INVALSI DATA IN EDUCATIONAL RESEARCH

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

edited by
Patrizia Falzetti

FrancoAngeli

INVALSI

INVALSI PER LA RICERCA
STUDI E RICERCHE

INVALSI PER LA RICERCA

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

 ISBN 9788835179498

Copyright © 2025 by FrancoAngeli s.r.l., Milan, Italy. ISBN 9788835179498

INVALSI DATA IN EDUCATIONAL RESEARCH

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

edited by
Patrizia Falzetti

FrancoAngeli

 ISBN 9788835179498

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

Assistant Editor: Francesca Leggi.

Isbn: 9788835171409

Isbn e-book Open Access: 9788835179498

Copyright © 2025 by FrancoAngeli s.r.l., Milan, Italy. ISBN 9788835179498

Index

Introduction

by Patrizia Falzetti

Over time, the seminars on INVALSI data have increasingly highlighted the importance of using the data published by the Institute, also in combination with data from other bodies and institutions, which in various ways provide important elements to shed light on the Italian school system. Indeed, the possibility of comparing different data sources is both a necessity and an opportunity to enrich and integrate information on education.

Evidence of this can be found in the contributions collected in this volume. They come from the seventh edition of the event "INVALSI data: a tool for teaching and scientific research" (Rome, 27-30 October 2022), in which – and not only – the data provided by INVALSI are tools to study the field of education from different perspectives.

In the first chapter, the authors present the results of an interdisciplinary project aimed at investigating the link between the INVALSI tests in mathematics and the teaching-learning processes in this subject.

The next chapter describes the use of RAV and INVALSI data in and for school self-evaluation processes.

In the third chapter, as part of the reflection on the personalisation of learning pathways, the possibility of using some of the tools proposed by the National Evaluation System, such as the INVALSI tests and the RAV evaluation rubrics, is considered.

Based on a reflection on the results of the OECD PISA survey on financial literacy, the fourth chapter highlights how Generation Z deals with the issues of sustainability and money management.

The exploratory analysis of some statistical indicators from different sources has allowed the authors of the fifth chapter to explore the issue of the general loss of skills among students in their final year of upper secondary school, as evidenced by the results of the 2020/2021 national surveys.

The sixth chapter compares the ability of standardised and non-standardised assessments to predict student performance at university. Here the authors focus on the role of the standardised assessment system in achieving the objective of measuring the overall health of the Italian education system.

Finally, the last chapter deals with a very important issue: early school leaving. It is well known that interrupting the educational process is an obstacle to a country's economic growth and employment, thus undermining the foundations of collective well-being. The aim of the authors is to develop an indicator of ESL that is capable of going into ever greater territorial detail in order to show the territorial diversity of the phenomenon.

As a statistical service, we hope that this volume will not only help to increase knowledge about the potential of the data, but also stimulate curiosity and new research questions.

1. An exploratory study on the connection between teachers' professional development and meta-didactical conflict

by Camilla Spagnolo, Valentina Vaccaro, Eleonora Faggiano

This exploratory study investigates the connection between mathematics teachers' professional development and the meta-didactical conflict arising from the interpretation and use of INVALSI test data. Based on a survey administered to 526 primary school teachers (grade 5), the research examines beliefs, knowledge, and teaching practices related to standardized assessment data. Preliminary findings show that initial and in-service professional development plays a more decisive role than years of experience in helping teachers identify the causes of students' mistakes. The meta-didactical conflict, as described by Arzarello and Ferretti, particularly emerges from the gap between everyday teaching practices and the theoretical frameworks underlying INVALSI test design. The study highlights the need for targeted professional development programs to enhance teachers' ability to consciously integrate standardized test data into formative classroom practices.

Questo studio esplorativo indaga il legame tra la formazione professionale degli insegnanti di matematica e il conflitto meta-didattico generato dall'interpretazione e dall'uso dei dati delle prove INVALSI. Attraverso un questionario somministrato a 526 insegnanti di scuola primaria (grado 5), l'indagine analizza credenze, conoscenze e pratiche didattiche relative alla lettura e all'uso dei dati standardizzati. I risultati preliminari evidenziano che la formazione iniziale e in servizio risulta essere un fattore più determinante rispetto agli anni di esperienza nel supportare gli insegnanti nell'individuazione delle cause degli errori degli studenti. Il conflitto meta-didattico, descritto da Arzarello e Ferretti, emerge in particolare nella distanza tra le pratiche didattiche quotidiane e i quadri teorici che guidano la progettazione delle prove INVALSI. Lo studio propone riflessioni sulla necessità di percorsi formativi specifici per potenziare la capacità degli insegnanti di

integrare consapevolmente l'uso dei dati delle prove standardizzate nelle pratiche formative quotidiane.

1. Introduction and rationale

Data from the *Large Scale Assessments* can be considered as tools that can be used by teachers, not only from a systemic perspective, but also for the design and implementation of meaningful teaching and learning activities. Through the practice of formative assessment, INVALSI data can also be used to return detailed information to students about their learning and thus build reflective and metacognitive pathways that are functional for effective competency-based teaching (Wiliam, 2010).

Our study is part of a broader line of research aimed at improving a close link between standardized assessments and Mathematics education. Its aim is to find an effective way to merge standardized assessments' results, methods, theoretical frameworks and tools – that are designed in order to impact at a systemic level – into actions of teachers and schools (Doig, 2006; Looney, 2011). In order to fully acknowledge the potentials and educational objectives of standardized assessments, we need effective theoretical tools to interpret the quantitative data they provide and the macro-phenomena that emerge from the complexity of educational systems.

The use of the standardized assessments can truly improve the teaching and learning of mathematics only if it is able to give refined, culturally wide-ranging and operational information to policymakers, teacher professional development programs, curriculum developers, principals and teachers (De Lange, 2007). Our research moves exactly within this stream of thought: we investigated teachers' beliefs on standardized assessments and tools, how they read and interpret standardized tests and data and, in particular, the central role that analysis of standardized assessment data may have for teachers' professional development.

In particular, a preliminary analysis of our data showed that a conflict is generated by teachers' difficulties in interpreting INVALSI tests and in using them coherently with the framework on which the tests have been designed. For example, this conflict is generated when teachers rely on their personal teaching experience and classroom assessments rather than on the theoretical constructs and competencies underlying the INVALSI framework. This misalignment may lead to misunderstandings in interpreting students' results, thereby limiting the potential of standardized assessments as tools for formative assessment.

Such conflict consists of three components. The first highlights that many teachers have a perception of the difficulty of the INVALSI tests that diverges significantly from the national trend. The second component points out that many teachers assess their students' difficulties or evaluate their responses to the INVALSI tests in ways that starkly contrast with the national data. The third component, stemming from the first two, involves the inconsistent interpretations teachers have regarding the purpose of the INVALSI tests and their relationship and impact on the teaching practices.

This conflict is considered meta-didactical by Arzarello and Ferretti (2021), because it involves discussions about didactic processes such as assessment, students' competencies, and mistakes, rather than focusing on mathematical concepts themselves, which is typical of the epistemic or didactic conflicts commonly described in the literature.

In this paper, we explore the connection between teachers' professional development and the meta-didactical conflict, correlating them with mathematics teaching and aspects related to general teaching. Our findings obtained from a questionnaire completed by 526 grade 5 teachers indicate that specialized pre-service teachers' professional development in a master's degree program for Primary Education significantly enhances teachers' ability to understand the reasons behind their students' errors. While years of service do play a role, the professional development received is a more influential factor in helping teachers identify these errors effectively.

2. Research background and problem statement

In Italy, the evaluation tests of the education system are prepared and administered annually by the National Institute for the Evaluation of the Educational System of Education and Training (INVALSI) for specific grade levels (grades 2, 5, 8, 10, and 13). The assessment of students' knowledge and skills and the overall quality of educational provision in schools is periodic and systematic (www.invalsi.it).

The framework adopted by INVALSI assessment tests (INVALSI, 2018) is strictly connected to the *Italian National Curicular Guidelines*, includes aspects of mathematical modelling adopted in PISA research, and is developed according to results provided by Mathematics education research literature. These facts show the link between INVALSI tests and results from research in Mathematics education.

The alignment of the INVALSI mathematics assessment framework (INVALSI, 2018) with the *National Guidelines for the Curriculum* has been

highly relevant for research (Bolondi and Ferretti, 2021; Bolondi *et al.*, 2021; Bolondi *et al.*, 2019; Looney, 2011).

The INVALSI institute releases a large number of items so that all stakeholders can analyze the statistical data, knowing the specific requirements of the proposed questions. The results are issued annually and can provide important information for categorizing students' errors at a macro-level.

Moreover, in the case of primary education, the tests are made fully available to everyone on INVALSI's institutional website: the results of the national sample of each item are published and the results relative to each grade are given back to individual schools in their reserved area of the INVALSI website.

As already shown in various research studies (e.g. Ferretti and Bolondi, 2019), the results of the INVALSI surveys highlight didactic macro-phenomena that can provide very useful information on learning/teaching processes.

Since 2017, the "INVALSI Group – Didactics and Disciplinary Knowledge" of the SIRD (Italian Society for Educational Research) Observatory on General and Disciplinary Didactics, consisting of disciplinary experts and pedagogists, has been conducting a broad research project with the aim of investigating teaching-learning processes in relation to the INVALSI mathematics tests. In particular, the focus was on beliefs, attitudes and classroom practices of grade 5 mathematics teachers regarding the data made available by INVALSI. This type of investigation allows, among other things, to identify the professional development needs of teachers at national level, fitting perfectly with the aims of the SIRD[1]. Among the broader objectives of the research, in line with INVALSI aims, we find the need to understand whether and how data from the Large Scale Assessments can be useful in the implementation of formative assessment.

The questionnare built within the project of the SIRD Observatory attempts to better understand what kinds of information teachers predominantly derive from INVALSI tests; how it is used, and valued, in classroom practices in everyday life. The guiding questions that inspired our work were: What are teachers' perceptions of INVALSI tests in mathematics? What are teachers' beliefs about the impact they can have in teaching? What assessment practices do teachers actually implement in the classroom? What is the distance between INVALSI tests and mathematics teaching and learning processes in classrooms?

These objectives were intended to identify the teacher's professional development needs, particularly with respect to formative assessment, which is

[1] https://www.sird.it/osservatorio-didattica-e-saperi/.

the overriding prerequisite of those teaching and assessment practices, capable of supporting pupils' mathematics learning in primary school (Truffelli and Vannini, 2021). The use of INVALSI tests in the classroom can guide and support the learning of the students, who are building specific mathematical knowledge and skills, thanks to the fruitful use of mistakes, which are considered a useful resource for deconstructing misconceptions and restructuring knowledge. It is crucial, however, that this process is supported by teachers' beliefs, that they gain active awareness of its value, showing that they appreciate the diagnostic and formative function of assessment.

