems Department). Graduated in Education, Master in Design and assessment of training processes (University of Perugia; Catholic University of Piacenza), over the years she has engaged in formal technical groups, research, and comparative analysis on the transparency and certification of competences, on IVET training standards, and on the learning outcomes-approach of qualifications and frameworks (EQF-NQFs). She has published articles, research findings and contributions for institutional reports.

***Ketty Savioli*** is primary school teacher with a master's degree in Mathematics from the University of Torino. Responsible for mathematics education and assessment projects, she collaborates with INVALSI and TIMSS for the learning evaluation. She is member of the Italian Commission for Mathematics Teaching.

***Lucia Scipioni***, primary school teacher with a PhD in Humanities and currently a Research Fellow at the Department of Education and Humanities, University of Modena and Reggio Emilia. Her main research interests are in the pedagogical field and particularly in the promotion of critical thinking and learning-to-learn competence.

***Silvia Servili***, primary school teacher, graduate in Education. Expert trainer OM 172/2020 and referent at regional level on issues of assessment at elementary school. Trainer in innovative teaching methodologies in particular debate.

***Alice Severi*** degree in Environmental Science and Agronomy, teaches science at ISIS Follonica (Gr). She is Instrumental Function for curriculum, didactics and assessment, coordinator of Invalsi, STEM projects and ICDL. She has published articles about didactics of STEM subjects and teaches a course on IBSE method, digital learning and inclusion.

***Giuseppina Maria Sfasciamuro***, Teacher of Primary school in Turin since 1990. Degree in Primary Teaching and Psycology; master degree in Adolescence Psycology. In addition to teaching, she deals with education and project management for Cooperativa Arcote, that operates in the field of culture, education, art and music therapy.

***Daniel Shephard*** leads the NORRAG project on Missing Education Data and is a PhD candidate in Comparative and International Education with a specialization in Sociology at Teachers College, Columbia University. He has done research in more than 20 countries and has published in a variety of international journals. He also holds a Master of Science in Evidence-Based Social Intervention from the University of Oxford and is a former teacher.

***Mara Soncin*** is Junior Assistant Professor at Politecnico di Milano, School of Management. Her research interests are in the educational field, with a focus on (i) digital learning, (ii) school management and (iii) the use of quantitative models for the evaluation of public policies.

***Camilla Spagnolo*** is a researcher of Mathematics Education at the Free University of Bozen-Bolzano. Her main research interests include argumentation processes in mathematics, didactics by competence and teacher training. She is a professor at the University of Bologna and a member of the Nucleus of Research in Didactics of Mathematics (N.R.D.). She is involved in teacher professional development courses for all school levels.

***Linde Stals*** is a doctoral researcher and teaching assistant at the Centre for Political Science Research at KU Leuven, Belgium. In her research, she draws on political socialization and political psychology literature to investigate the meaning of political trust and distrust and how it relates to modes and degrees of political participation among adolescents.

***Luca Stella*** is junior professor at the Freie Universität and the John F. Kennedy Institute for North American Studies.

***Francesca Storai*** works at Indire, she is responsible for the research structure Evaluation of innovation processes. her research interests concern the evaluation of competences with particular reference to transversal ones and the evaluation of innovation and improvement processes. She is an adjunct lecturer at IUL- Telematic University of Studies, for the teaching: Early childhood activities.

***Maurizio Strazzeri*** is a postdoctoral researcher at the Department of Economics of the University of Bern. He received his PhD in Economics from the Graduate School of Decision Sciences of the University of Konstanz. His research focus is on education, migration and labor markets.

***Andres Strello*** is an IEA Research Analyst with expertise in education, inequality, and achievement, coupled with rich knowledge in statistical analysis of large-scale data. His global research publications reflect his dedication to promoting equity in education.

***Rolf Strietholt*** is the co-head of IEA's Research & Analysis Unit, overseeing analysts for studies like ICILS and TALIS. His research focuses on survey design, measurement, and effectiveness, published in journals. He edits reports, secures research funding from various sources, including the European Commission, German and Swedish funding bodies, foundations, and IEA R&D fund.

***Cristina Stringher***, PhD, INVALSI researcher, led a team for the study of Learning to learn (L2L) in Italy, Europe and Latin America, first international project by INVALSI. She also managed the Italian experimentation on preschool quality indicators for internal evaluation. Past member of the Informal Advisory Group and the Italian manager for the OECD Study on Social and Emotional Skills and Chair of the Working Group on Data Development at the OECD ECEC Network, she was invited expert by the European Commission in the working groups on ECEC quality indicators, on the EU Key Competencies and on L2L Frameworks.

***Roberta Strocchio***, Teacher of Italian, Latin and History since 1992. Ph. D. in Greek - Latin philology. Member of the work group of professor Italo Lana (Faculty of classics at the University of Turin): Seneca's bibliography of XX century; Seneca and youth. Publications: The meanings of silence inside Tacito's works, The dissimulatio of Tacito's works ...

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

***Adriana Timpone***, graduated in Education and training science at "La Sapienza" University of Rome. She is a research assistant at INVALSI and deals with the development and review of the English Listening and Reading standardised tests within the context of the INVALSI large-scale National Assessment.