The main findings of this research project have been object of several works having national and international dissemination. For example, drwaing on the Try out of the study, Arzarello and Ferretti (2021) presented in the Fourth Seminar "INVALSI data: a tool for research" the notion of meta-dadactical conflict. Using the data obtained through the Try out, some questions have been modified and the questionnaire was administered again to a larger convenience sample (Main Study). Data from both the Try out and the Main Study were further processed using a statistical software for data analysis.

For the current study, the purpose of our correlational analysis is to test whether there is a link between the meta-didactical conflict – highlighted by Arzarello and Ferretti's (2021) and confirmed by Faggiano and colleagues (2023) with the Main Study data – and the number of years of service of the survey participants. More precisely we are interested in understanding if the professional development received is an influential factor in helping teachers to correctly interpret the INVALSI tests results and in using them coherently in their teching practices.

3. The research: design and administration of the questionnaire

For the broader study, the purpose of the empirical, descriptive and correlational survey was to analyse primary school teachers' knowledge, teaching experiences and beliefs in reading and interpreting INVALSI items and data, particularly in the area of mathematics.

The chosen research tool to collect data on both knowledge, beliefs and behaviour statements, consisted mainly of a structured questionnaire and allowed for a combination of the three points of view present within the research team: the one of general education, the one of disciplinary didactics and the one of the teacher-researcher.

Specifically, the aim of the questionnare was to investigate teachers' beliefs about the knowledge and skills measured by INVALSI standardized

tests; to explore the proximity/distance between the functions and content of INVALSI items, on the one hand, and teachers' beliefs and statements about teaching practices, on the other.

To this aim, we have identified three areas of variables according to which the questionnaire has been built: one specifically of Mathematics Teaching, one related to aspects of General Teaching, and one related to Contextual Data. In addition, there were also questions about teachers' ability to read INVALSI data; these variables are the dependent variables. The independent variables are those related to aspects of General Didactics that refer to constructs aimed at detecting the attitudes of teachers.

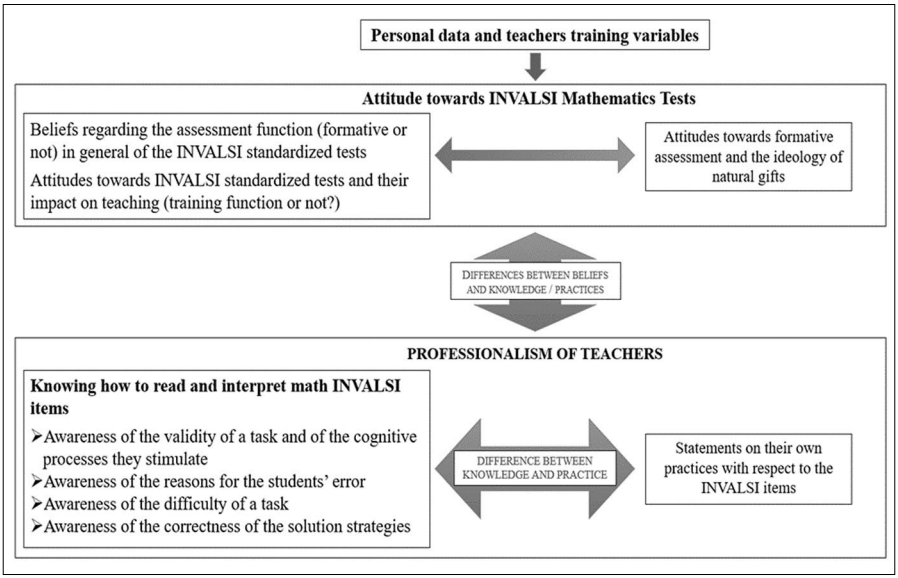


Fig. 1 – The framework of the research variables

In order to specify the different research variables in which we were interested, and the research hypotheses regarding the relationships between the variables, we constructed the framework of the variables under research represented in Figure 1.

The questionnaire consists of 51 questions. Most of them are closed-ended or ranking scales: some attitude scales come from previous validation studies, other batteries were constructed from scratch. Some open-ended questions allowed for some in-depth qualitative analysis.

According to the framework of research variables, the questionnaire consists of the following three sections:

– a first section related to mathematics teaching, which aims to investigate how teachers interpret (and know) INVALSI items and their results and how, if at all, they use them in their teaching practices;
– a second section related to teachers' beliefs and attitudes toward INVALSI tests and toward the use of assessment with diagnostic-formative functions;
– a third section collecting personal data and information on respondents' educational and professional backgrounds.

The first section presents seven INVALSI items, taken from the national tests for grade 5 (corresponding to the fifth and last year of primary school), in their original wording. For each of them, questions to teachers are designed to detect their pedagogical knowledge of mathematical content, the so-called and well known Pedagogical Content Knowledge, PCK (Shulman, 1986): misunderstandings, recurring errors, level of difficulty. Questions are also proposed on the proximity/distance that teachers perceive between the seven items and daily teaching practices and, also, the *National Guidelines for the Curriculum* (MIUR, 2012). Finally, respondents' perceptions of effectiveness were investigated in this section with respect to the same seven INVALSI items (do such items really succeed in testing certain skills?).

The second section proposes three sets of questions regarding: teachers' opinions on the INVALSI assessment program; the perceived educational usefulness of INVALSI tests, attitudes toward a diagnostic-formative use of classroom assessment and toward the ideology of natural giftedness (Ciani and Vannini, 2017).

The data collected in the third section refer to professional development as well as personal data of respondents. For example: seniority, initial training, continuing education, and institutional duties currently performed in the school.

The questionnaire was initially administered to 105 teachers (Try Out phase). This first phase allowed the questionnaire questions to be tested and the necessary modifications to be made in order to make it more valid and reliable. The new version of the instrument was administered to 427 teachers (Main Study), with a final total number of 526 valid cases. The geographical origin of the 526 primary school teachers who voluntarily took part in the research by answering the questionnaire[2] is mainly located (71%) in the North-Eastern part of the Peninsula. The data collected in the two administrations were coded and analysed using a statistical software for data analysis (IBM® SPSS® Statistics 27). The analysis conducted is descriptive and correlational, and the initial results are presented and discussed below.

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

In particular in this chapter we discuss our data with the aim to explore the connection between teachers' professional development and the meta-didactical conflict.

4. Example of teachers' perception of INVALSI item difficulty

Herein we present some early results of teachers' responses to the first section of the survey. INVALSI items and survey questions were translated by the paper authors from Italian into English for the ease of the reader.

To give a first insight into the richness of information gathered through the survey, we start with an example of INVALSI item within the questionnaire (Figure 2) that in 2009 resulted to be quite a difficult item for 5th grade students: indeed, only 33% of Italian students gave the correct answer.

D10. To which number does "12 tens, 7 tenths, 2 thousandths" correspond?

A. 12.702
B. 120.702
C. 12.72
D. 120.72

Fig. 2 – Item 10, Grade 5 Mathematics INVALSI test (2009)

We were interested in investigating teachers' understanding of the difficulty of this item. For this purpose, without informing participants about the percentage of the correct answers given by the students, we asked: "On a 1 (very easy) to 10 (very difficult) ranking, how difficult do you think the item is for 5th grade students?".

The 79.5% of teachers estimated the difficulty to be at most 5, hence, although the item required to manage a non-trivial conversation transformation between two different semiotic registers (Duval, 1993), we can say that this item was not considered to be difficult.

This result confirms what was found in the Try Out (Arzarello and Ferretti, 2021): teachers' perception of students' difficulties does not correspond to the INVALSI national data. Despite this discrepancy, results also confirm that the item is, among the seven used in the survey, the one which is considered the "most suitable for assessing learning" (with 86.2% of teachers which evaluate its suitability ranking it 3 or 4 over 4) and one of the "most

commonly used in assessment tests" (with 87.6% of teachers stating to be using this type of item in their assessment tests – ranking it 3 or 4 out of 4 in the relevant question).

5. The connection between teachers' professional development and the meta-didactical conflict

Thanks to the analysis of the third section of the Questionnaire (personal data and context information) it is possible to outline the characteristics of the sample. Although the sample, even large, cannot be considered to be representative (for example for its geographical origin), the data collected can give us a wide range of different information to reflect on.

An important feature of the sample is that 90% of the respondents are tenured teachers. Among the data collected in the administering stages are biographical age and years of service of the teachers involved. The correlation between biographical age and years of service is such that the years of service can be used for our analysis.

One of the components of meta-didactical conflict concerns teachers' ability to identify the reasons for students' errors, and it is this component that has a correlation, highly significant but with a Pearson coefficient that is not particularly high, with the years of service of the teachers involved in the study (see Table1).

Tab. 1 – Correlation of years of service with awareness of reasons for students' mistakes at INVALSI items

		Years of Service Category
Awareness of reasons for students' mistakes in INVALSI test	Pearson correlation	-0.143**
	Significance (two-tailed)	< 0.001
	N	526

Table 1 presents the correlation between teachers' years of service and their awareness of students' errors in the INVALSI mathematics tests. Years of service correlates inversely (as years of service increase, teachers' awareness of students' errors decreases) with students' awareness of the reasons for errors at INVALSI items in the related questions of the questionnaire, with significance less than 0.001. The negative Pearson correlation coefficient (-0.143, $p < 0.001$), indeed, suggests that as teachers' years of service increase, their ability to identify the reasons behind students' mistakes tends to decrease. Although Pearson's coefficient is -0.143 we find it significant to

focus on teachers who have been in service for a long time. While the correlation is statistically significant, its magnitude is relatively low, indicating that other factors, such as initial teacher education and professional development, may play a more critical role in shaping teachers' diagnostic skills (Viola *et al.*, 2024).

Moreover, in the analysis conducted the variable regarding the pre-service training of the interviewed teachers was also considered. The data collected allowed us to distinguish between respondents from the university route (specifically designed for access to teaching in primary school) and those who obtained a master's degree. In fact, in the third section of the questionnaire we asked teachers how they had gained access to teaching. The data confirm that specific pre-service training through the master's degree in Primary Education has a positive impact on teachers' ability to interpret the reasons for their students' errors. Years of service, therefore, would represent only one of the variables affecting the determination of meta-didactical conflict. Indeed, on the ability of teachers to identify the reasons for students' errors, initial training and professional development seem to have an impact above all.

The preliminary analysis shows, therefore, that the newly trained teachers, seem to have a greater awareness of identifying the reasons that mislead students. However, this finding could also be interpreted by taking into account how, over the past two decades, the pre-service teachers' training of primary school teachers in Italy has changed radically (Bolondi and Ferretti, 2021; Bolondi *et al.*, 2019; Looney, 2011). It is well known, in fact, that teachers trained before 2002 access primary school teaching through the "Diploma magistrale" (i.e., a high school-leaving certificate, ISCED level 3), while those trained after 2002 access primary school teaching through a master's degree in Primary Education (Truffelli and Vannini, 2021).

Our findings align with broader research suggesting that more recent teacher professional development, particularly those including structured training in standardized assessment interpretation, may contribute to a better understanding of student errors (Viola *et al.*, 2024). Therefore, professional development programs should consider integrating targeted training on assessment literacy to bridge this gap.

The interpretative hypotheses formulated here are currently being tested through further quantitative and qualitative investigations.

6. Concluding remarks

In this work we have discussed preliminary findings of a study on meta-didactical conflict indicating that (pre-service and in-service) professional development is a more influential factor in helping teachers identify students' mistakes effectively than years of service. This connects with meta-didactical conflicts teachers face when using and interpreting INVALSI tests.

This work is ongoing, between both quantitative and qualitative investigations. It seems clear that the debate between disciplinary didactics and experimental pedagogists, as well as the discussions with teachers and stakeholders, is producing interesting interpretations and developments, both from the point of view of the analysis of terms and concepts born in different cognitive and experiential fields, and from the point of view of understanding the meanings to be attributed to mistakes and the planning of intervention strategies that are shared, and not sector-specific.

It seems of fundamental importance not to lose, but to strengthen, the collaboration with teachers engaged in the field, through cognitive tools of different kinds: questionnaires, but also in-depth interviews to explore the thoughts, doubts, and contradictions of teachers, in order to understand more deeply the reasons for the relationship, which sometimes appears contradictory, between teaching practices and INVALSI tests.

It is, therefore, worth reflecting, from an educational point of view, on the conditions that make it possible for INVALSI tests to be used in the classroom in a way that is useful for promoting the development of competence in learners, in terms of real knowledge and skills in the field of mathematics. In addition, the focus on the use of INVALSI items as a tool for formative assessment in the classroom is one of the prerequisites for setting up appropriate teaching situations to promote learning at a good level and in an equitable manner, as well as to monitor its development as it goes along.

Finally, it should be emphasized that all the results that are gradually emerging from this survey will enable the interdisciplinary group of researchers to reflect and hypothesize ad hoc paths for teacher professional development, capable of considering all aspects – disciplinary, pedagogical, evaluative – useful for developing a teachers' awareness, that is not compartmentalized of the processes put in action, but networked and virtuous.

Acknowledgements

Authors are grateful to the voluntary teachers who participated in the survey and all the participants to the research project group "INVALSI – Disciplinary Didactics" of the SIRD (Italian Society for Educational Research). This group consists of disciplinary experts (belonging to the University of Turin, the Free University of Bolzano, the State University of Milan, the University of Ferrara, and the University of Bari Aldo Moro) and pedagogists (belonging to the Alma Mater Studiorum University of Bologna, the University of Turin and the Sapienza University of Rome) coordinated by Prof. Arzarello and Prof. Vannini.

References

Arzarello F., Ferretti F. (2021), "The connection between the mathematics INVALSI test and the teaching practices: an explorative study", in P. Falzetti (*ed.*), *INVALSI data to investigate the characteristics of students, school, and society. IV Seminar "INVALSI data: a research and educational teaching tool"*, FrancoAngeli, Milano.

Balconi B., Truffelli E., Maccario D. (2024), "The INVALSI tests for teaching: teachers' beliefs about the validity of INVALSI tests and teaching practices implemented", in P. Falzetti (*ed.*), *INVALSI Data Support in Policies and Schools. VI Seminar "INVALSI data: a tool for teaching and scientific research"*, FrancoAngeli, Milano.

Black P., Wiliam D. (2009), "Developing the theory of formative assessment", *Educational Assessment, Evaluation and Accountability* (formerly: *Journal of Personnel Evaluation in Education*), 21 (1), pp. 5-31.

Bolondi G., Ferretti F. (2021), "Quantifying solid findings in mathematics education: loss of meaning for algebraic symbols", *International Journal of Innovation in Science and Mathematics Education*, 29 (1), pp. 1-14.

Bolondi G., Ferretti F., Santi G. (2019), "National standardized tests database implemented as a research methodology in mathematics education. The case of algebraic powers", in U. T. Jankvist, M. van den Heuvel-Panhuizen, M. Veldhuis (*eds.*), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education*, Freudenthal Group & Freudenthal Institute, Utrecht University and ERME, Utrecht.

Bolondi G., Ferretti F., Spagnolo C. (2021), "Argomentare in Matematica: analisi di protocolli di studenti su catene di quesiti INVALSI proposti in diversi gradi scolastici", in P. Falzetti (a cura di), *I dati INVALSI: uno strumento per lo sviluppo delle competenze trasversali. III Seminario "I dati INVALSI: uno strumento per la ricerca"*, FrancoAngeli, Milano.

Ciani A., Vannini I. (2017), "Equità e didattica. Validazione di scale sulle convinzioni di insegnamento democratico", *Cadmo*, 25 (2), pp. 5-32.
De Lange J. (2007), "Large-scale assessment and mathematics education", S*econd handbook of research on mathematics teaching and learning*, 2, pp. 1111-1144.
Doig B. (2006), "Large-scale mathematics assessment: looking globally to act locally", *Assessment in Education: Principles, Policy & Practice*, 13 (3), pp. 265-288.
Duval R. (1993), "Registres de représentations sémiotiques et fonctionnement cognitif de la pensée", *Annales de Didactique et de Sciences Cognitives*, 5, pp. 37-65.
Faggiano E., Ferretti F., Arzarello F. (2022), "How do primary teachers interpret and use standardized assessment: the case of the crochet placemats", in *Twelfth Congress of the European Society for Research in Mathematics Education (CERME12)*, Feb., Bozen-Bolzano, Italy.
Faggiano E., Monaco A., Rizzo O.G., Vaccaro V. (2022), "An exploratory study on the connection between INVALSI assessment and mathematics teaching and learning processes at the Primary School level", in P. Falzetti (*ed.*), *The school and its protagonists: the teachers. V Seminar "INVALSI data: a tool for teaching and scientific research"*, FrancoAngeli, Milano.
Ferretti F., Bolondi G. (2019), "This cannot be the result! The didactic phenomenon "the Age of the Earth", *International Journal of Mathematical Education in Science and Technology*, 52 (2), pp. 194-207.
INVALSI (2018), *Quadro di riferimento delle prove di INVALSI di matematica*, retrieved on April 29, 2025 from: https://invalsiareaprove.cineca.it/docs/file/QdR_MATEMATICA.pdf.
Looney J.W. (2011), *Integrating formative and summative assessment: progress toward a seamless system?*, OECD Education Working Papers, No. 58, http://dx.doi.org/10.1787/5kghx3kbl734-en.
MIUR (2012), *Indicazioni nazionali per il curricolo della scuola dell'infanzia e del primo ciclo d'istruzione*, Roma, https://www.mim.gov.it/documents/20182/51310/DM+254_2012.pdf.
Shulman L.S. (1986), "Those who understand: Knowledge growth in teaching", *Educational Researcher*, 15 (2), pp. 4-14.
Truffelli E., Asquini G. (2022), "Concezioni della valutazione e dell'apprendimento in insegnanti di matematica nella scuola primaria", *Cadmo*, 1, pp. 51-66.
Truffelli E., Vannini I. (2021), "Convinzioni e atteggiamenti degli insegnanti di scuola primaria italiani e orientamento all'uso formativo delle prove INVALSI di matematica", in L. Lucisano (a cura di), *Ricerca e didattica per promuovere intelligenza comprensione e partecipazione. Atti del X convegno della SIRD*, Pensa Multimedia, Lecce, pp. 376-395.
Vaccaro V., Faggiano E., Ferretti F. (2021), "Consapevolezza degli insegnanti delle ragioni degli errori degli studenti in matematica: uno studio esplorativo nella scuola primaria", in L. Lucisano (a cura di), *Ricerca e didattica per promuovere intelligenza comprensione e partecipazione. Atti del X convegno della SIRD*, Pensa Multimedia, Lecce, pp. 411-430.

Vannini I. (2022), “Una valutazione che analizza e ricostruisce”, in *Rileggere oggi Didattica e docimologia*, Armando, Roma.
Viola G., Ferretti F., Gambini A., Martignone F., Soldano C., Spagnolo C. (2024), “Teachers’ professional development and mathematics LSA: first result of national project”, in T. Evans, O. Marmur, J. Hunter, G. Leach, J. Jhagroo (*eds.*), Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education, PME, Auckland, vol. 1, pp. 223-224.
Weiner B. (1985), “An attributional theory of achievement motivation and emotion”, *Psychological Review*, 92 (4), pp. 548-573.
Wiliam D. (2010), “An integrative summary of the research literature and implications for a new theory of formative assessment”, in H. Andrade, G.J. Cizek (*eds.*), Handbook of formative assessment, Routledge, New York, pp. 18-40.

2. *Achievement & Self-evaluation Questionnaires: A Tool for School Improvement*

by Marilena Salsano

Following the introduction of the National Assessment System, it became compulsory for schools to draw up strategic documents such as the Self-Assessment Report (RAV), the Improvement Plan (PdM), and the Educational Offer Plan (PTOF). The self-assessment phase is the principle of the entire process and it is fundamental to its success. Starting from the effectiveness factors (Scheerens, 2018), satisfaction/self-evaluation questionnaires were created to provide schools with further indicators. Questionnaires were administered to managers, teachers, ATAs, students (from the 5th year of primary school), and parents of 30 schools in the province of Cosenza. The results of the questionnaires were then used by the schools to draft the new RAV 2022-2025.

In seguito all'introduzione del Sistema nazionale di valutazione è diventata obbligatoria per le scuole la stesura di documenti strategici quali il Rapporto di autovalutazione (RAV), il Piano di miglioramento (PdM) e il Piano dell'offerta formativa (PTOF). La fase di autovalutazione, principio di tutto il processo, è fondamentale per la sua riuscita. Gli indicatori del RAV poche volte sono integrati da quelli delle scuole che invece potrebbero avere informazioni dai questionari di gradimento che somministrano a fine anno scolastico. Partendo dai fattori di efficacia (Scheerens, 2018) si sono creati questionari di gradimento/autovalutazione per dare ulteriori indicatori alle scuole. Sono stati somministrati questionari a dirigenti, docenti, ATA[1]*, studenti (dal V anno di scuola primaria) e genitori di 30 scuole della provincia di Cosenza. I risultati dei questionari sono stati poi utilizzati dalle scuole per la stesura del nuovo RAV 2022-2025.*

[1] ATA is the acronym for the administrative, technical and auxiliary staff of schools and is made up of the administrative and supervisory staff of pupils.