***Andrea Tinterri*** is research collaborator for the IUL Telematic University and adjunct professor at the University of Modena and Reggio Emilia. He has a PhD in developmental neuroscience and M.Sc. in cognitive neuroscience. His main research topics are educational neuroscience and game-based learning.

***Enrico Toffalini*** Assistant Professor of Psychology at the University of Padova, with a PhD in 2016. Expertise in intelligence, learning disorders, memory, and quantitative methods for psychology. Published 75 articles in international peer-reviewed journals, including top-tier ones, and has a long track of presentations at national and international conferences.

***Marco Tonello*** holds a degree in Economics at the University of Milano-Bicocca. He obtained his PhD at the Chatolic University of Milan. From 2012 to 2017 he worked as economist at the Research Department of the Bank of Italy. Since 2017 he is an economist at the Regional Economic Analysis Division of the Bank of Italy in Florence.

***Solveig Topstad Borgen*** is a Postdoctoral Research Fellow at the University of Oslo. She is actively involved in multiple education projects, including the Children of Immigrants Longitudinal Study in Norway, the Fairness of Educational Attainment and its Measures in Norway, and the Intergenerational Mobility and Labor Market Inclusion. Her research findings have been published in esteemed journals like Social Forces, Demography, and the European Sociological Review.

***David Torres Irribarra***, Psychologist from the Pontificia Universidad Católica de Chile, holds a PhD from the School of Education at the University of California, Berkeley. Before embarking on his doctoral studies, he served as a researcher and analyst for national teacher assessment projects in Chile. His doctoral research was centered around the utilization of statistical models for measurement in the social sciences, with a he emphasis on the use of item response theory in educational contexts. In his professional career, David held the position of Information Technologies Director at the Berkeley Evaluation and Assessment Research Center and served as a researcher at the MIDE measurement center. Currently, he is an Associate Professor at the Pontificia Universidad Católica de Chile. In the realm of educational and social measurement, he has made significant contributions as a senior psychometrics consultant for several major national Chilean assessments. He also has international experience, having been involved in research projects and consultancies in Honduras, Peru, Mexico, and the United States. His primary research areas include the application of latent variable models to measurement contexts, with his work published in journals such as Educational Evaluation and Policy Analysis, and Measurement, the theoretical foundations of social sciences measurement, an area in which he authored the book "A Pragmatic Perspective of Measurement" (2021, Springer), and the use of digital technologies to enhance educational assessment, a field where he is actively involved in the development of the Berkeley Assessment System Software.

***Daniela Torti*** master's degree in Developmental Psychology, Education and Wellness and PhD in Social science. Since January 2022 she has been a technical collaborator at the Research Area School Evaluation from INVALSI.

***Valeria F. Tortora*** is a researcher at INVALSI, where she is National Manager for the International Association for the Evaluation of Educational Achievement (IEA). She is PhD in Comparative Education, the most recent research concerns the study of social inequalities, the variables connected to educational performance of students.

***Alessio Trevisan***, Italian, History and Geography teacher in Turin since 2019. Degree in Literature, philology and Italian linguistic (University of Turin) and in School Leadership and Clinical Pedagogy (University of Florence). Since 2017 he deals with education, orientatation, teaching studies. Co-founder of "EDB - Esperienze Didattiche dal Basso".

***Moris Triventi*** is Associate Professor in Quantitative Sociology at the University of Trento (with Italian habilitation for the role of Full Professor) and a member of the CSIS (Center for Social Inequality Studies). He is also Deputy Director of the PhD program in Sociology and Social Research and Research Delegate for the Department of Sociology and Social Research. He is Principal Investigator of the international project INEQUALITREES - A Novel Look at Socio-Economic Inequalities using Machine Learning Techniques and Integrated Data Sources and local PI (Unitn) of the PRIN project GESI - Geography and Social Inequality in Italy.

***Valentina Vaccaro*** is a PhD student at the University of Oviedo with a thesis in mathematics education. She is research collaborator at INVALSI since 2018. She carries out research and training in mathematics education. Her research interests concern the use of new technologies and games in the teaching-learning of mathematics.

***Lorenzo Vaiani*** is research fellow at Itinerari Previdenziali.

***Ester Valloreo*** is a professor of Mathematics and Physics. As a contact person of the school for Invalsi, she is engaged in the analysis and interpretation of data with the aim of improving the students' learning levels.

***Agnese Vezzani*** is a principal, now seconded to the University of Modena and Reggio Emilia, as internship supervisor. PhD in Humanities, her main research interests are didactic strategies focused on collaboration among students, conversation analysis and reading comprehension.

***Vettori Giulia***, PhD in Psychology, research fellow and contract lecturer at the University of Florence. She carries out teaching activities and research with a specific interest in the study of literacy processes, learning processes and reducing school drop-out.

Via Ippolito Nievo, 35 – 00153 – Tel. 06941851
Mail: protocollo@INVALSI.it PEC: protocollo.INVALSI@legamail.it Website: www. INVALSI.it
C.F. 92000450582