1. The school self-evaluation process

Organisational, social, economic, and geographical differences in individual schools influence the educational outcomes of students (Banerjee and Duflo, 2012). The disadvantaged situation of some educational institutions, in terms of digital devices, teaching levels, and geographical differences has also been confirmed by the pandemic and has proven the need for policy choices financed by the National Recovery and Resilience Plan (NRP) and the National Operational Programme PON "Per la Scuola" 2014-2020 – European Regional Development Fund.

These resources support an education system that should be characterised by managerial, organisational, and didactic autonomy, "good practices" of governance, flexibility, managerial efficiency, and the optimisation of so-called human capital. The following are important in ensuring quality improvement and quantitative growth in education and training settings: presence of the school headmaster/headmistress at the top of the school/college, with managerial tasks and objectives linked to standards of performance measures; introduction of the three-year formative offer plan; emphasis on the need to respond to the expectations expressed by families, local authorities, and the social, cultural and economic contexts of the local area; rewarding of teachers for organisational support or for "merit".

The demand for measurement and evaluation of the effectiveness of education systems and individual schools, accompanied by accountability is a direct consequence of autonomy, which requires transparency of the choices made at a management and teaching level, both by individual educational institutions and the school system. Schools have been given support tools to help them define the vision, mission, strategic choices, improvement plans, and analysis of results for continuous improvement, to positively affect pupils' performance. Through these process improvements, the schools should be able to become more resilient and prepared to respond more efficiently to challenging situations.

The evaluation system (SNV) requires schools to implement and document their improvement through performance procedures and tools and entrusts school leaders not so much with administrative tasks but with shifting the focus from day-to-day practices to improving teaching and learning. It is therefore up to DSs to decide whether to be rule-abiding bureaucrats or to be initiators of the processes of implementing educational policies for school improvement (Lipsky, 1980; Mintrop and Trujillo, 2007; Demerath and Louis, 2017). Then, in school improvement, teachers are involved with their professionalism and skills through a continuous process of transformation, cultural growth, involvement, and empowerment that invests the whole school.

The school seen as a public service with the capacity for autonomous planning through the tools provided by the regulatory authorities must strive for educational effectiveness. If these tools are lacking, the school has to be able to try to create them and verify them in the field. To self-evaluate effectively, it is important to understand the goals to pursue, the standards to meet, and, most importantly, the factors that enhance school quality and how individual schools measure up against those indicators. Self-evaluation requires the formulation of specific relevant questions based on which a reliable valid data collection must be set up (Hofman, Dijkstra and Hofman, 2009; Schildkamp, Visscher and Luyten, 2009; Vanhoof, De Maeyer and Van Petegem, 2011).

Schools, having them as a reference, can work on ineffectiveness, understand how to "do better" (Palumbo, 2019), and self-evaluate themselves to design paths of improvement. A correct self-assessment analysis needs a model to strive for and has educational effectiveness as its ultimate goal.

2. Educational effectiveness factors as a reference for self-evaluation

The general promotion factors influence student achievement at the level of the educational institution (Edmonds, 1979; Levine and Lezotte, 1990; Scheerens and Bosker, 1997, Sammons, 1995, 1999; Marzano, 2000, 2003) and have been grouped (Scheerens, 2018) into:

- performance orientation and high expectations;
- educational leadership;
- consensus and cohesion among staff members;
- quality of the curriculum/learning opportunities;
- school environment;
- evaluative potential;
- parental involvement;
- feedback and reinforcement.

The indicators are made available to educational institutions and serve to provide elements of reflection in a self-assessment key for the RAV. They include the following kinds of data:

- data processed and returned by INVALSI to the school based on its answers to the school questionnaire;
- data collected from the administration itself and from the SIDI system;
- data derived from statistical sources (ISTAT, Ministry of the Interior, etc.);
- data derived from the results of standard surveys on learning levels carried out by INVALSI;

– possible supply of data (indicators) proposed and processed by the school itself.

Considering the priority of external indicators, the limitation of the RAV thus constructed is the focus evaluation on these indicators and the difficulty of carrying out self-assessment and improvement processes. These processes invest the entire organisation: a shift of attention towards the schools, and towards their specificity and ability to analyse processes, is therefore necessary.

The difficulties in implementing self-evaluation at the school level need to be found in the evaluation culture present in schools, in the socialised models of teachers' professional interpretation, and in the relations between the school and the social stakeholders directly concerned.

The teachers often are not able to find themselves in the RAV as a representation of their organisation's identity: the educating community does not construct the strategic documents as a priority commitment to pursue the path of self-assessment and improvement.

The indicators of the RAV cannot analyse the totality of the effectiveness factors and do not always succeed in promoting a collective self-diagnostic reflection in the school organisation.

It is necessary to provide schools with tools that give them the possibility to:
– have indicators that cover all the sub-components of the educational effectiveness factors;
– involve all stakeholders in the self-evaluation process;
– promote shared and participative self-evaluation.

Self-evaluation is an ongoing process of transformation and cultural growth. The central role entrusted to INVALSI and the lack of involvement of schools in the design, management, and use of learning surveys, has accentuated the mistrust and resistance of school operators: schools tend to be subjected to the demand for assessment. This demand is often not recognised by the school as an opportunity for growth and improvement (Castoldi, 2010).

The following analysis investigates the general Scheerens effectiveness-promoting factors at the level of educational institutions, deliberately leaving out those at the classroom level.

In the following tables, the first column (component of efficacy factors) correspond to Scheerens, while in the second (sub-components and exemplary items), the items have been modified exclusively to adapt them to the Italian school context. This confirms that the factors concerning education have almost the same components, regardless of the education system to which they refer.

The third column (RAV indicator/questionnaire) shows the corresponding RAV indicator, and the fifth column shows the reference index.

Tab. 1 – Correspondence between the effectiveness factors and the indicators of the RAV/Satisfaction-self-assessment questionnaires for example the component 3.2 by factor 3 (Consensus and cohesion among staff members)

Components	*Sub-components and exemplary items*	*RAV Indicator/Questionnaire*	*Index*
3.2 Importance related to cooperation	Measured by three scales: time devoted to collaboration between teachers (time), improvement of teaching through discussion (innovation), promotion of teacher participation (participating)	3.5.b Human resources management	>25%-50%
		3.5.b.1 Degree of participation in the organizational model	>50%-75
		Teachers Questionnaire D19 Do you agree with the following statements?	>75%-100%
		The school's priorities and goals are clear	
		The school's priorities and goals are shared	
	The teachers and the head teacher are clear about the objectives and values of the school, and how they are implemented through teaching and learning	The teachers of this school are committed to implementing school goals and values	Yes/No
		Teachers follow the same set of rules about pupil behaviour	
		There is discussion among teachers on effective teaching methods	
		There is a unanimous view on the behaviour that teachers and pupils should engage each other	

Tab. 2 – Indicators by efficacy factors divided between those present in the RAV and those given by the questionnaires

Factors	*N. components and exemplary items*	*RAV indicators*	*Headmaster N. of questions*	*Teacher N. of questions*	*ATA N. of questions*	*Student N. of questions*	*Parent N. of questions*
Performance orientation and high expectations	4	17	0	2	1	0	0
Educational leadership	7	21	8	15	0	1	1
Consensus and cohesion among staff members	3	14	0	1	0	0	0
Quality of the curriculum/ learning opportunities	3	11	0	1	0	0	0
School climate	12	20	5	9	4	10	2
Evaluative potential	6	14	0	8	1	0	5
Parental involvement	2	4	8	3	0	0	2
Feedback and reinforcement	1	0	0	2	0	2	2
Total	38	101	21	41	6	13	12

The indicators do not manage to cover all factors of effectiveness so, where necessary, some questions that derive directly from the sub-components and exemplary item were collected from Scheerens and added.

The five questionnaires should serve schools to analyse all those aspects that would otherwise remain unexplored and would not allow for a correct, participatory and complete self-analysis.

By linking the available indicators to the factors of effectiveness, it can be seen that the whole part concerning the involvement of stakeholders in the educational process remains "uncovered". This leads to a fundamental element of analysis, which the schools still do not investigate sufficiently. In the set of indicators of the RAV there are no indicators deriving from the direct participation of the stakeholders. The task of constructing further indicators is left to the culture and capacity of the individual schools.

In particular, indicators to investigate parental involvement, feedback and reinforcement of learning are almost completely missing.

Not all effectiveness factors need the same number of questions to be analysed by schools. We can divide them into:

– factors widely covered by indicators for which they do not need significant additional coverage (performance orientation and high expectations; consensus and cohesion among staff members; quality of the curriculum/learning opportunities);
– factors that despite the large number of indicators provided to schools still need to be extensively investigated (educational leadership, school climate, and evaluation potential);
– factors that are characterised by few indicators and therefore need to be investigated with questionnaires (parental involvement and feedback and reinforcement).

Five questionnaires (DS, teachers, ATA staff, students, and parents) were created and used to support self-evaluation in schools.

The questionnaire was chosen for two reasons:

1) no new tasks for the schools; the questionnaire was simply provided/used in place of the school's end-of-year evaluation questionnaires;
2) it is a simple and straightforward tool.

At the end of each school year, the schools administer satisfaction questionnaires, which are set out in the Service Charter. "In order to collect useful elements for the evaluation of the service, a survey is carried out using appropriately calibrated questionnaires, sent to parents, staff, and – especially in secondary schools – students. The questionnaires, which deal with organisational, didactic, and administrative aspects of the service, must provide for the grading of evaluations and the possibility of formulating proposals. In

formulating the questions, indicators provided by the school administration themselves and local authorities may be used".

Subsequently, the feedback from the satisfaction questionnaires is used by the teachers to draw up a report on the school's educational activities; this must then be approved by the board and submitted to the circle or institute council or to the extraordinary commissioner.

Questionnaires are administered at the end of the school year. It is preferred to carry out the administrations in May, with the deadline set for the end of the term. With the beginning of the summer holidays, it would be very difficult to receive feedback even from the staff.

However, the last period of the school year is particularly busy with textbook orders, the verification and review of planning, departmental meetings, the teachers' board meeting, the final verification of the Individualised Educational Plan (PEI), the verification of teaching/discipline progress, the final exams (both State and internal).

The questionnaires are filled in online, generally with a link connected to a google form and in some occasions a QR code that connects to the questionnaire. They do not follow a common format, except in particular cases of schools that have carried out practical experiences of self-evaluation. The questionnaires do not follow a defined structure but are autonomously constructed by the educational institutions and cover the didactic, organisational, relational, and finally educational areas, focusing essentially on the degree of satisfaction with the services. To ensure data standardization, we used the same forms across all schools, unlike previous questionnaires. The forms were carefully designed to encompass all the necessary assessment of the different groups (school headmaster, teacher, ATA staff, primary and secondary school students and parents).

Tab. 3 – Questions of each questionnaire divided by type

Type of questionnaire	*Total questionnaire inquires*
School Headmaster Questionnaire	29
Teacher Questionnaire	33
ATA Staff Questionnaire	14
Primary and secondary school students forms	16
Parents forms	18
Total	110

The questionnaires, based on Scheerens' factors of educational effectiveness, were administered to schools in the province of Cosenza. These schools were chosen for a number of reasons:

– lack of schools that are active in self-evaluation, particularly in the integration of the RAV indicators;
– particularly negative results in standardised tests;
– gap between the results obtained in State examinations (more excellent pupils with a grade of 100 and honours) and the INVALSI results;
– geographical area – vast.

The forms, together with a concise explanation of the project and a link, were sent to 142 schools in Cosenza province. The schools were divided into 4 areas, according to their grades. Only 30 schools decided to participate and administer the questionnaires (20 primary schools and 10 secondary schools).

The results of the questionnaires could be viewed in real-time via Google Moodle. Each participating school received three reports (from the end of May, close to the end of teaching activities) on how many questionnaires had been completed, divided by type (DS, teachers, ATA, students, and parents).

From the sending of the links at the beginning of May until the end of the administration, which was announced on the last day of school (9 June 2022), the schools received periodic reports on how many questionnaires had been filled in by all those involved in their school.

At some schools' request, the deadline for the questionnaires was extended from the last day of school (9 June 2022) to 20 June to allow more time for the administration. Each school received its data and an accompanying file to better "understand" how to read them:

The administered questionnaires were:

1) School Headmaster Questionnaire (23 filled forms each consisting of 29 questions);
2) Teacher Questionnaire (1,342 filled forms each consisting of 33 questions);
3) ATA Staff Questionnaire (174 filled forms consisting of 14 questions each);
4) Primary and secondary school students forms (864 filled forms consisting of 16 questions each);
   4a) High school students forms (1,217 filled forms consisting of 16 questions each);
5) Parents forms (1,720 forms consisting of 1246 filled for the first grade and 474 for the second grade, each with 18 questions).

The schools with the highest percentage of completion of questionnaires were those in which it was possible to have more direct relations. In particular, it was verified that it is not enough to publish the notice of administration and the links to access the completion of the questionnaires on the institutional website page, but the schools must involve all the participants in the process and define the most effective procedures.

At the end of June, each school had its own data available to self-analyse and support the evaluation analysis required by the National Evaluation Service for the three-year reference period 2022-25. A file was sent for reading and understanding the data, linking the questions from the evaluation questionnaires to each effectiveness factor.

3. Data analysis

The schools taking part in the research received a reference outline for data analysis, accompanied at the end of each factor with a short analysis.

3.1. Performance orientation and high expectations

Almost all teachers like their job. In particular, despite that teaching at primary and pre-primary school is demanding, it is precisely at this level of education that we find the highest percentage of teachers who respond positively; in secondary school, on the other hand, we have 8.6% of teachers who like their job "quite a bit": in this category where the training is more specific and focused on set subjects, the teaching profession is considered by some to be a fallback.

The teaching staff believe it is very important for pupils to achieve good results in their studies and also in tests and examinations. However, a small part still prefers his pupils studying without considering the connection between teaching and the acquisition of certified competencies.

3.2. Educational leadership

At least one-third of the headmasters have taken action for the inadequate performance of the teachers, through organisational changes, and only as the last option through disciplinary measures. The school's main priority is the quality of teaching and learning, as well as the well-being of the pupils and a positive school environment: these are not regarded as means but as the purpose of educational effectiveness.

While schools schedule meetings for evaluating, planning, and setting priorities and objectives to be achieved, in one-third of the cases they fail to do the same for the teaching methods and behavioural rules. On the other hand, this is an area that is very much felt by all the components who share

the importance of the rules but that underline a lack of compliance with them.

The staff feel that the headteacher supports them and that his actions serve to improve the school. His positive contribution, not only in coordinating but also in participating in activities, and dispensing the correct information, makes it possible to nurture a relaxed environment as well as shared values.

One aspect that can be improved is being able to share objectives not only with all the staff but also with the students. The percentage of students who feel supported in their school education career drops dramatically from the first grade (91%) to the second grade (75%).

The school head dedicates most of his time to security, administrative, organisational, and management activities and devotes less time to activities related to the educational sphere and involving direct contact with pupils. The relationship with institutions is characterised by the organisation of activities and shared decisions, but only in 43% of cases this is considered fruitful.

3.3. Consensus and cohesion among staff members

Both teachers and ATA staff feel more satisfied with their working hours than with the appreciation they receive for their work. Priorities and objectives are clear but not always shared.

3.4. Quality of the curriculum/learning opportunities

On the curriculum there is the greatest convergence, there is no doubt that the one proposed is suitable for the educational success of pupils. Extra-curricular activities, which absorb considerable resources in Italian schools, are considered useful, but only 61% of the schools considered them very useful.

3.5. School environment

All participants (also pupils) agree on the existence of a behavioural problem, there is a lack of shared rules of behaviour in schools for pupils to follow.

Pupils are satisfied with the relationships they have with their classmates but relatively less satisfied with their behaviour. This gives further food for thought. Together with the respect for rules required by all stakeholders, makes

it necessary for schools to address this issue with the tools available (co-responsibility pact, internal regulations, shared analysis of problems), without wanting to conceal uncomfortable situations for the sake of a quiet life.

The school promotes positive situations through a welcoming environment and positive interpersonal relationships. A factor on which pupils and parents agree is the fact of rewarding virtuous behaviour but penalising erroneous behaviour with the same commitment. Parents are the least satisfied with the behaviour of pupils, safety issues, and actions taken by schools against vandalism, drugs, alcohol, and smoking.

Despite these situations, the climate in schools is still considered good, even if the environment within classes is appreciated more than that of the whole school.

While the staff state that they have positive relations with each other, the students on the other hand are quite critical of the relations they have with the school staff with values differing between school staff, teachers, the headmaster, and secretarial staff. This highlights a detachment between the secretary's office, whose administrative role appears distant from the students' daily lives, and the headmaster who, as already pointed out, has little relationship with the students and devotes little time to educational and teaching activities.

The students' questionnaire shows its effectiveness especially in analysing the differences between the results of the first- and second-grade schools. Students' commitment is not always appreciated and in case of difficulties, learning objectives are not always set. Highly gifted pupils are not always appropriately valued through learning objectives dedicated to them: it is interesting to note that the best results are clearly in first-grade classes where pupils state that they are encouraged, helped in difficulties, and valued with much higher percentages than their older counterparts.

What the schools and the national school system itself have to reflect on, is that the data shows that students' satisfaction with teaching (94% cycle I, 80% cycle II), the support they receive in their school education (91% cycle I, 75% cycle II), and their participation in school decisions (54% cycle I, 38% cycle II) all decrease over the course of their education, along with their performance.

Fewer than half of the students feel involved in school decisions, while the remainder feel involved only partly or not at all. In high schools where students' representativeness is ensured by a whole set of bodies set up for it, the figure is much lower. For the students, the school has a lot of things to improve, especially laboratories, sports activities (and consequently gyms), building maintenance as well as school participation in competitions and events. Even projects for most pupils could be more fruitful and with all the resources spent by schools, further investigation would be appropriate.

3.6. Evaluative potential

There is not always a comparison between the home school and the next grade. This point is further confirmation of the schools' attempt to bind themselves up in their self-regulation: collaboration between educational institutions belonging to different grades should be the practice, aiming at improving not the number of enrolments but the educational offer. In addition, the use of data to compare with similar schools could lead to a process of self-analysis and subsequent improvement, which seems still not understood.

Teachers discuss the evaluation of their pupils both in formal meetings but also informally with their colleagues, but even though they discuss the results of their pupils, 25% do not adjust their educational, learning, or teaching strategies.

Almost three-quarters of the questionnaires were filled in by first-grade parents, further confirming the more active participation by this cohort. In this group, it was noted that the register check takes place at least once a week and, in addition to checking mainly grades, report cards, and marks, the parents of the youngest pupils look also at the homework assigned. The second-grade parents instead look at absences, lateness, and tests results.

Parents manage to communicate with the school without too many problems mainly through the electronic register but also by looking at information on the institutional website or by reading circulars, notices sent home, or by meeting teachers and through messages in chat rooms. Few schools use Facebook as a communication channel.

Although the Three-Year Educational Offer Plan should be the founding document of the school's entire offer, parents are more familiar with the Schools in Clear portal and the Self-evaluation Report.

3.7. Parental involvement

The school proposes activities with the participation of parents, organises four-monthly school-family meetings, and takes active steps to ensure that parents participate in the life of the school. This happens because the school is aware of the parent's influence on pupils' performance, but it is important to take care not to be influenced by the choice of content and organisation.

Although parents are aware that participation in their children's school life is very important and although they feel involved in the school's educational

process, they do not always take part in school-related activities or learn about what concerns their children's education, especially in the second grade.

3.8. Feedback and reinforcement

Class performance is discussed at class council meetings but even informally: this proves that teachers also discuss their pupils outside the bodies set up for this purpose. In case of bad performance, the parents are the first to know, and, especially in secondary school, the students are also directly involved.

The feedback is appreciated by the families who, should they need it, can discuss their child's progress with the school, as well as be contacted in case of difficulties and informed according to the activity plan. If the feedback is appreciated by parents, it is not the same for students who do not discuss their results with teachers. Students essentially receive feedback when teachers return tests despite the formative nature of assessment.

Pupils in both grades also agree that, if they do not understand a topic, the teachers seldom make a personalised pathway – in almost half of the cases the teachers provide the same explanation for all topics.

4. Conclusions

The questionnaires effectively enabled the participation of all the people involved in the process of school self-evaluation and improvement. Moreover, the data highlighted in the questionnaires is interesting at the level of the educational institution because they can give schools elements of reflection in a self-evaluation key also through the comparison of their findings with those of other schools. At a system level, on the other hand, in addition to confirming what the studies highlight, they can provide further policy indications.

School autonomy, which is now well defined from a regulatory point of view, must be understood and supported by all the staff who are asked not only to be prepared didactically and methodologically but also to feel part of an organisation. The aim of the school is not only the imparting of knowledge but also of skills that lead to the pupil's educational success.

The school organisation, however, needs the full participation of all its stakeholders, through the use of a tool such as the self-assessment questionnaire. It is an instrument that can be adapted to its context but that can also provide, through its foundation in the principles of educational effectiveness, useful elements for individual schools but also for the educational system as a whole.

References

Banerjee A.V., Duflo E. (2012), *L'economia dei poveri. Capire la vera natura della povertà per combatterla*, Feltrinelli, Milano.

Castoldi M. (2010), "Valutazione e autovalutazione delle scuole", in M. Spinosi, *Speciale Valutazione. Studenti, scuole, professionalità*, Tecnodid, Napoli.

Demerath P., Louis K.S. (2017), "US contexts of/for educational leadership", in D. Waite, I. Bogotch (*eds.*), *The Wiley international handbook of educational leadership*, Wiley, Hoboken, pp. 453-470.

Edmonds R. (1979), "Effective Schools for the Urban Poor", *Educational Leadership*, 37, pp. 15-24.

Hofman R.H., Dijkstra N.J., Hofman W.H.A. (2009), "School self-evaluation and student achievement", *School Effectiveness and School Improvement*, 20 (1), pp. 47-68.

Levine D.U., Lezotte L.W. (1990), "Unusually Effective Schools: A Review and Analysis of Research and Practice", *School Effectiveness and School Improvement*, 1 (3), pp. 221-224.

Lipsky M. (1980), *Street level bureaucrats*, Russel Sage Foundation, New York.

Marzano R. (2000), *Transforming Classroom Grading. Association for Supervision and Curriculum Development*, Association for Supervision and Curriculum Development, Alexandria.

Marzano R., Marzano J.S. (2003), "The Key to Classroom Management", *Educational Leadership: Journal of the Department of Supervision and Curriculum Development, NEA*, 61 (1), pp. 6-13.

Palumbo M., Pandolfini V. (2019), "School and inequalities: learning by evaluation?", *Sociologia Italiana*, 13, pp. 113-132.

Sammons P., Hillman J. Mortimore P. (1995), *Key characteristics of effective schools: A review of school effectiveness research*, B & MBC Distribution Services, Ringwood.

Sammons P. (1999), *School Effectiveness*, CRC Press, Lisse.

Scheerens J. (2018), *Efficacia e inefficacia educativa. Esame critico della Knowdlege Base*, Springer, Dorchrect.

Scheerens J., Bosker R. (1997), *The Foundations of Educational Effectiveness*, Pergamon Press, Oxford.

Schildkamp K., Visscher A. (2009), "Factors influencing the utilisation of a school self- evaluation instrument", *Studies In Educational Evaluation*, 35, pp. 150-159.

Vanhoof J., De Maeyer S., Van Petegem P. (2011), "Variation in the conduct and the quality of self-evaluations: a multi-level path analysis", *Educational Studies*, 37 (3), pp. 277-287.

3. Analysis of student needs and personalisation of learning: tools and strategies for self-evaluation and instructional design

by Sara Mori, Alessia Rosa, Jessica Niewint*

The personalization of learning pathways is a process that implies the recognition of students' specificities and the design of appropriate teaching interventions. The aim of this paper is to reflect on the possibility of using the tools proposed by the National Assessment System by asking two questions: Do teachers consider the national tests as a possible tool for identifying pupils' potential? In which areas do the RAV rubrics describe, through the guiding questions and indicators, the complexity of the didactic actions implemented by teachers for the personalization of educational pathways?

An exploratory study carried out through an online questionnaire and interviews showed that the participants did not consider the INVALSI tests as a reference tool for identifying students' potential in their daily work, compared to the tests created within the institutes or purchased for this specific purpose. Regarding the RAV evaluation rubrics, in addition to the "Inclusion and differentiation" rubric, descriptors and guiding questions relating to the organizational and methodological dimension of the "Learning environment" area can be useful in describing personalized teaching practices.

La personalizzazione dei percorsi di apprendimento è un processo che implica il riconoscimento delle peculiarità degli studenti e la progettazione di interventi didattici adeguati. Il presente contributo intende riflettere sulla possibilità di utilizzare gli strumenti proposti dal Sistema nazionale di valutazione attraverso due domande: le prove nazionali vengono identificate dai docenti come possibile strumento per l'identificazione delle potenzialità degli studenti? In quali aree le rubriche del RAV descrivono attraverso le

* The contribution was planned and designed jointly by the authors. Specifically, paragraphs 1 and 2 are attributable to Alessia Rosa; paragraphs 3, 4, 5 and 8 to Sara Mori; and paragraphs 6 and 7 to Jessica Niewint.

domande guida e gli indicatori la complessità delle azioni didattiche messe in campo dai docenti per la personalizzazione dei percorsi didattici?

Da un'indagine esplorativa condotta attraverso un questionario online e delle interviste emerge che i partecipanti non considerano le prove INVALSI uno strumento che hanno come riferimento per identificare le potenzialità degli studenti nel loro lavoro quotidiano, rispetto a prove create all'interno degli istituti o acquistate per questo specifico obiettivo. In merito alle rubriche di valutazione del RAV emerge che oltre alla rubrica di valutazione "Inclusione e differenziazione" posssono essere utili per descrivere le pratiche di personalizzazione della didattica anche descrittori e domande guida che fano riferimento alla dimensione organizzativa e metodologica dell'area "Ambiente di apprendimento".

1. Introduction

There is a wide variety of cognitive functions against which the school is called to recognize (Amstrong, 2012). It's important to value the attitudes of each child trying to create a differentiated learning environment to safeguard everyone's right to educational success and to develop their potential. Miliband (2006) considers personalized learning as the solution to adapt learning according to the needs and previous experiences of individuals, to allow everyone to reach their potential (Hsieh and Chen, 2016; Lin, Yeh, Hung and Chang, 2013). Analysis of international literature has shown that personalized learning is a multi-layered construct (Schmid and Petko, 2019) with numerous definitions and various forms of implementation. Personalized education includes what is taught, how it is taught and the pace at which it is taught, thus managing to meet the individual needs, interests, and circumstances of the students. In previous projects such as PQM (Plan of Quality and Merit) the standardized tests were the starting elements for structuring courses of enhancement and recovery of Mathematics and Italian (Mori *et al.*, 2011; Meroni and Abbiati, 2016). INDIRE researchers are continuing to investigate which tools can be useful for the identification of students' needs and the personalization of learning, within a synergic perspective of intervention between INDIRE and INVALSI. For this reason, this study analyses the possibility of using the tools proposed by the National Assessment System to improve personalization processes within the classroom.

The Self-Assessment Report (RAV) describes the area "Inclusion and differentiation" like «the strategies adopted by the school for the promotion of inclusion processes and respect for diversity and the adaptation of teaching and learning to the training needs of each student in classroom work and in

other educational situations» (INVALSI, 2017, p. 29). From this definition, it is possible to explore the concept of differentiation and personalization of teaching to offer the possibility of expanding the proposed guiding questions and descriptors to guide schools' reflection in the improvement process.

2. Personalization, differentiation and individualization

Before proceeding with the definition of the research plan we believe it is appropriate to clearly define some terms that are sometimes confused. We take up the distinction proposed by the US Department of Education in the 2010 Education Technology Plan because it is noticeably clear and straightforward. "Individualization refers to instruction that is paced to the learning needs of different learners. Learning goals are the same for all students, but students can progress through the material at different speeds according to their learning needs. For example, students might take longer to progress through a given topic, skip topics that cover information they already know, or repeat topics they need more help on. Differentiation refers to instructions tailored to the learning preferences of different learners.

Differentiation refers to instruction that is tailored to the learning preferences of different learners. Learning goals are the same for all students, but the method or approach of instruction varies according to the preferences of each student or what research has found works best for students like them. Personalization refers to instruction that is paced to learning needs, tailored to learning preferences, and tailored to the specific interests of different learners. In an environment that is fully personalized, the learning objectives and content as well as the method and pace may all vary (so personalization encompasses differentiation and individualization)"[1]. In summary, "personalizing" learning refers to placing a more organized emphasis on each student or child learning to increase involvement, achievement, and progress. The study also emphasizes that every student has the right to assist that caters to their needs, interests, and skills. Briefly, personalized learning is a strategy developed to support students in identifying and building upon the knowledge they acquire in the classroom in a way that is appropriate for their talent and competencies. Bray and McClaskey (2012) schematized the differences between the three educational principles to elaborate on various phases (Fig. 1).

[1] http://www.ed.gov/technology/draft-netp-2010/individualized-personalized-differentiated-instruction.

Personalization	Differentiation	Individualization
Starts with the learner	Starts with groups of learners	Starts with the need of an individual learner
Connects with interests, passions, and aspirations	Adjusts to learning needs of groups of learners	Accommodates learning needs of the individual
Learners actively participate in the design of their learning	Explicit instruction based upon the learning needs of groups of learners	Explicit instruction based upon the learning needs of an individual learner
Learners have a voice and choice on what they learn	Teachers create or adapt instruction and choose roles for learners based on different needs of learners	Teachers customize lessons and tasks for learners based on individual needs
Different objectives for each learner	Same objectives for groups of learners	Same objectives for learners with specific objectives for individuals who receive one-on-one support
Learner selects appropriate technology and resources to support their learning	Technology and resources are selected to support the learning needs of groups of learners	Technology and resources are selected to support the learning needs of an individual learner
Learners build a network of peers, experts, teachers, and paraprofessionals to guide and support their learning	Learners are reliant on the guidance of teachers to support their learning	Learners are dependent on individual teachers or paraprofessionals to support their learning
Competency-based models where the learner demonstrates mastery	Based on Carnegie unit (seat time) and grade level	Based on Carnegie unit (seat time) and grade level
Assessment AS learning	Assessment FOR learning	Assessment OF learning
Teachers develop capacity to create independent learners who set goals, monitor progress, and reflect on learning and summative assessments based on student mastery	Assessment involves time-based testing and teachers provide feedback to advance learning	Summative assessment is grade-based and involves time-based testing which confirms what learners know and don't know

Fig. 1 – Difference between personalization, differentiation and individualization

According to Leadbeater (2004), the script of a system characterized by personalized learning starts from the premise that learners should be actively engaged in setting their own targets, devising their own learning plans and goals, and choosing from a range of different ways to learn. By making learning the guiding principle of the system, personalization challenges some of the current divide and boundaries that exist – for example between formal and informal learning, between academic and vocational learning and between

different ages and types of learners (Leadbeater, 2004a, p. 16). Numerous studies have also suggested that personalization can improve students' attitudes toward learning (Higgins *et al.*, 2008), perceptions of empowerment, and motivation (Jones *et al.*, 2013; Pintrich, 2003; Järvelä, 2006). This is especially true when students are given opportunities to foster their creativity and curiosity. Personalized learning, according to Leadbeater (2004a, 2005), encourages children from an early age and across all backgrounds in different learning areas, in school contests, and in everyday life. Finally, individualized learning supports knowledge exchange and development through networking and cooperation (Järvelä, 2006), and the ability to choose (Leadbeater, 2004b, 2005). Then, personalized learning is not meant to be a solitaire educational proposal, but, on the contrary, it takes place in a more complicated dynamic learning environment. Given the central role of each student's peculiarities in the customization processes, it is crucial to identify useful tools for identifying these aspects. Through such identification tools, it is possible to define truly customized projects and spillover assessment techniques.

3. RAV and evaluation rubrics on personalization

The School Self-assessment Report, that is RAV (INVALSI, 2017), includes the area of "Inclusion and differentiation" within the processes describing educational and teaching practices, defining it as described as above in the introduction.

The area is divided into two sub-areas: one specifically dedicated to inclusion, and the other focused on remedial support and learning reinforcement. As far as inclusion is concerned, particular emphasis is placed on the strategies used for students with special educational needs and on actions aimed at valuing and managing differences. The school is invited to reflect on how to monitor the objectives defined in the Individualized Educational Plans (PEI)[2] and to verify the achievement of the objectives contained in the Annual Plan for Inclusion (PAI)[3], but also on the reception of foreign pupils and students with disabilities. This area therefore asks the school to reflect on how it monitors and implements the main processes and documents required to facilitate inclusive practices within the school. As far as "remedial and

[2] A tool designed to plan the educational intervention for students with disabilities or special educational needs, in accordance with Italian school regulations.

[3] A programmatic document adopted by Italian schools to promote the educational inclusion of all students, with particular attention to those with special educational needs.

reinforcement" is concerned, the indicators refer specifically to interventions implemented to respond specifically to the needs of pupils with learning difficulties or for students with disciplinary aptitudes.

Level seven (it is the maximum level which means an excellence…) of the evaluation rubric for this area states: «Various parties (curricular teachers, support teachers, tutors, families, local authorities, associations) including the peer group are actively involved in inclusion activities. [...] The differentiation of educational pathways according to the educational needs of individual students is effectively structured at school level; activities aimed at different groups of students reach all potential recipients» (INVALSI, 2017, p. 31).

To complement this area, the one analyzing "Learning environments" deals with «The school ability to create a learning environment for the development of students' competencies. The care of the learning environment concerns both the material and organizational dimension (management of spaces, equipment, times, and schedules), the methodological dimension (dissemination of teaching methodologies), and finally the relational dimension (attention to the development of a positive learning climate and transmission of shared rules of behavior)» (INVALSI, 2017, p. 24).

The indicators of the organizational dimension stimulate reflection on the flexibility in the use of space and time in teaching functions with a particular focus on laboratories, the library, and technological and other resources. The methodological dimension investigates the teaching methodologies used by teachers and how they compare with their adoption. The relational dimension investigates the school climate both in terms of internal relations and with the parental component; special attention is given to the ways in which rules are shared, conflicts are resolved, and problematic behavior is responded to. Level seven of the evaluation rubric for this area states: «Organization of space and time responds optimally to the students' learning needs. Laboratory spaces, technological equipment and the library are used with high frequency by all classes. [...] Relations between students and between students and teachers are positive. Conflicts with students are managed in effective ways» (INVALSI, 2017, p. 27).

The division into different areas is necessary to describe the complexity of the school organization. However, it is necessary to expand the activities under the heading of differentiation activities, as personalization of learning does not only include remedial or reinforcement actions and does not only concern students with learning difficulties.

In fact, we would like to consider under the heading of inclusion and personalization also all the interventions directed at the class group (and not only at individuals); it would also be important to guide the school in a re-

flection on interventions for students who present potentialities and specific aptitudes in certain disciplines and fields. The words of the teachers on these issues can guide a reflection in this sense, through the different evaluation rubrics of the RAV, to promote a broader reading of the phenomena of personalization of teaching.

4. The research questions

This study aims to answer two research questions. The first one: "What tools do teachers use to identify the needs and aptitudes of their students? Are INVALSI tests considered by teachers as a diagnostic tool? The second one: "Which aspects can be considered by teachers as elements of self-evaluation of the school personalization processes?".

The results will be discussed with two objectives: to understand how teachers declare analyzing students' needs and aptitudes; to explore how the indicators and descriptors of the RAV and the standardized tests can contribute to reflecting on the analysis of students' needs and on personalized education.

It is assumed that both the national tests and the RAV could be crucial elements of information to personalize the activities in class. The results may provide us with interesting indications as to the INVALSI test and RAV can help guide teachers' reflection on their own practices and improvement at both the classroom and organizational levels.

5. Methodology

5.1. Participants

A total of 206 teachers participated in the study. As far as gender is concerned, the sample consisted of 197 females (96%) and 9 males (4%). Regarding the school level of work, 11 teachers worked in pre-school, 89 of them teach in primary school, 43 in lower secondary school, and 53 in upper secondary school. Finally, 10 teachers worked in differently classified contexts (e.g., PCAE, Provincial Centres for Adult Education). The average age of the professionals involved was 45.3 years, range: 32-61 years). Schools of the whole national territory were involved in the research, except for three territorial areas (Valle d'Aosta, Molise, and Basilicata); the major part of respondents come from Lombardy (N = 26 teachers) and Sicily (N = 26 teach-

ers). This contribution investigates the feedback of 27 teachers who declared that they use standardised tests to measure students' potential; among them, the research group also expected to find INVALSI tests in primary schools as a standardized tool. Teachers' school grades were kindergarten (1), 11 from primary, 4 from secondary low, and 5 from secondary high. 8 of them were willing to be interviewed online to learn more about the methods of analysing students' needs and what they understood through standardized tests.

5.2. Tools and procedures

In this survey, the design used a mixed-methods approach (Creswell and Plano Clark, 2011) of an explanatory sequential type, characterized by an initial quantitative data collection allowing a starting set of data, which will be deepened using a qualitative survey.

The online questionnaire was disseminated through INDIRE social networks and mailing.

It consists of 20 closed and open questions. The first part is aimed at collecting the teachers' biographical information; the second part aims to investigate what practices teachers use to understand students' characteristics and attitudes; the third part investigates the methodologies and tools that teachers put in place to personalize learning, with a particular focus on the role of innovative technologies in this process.

In order to meet the first objective, it was investigated which instruments teachers use to identify the aptitudes and potential of their students (e.g., entrance tests, competency tests, observations, documentation of previous classes, etc.): those who reported the INVALSI tests or with the help of the indicators proposed by the RAV were asked to be available for an in-depth interview. The interview conducted online was aimed at understanding what tools or moments teachers referred to when stating that they use "standardized tools" for student needs analysis.

To achieve the second objective, the answers to a question of questionnaire aimed at investigating the customization activities carried out in the classroom. By identifying recurrences and groups of actions, it was possible to initiate a reflection on whether the proposed guiding questions of the RAV under the heading "Differentiation and inclusion" are sufficient to understand the categories that emerge.

5.3. Data analysis

The strategies used for identifying the aptitudes of students and to personalize learning were analyzed by a text analysis of the open answers of the survey. The present study utilized QDA Miner Lite, a freely available software program that provides both quantitative and qualitative tools for text analysis, to analyze open-ended responses collected from the questionnaire. The analysis was conducted through text coding, followed by word frequency analysis and an interactive word cloud. Following a grounded theory approach, thematic categories were defined, with each category containing one or more sub-categories (Kuckartz, 2014). Specifically, the open-ended answers were coded for two subject areas, namely strategies for identifying the aptitudes and potential of students, as well as strategies for personalizing teaching and learning approaches. Finally, key categories were extracted based on the frequency and consistency of the observed phenomenon. The answers given by teachers in the interview were analyzed in a qualitative way, searching for evidence in using specific strategies to identify students' potential.

To investigate on the first research question, about what kind of tools do teachers use to identify the needs and aptitudes of their students and if INVALSI or other standardized tests are used as a diagnostic tool the open question the answers to the open question "How do you identify the aptitudes and potential of your students?" were analyzed and followed by a further investigation interviewing some of the participants who included "standardized test" in their answers (section 6.1 and section 6.2).

To investigate about which aspects can be considered as elements of self-evaluation of the school personalization processes the open answers to the question "What strategies do you use to personalize instructional pathways?" were analyzed (section 6.3).

6. Results

6.1. Strategies to identify aptitude and potential

In the context of the tools used to identify the aptitudes and potential of students, the answers to the open-ended question "How do you identify the aptitudes and potential of your students?" were analyzed in a qualitative approach and the 475 keywords found divided in six main categories: observation, entry test, documentation, competency test, standardized test,

and reality task. Moreover, the data was grouped based on the participants' school type, which included kindergarten, primary, lower secondary, upper secondary, as well as other categories that encompassed adult education and secondary education.

The figure "Strategies for identifying aptitude and potential" (Fig. 2) presents data on the percentage of participants who named different methods to identify aptitude and potential. The highest percentage of participants (44%) used observation as a strategy to identify aptitude and potential, while the lowest percentage (2%) used reality tasks.

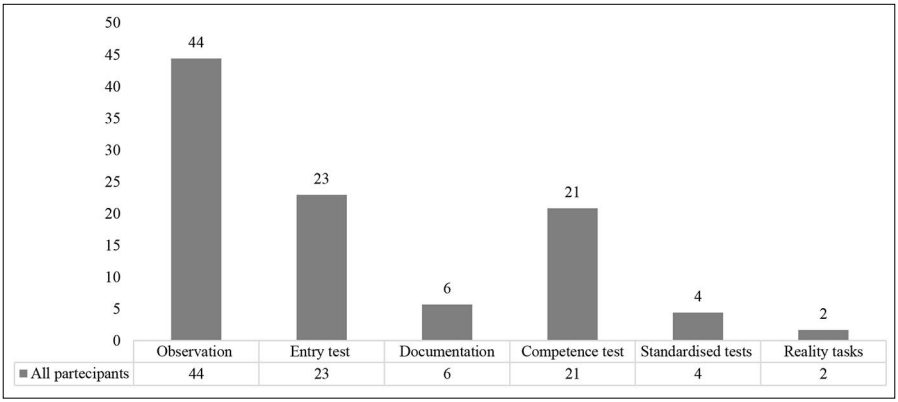


Fig. 2 – Strategies for identifying aptitude and potential

The data regarding the strategies utilized to identify the aptitudes and potential of students was further analyzed and grouped according to the participants' school grade. Observation, entry test and competence test were the strategies named most frequently by the participants (Fig. 3).

The most frequently named category for identifying aptitude and potential in kindergarten students is "observation" with a percentage value of 50%, reflecting the age-appropriate approach of observing children's interactions and behaviors. Competence tests are also utilized (20%), suggesting that basic skill assessments are sometimes applied. Tools like Entry tests and standardized tests are less frequently used (10%), while documentation is not named at all.

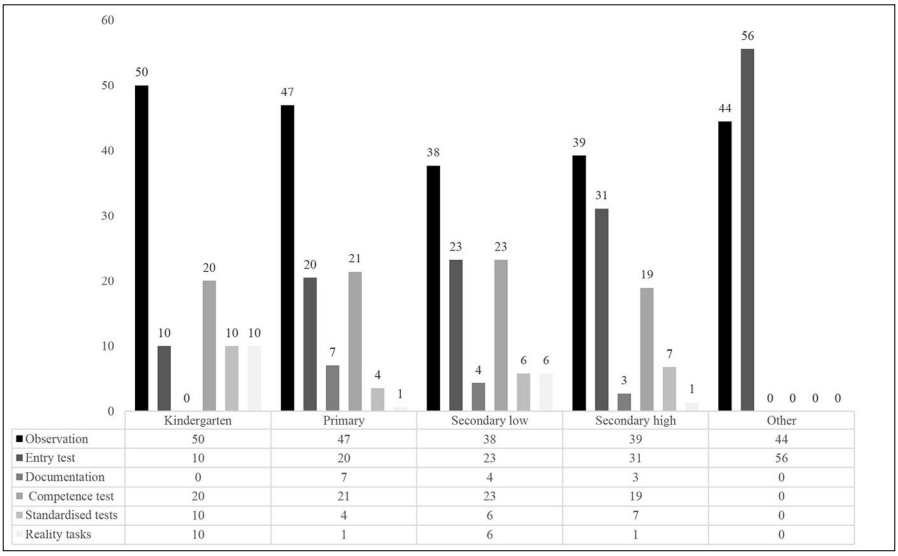
Observation remains the most frequently used strategy (47%), though competence tests (21%) and entry tests (20%) see a noticeable increase. documentation is applied in a moderate number of cases (7%), with standardized tests (3.5%) and reality tasks (0.6%) used less frequently. This distribution suggests a balance between unstructured and structured assessment methods.

In low secondary school observation is still prevalent (38%), though the usage of competence tests and entry tests both increase to approximately 23%, indicating a gradual shift toward more structured assessments. Documentation and standardized tests are less common (4% and 6%, respectively), while reality tasks account for about 5.8%.

While observation remains common (39%), in secondary high entry tests gain importance (31%), likely due to increased academic tracking. competence tests are used at 19%, and standardized tests rise to nearly 7%, reflecting a stronger reliance on standardization in later grades. Reality tasks as a tool to identify aptitude and potential are used only by 1%.

For the participants of the other school type group, entry tests dominate (56%), with observation following at 44%, suggesting a strong emphasis on initial assessment to determine placement or skill levels. No other strategies are used in this category.

Observation is the most common strategy across all school grades, supporting its value in assessing aptitude and student's potential in the classroom. However, as students progress to higher grades, there is a noticeable shift towards more formalized assessments, such as entry and competence tests. The data suggests that while observational methods are foundational, there is an increased emphasis on quantifiable assessments as students advance, potentially to align with educational standards and curricular goals.



	Kindergarten	Primary	Secondary low	Secondary high	Other
Observation	50	47	38	39	44
Entry test	10	20	23	31	56
Documentation	0	7	4	3	0
Competence test	20	21	23	19	0
Standardised tests	10	4	6	7	0
Reality tasks	10	1	6	1	0

Fig. 3 – Frequencies of the six key categories of the "Strategies used to identify" dimension group by type of school

6.2. Experiences of identifying specific needs or aptitudes of students

Twenty-one teachers declared in the survey to use standardized tests to identify specific needs or aptitudes of their students: 1 from kindergarten, 11 from primary school, 4 from secondary low and 5 from secondary high. Eight of these teachers were willing to be interviewed online to learn more about the methods of analyzing students' needs and what they understood by standardized tests. Following are quoted some of the answers given regarding the single experiences of identifying specific needs or aptitudes of students.

The kindergarten teacher reports a case in which: "A child with great difficulty in number-quantity correspondence but was able to read at the age of 3". The teachers apply strategies to identify the single child's potential and difficulties and stretch them into specific reinforcement paths, embedded in the learning paths of the whole group.

One primary teacher stated that the first time she noticed the potential of her student in "Maths activities on fractions" observing that the student executed the activity of "measuring with unconventional measures". Formative testing strategies seemed to also provide a resource for identifying students' needs, as another primary teacher stressed: "With prepared final examinations, I realized that my pupils need more targeted and more consistent formative verification strategies".

A teacher of lower secondary school stated that she assesses students' knowledge by "parallel tests at the end of each term» and often offers «reality tasks with the use of new technologies» in class to apply the new learned knowledge. She provides «diversified and multimedia materials, written test with choice table, final product on authentic task or oral tests on topics chosen by the students with questions on personal reflections (problem solving, critical thinking...)".

One of the secondary high teachers used to share the objectives of teaching, the necessary timeframe and the evaluation grids with her students. She stated also to apply often an assignment of tasks of varying difficulties and to use strategies of reinforcement within the student's reach of competences to create more challenging tasks. Another teacher promotes peer tutoring strategies as a differentiation strategy: "I propose group or pair activities in which I pair weaker pupils with more gifted one".

Overall, it can be noticed that none of the teachers stated to use the assessment results of standardized test, like INVALSI, to investigate the potential of their students and as a source of information for differentiation of the learning content.

6.3. Strategies used to personalize classroom activities

A total of 238 keywords of the answers to the open question "What strategies do you use to customize learning paths?" were grouped into two main categories: "Methodology and approaches", with a total of 191 keywords, and "Tools", with a total of 47 words. Table 1 shows the frequencies in percentage of the keywords for the two main categories.

Tab. 1 – Frequencies in % of keywords in the two main categories for strategies used to customize learning paths

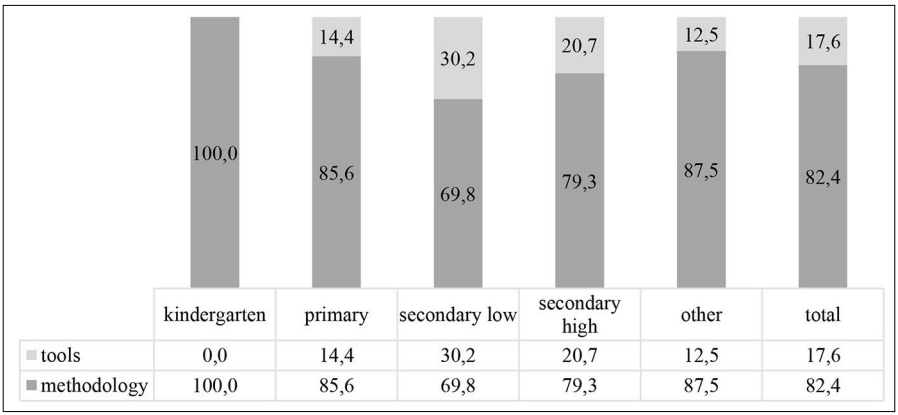
Category	*Keywords*	*Count*	*% for sub category*	*% over all*
Methodology and Approaches				
	Cooperative learning	27	14	11
	Learn to learn	4	2	2
	Diferentiantion	66	35	28
	Laboratory activities	18	9	8
	Problem based learning	7	4	3
	Project based learning	2	1	1
	Group work	31	16	13
	Peer tutoring	26	14	11
	Spaced learning	2	1	1
	Multiple intelligences	3	2	1
	Learning styles	5	3	2
		191	100	80
Tools				
	Compensative tools	4	9	2
	Maps	5	11	2
	Digital tools	20	43	8
	Reality tasks	8	17	3
	Gamification	5	11	2
	Metacognitive tools	5	11	2
Sub categories		47	100	20
Over all		238		100

The "Methodologies and approaches" category, which accounts for 80% of all entries, shows a strong focus on pedagogical strategies. Differentiation is the most commonly used method, accounting for 35% of this category, reflecting a commitment to personalised learning. Other significant methods include group work, cooperative learning and peer tutoring, emphasising the collaborative learning category. In addition, laboratory activities, prob-

lem-based learning and learning styles are also used, although less frequently, to support hands-on and individualised learning.

The category of "tools", which accounts for 20% of the total entries, indicates a minor focus on specific pedagogical tools compared to the named methodologies. Digital tools are the most common, accounting for 43% of this subcategory, reflecting a trend towards digital integration in education. Reality tasks, which focus on real-world applications, follow with 17% and support experiential learning. Other tools such as maps, gamification and metacognitive tools each account for 11%, promoting visual aids, game-based learning and reflective practice. Compensatory tools are the least represented at 9%, suggesting that they are more specialised or complementary within the curriculum.

The data shows a strong emphasis on methods and approaches (80%) over tools (20%), highlighting a focus on structured, inclusive teaching methods. Differentiation is particularly prominent, reflecting a commitment to personalised teaching for diverse classrooms. Digital tools are the most common, supporting an increasing reliance on technology, while collaborative methods such as group work and peer tutoring emphasise the focus on social learning approaches.



	kindergarten	primary	secondary low	secondary high	other	total
tools	0,0	14,4	30,2	20,7	12,5	17,6
methodology	100,0	85,6	69,8	79,3	87,5	82,4

Fig. 4 – Frequencies of key word for the "Strategies used to customize learning paths" grouped by the two main categories tools and methods and school grade

Grouping the keyword frequencies for school grade (Fig. 4) emerges that in kindergarten 100% of the strategies used to personalize learning paths are part of the category "methods". The data reveals two primary categories, "methodology" and "tools," distributed across various educational levels: kindergarten, primary, secondary low, secondary high, and other, with

an overall total summarizing their representation. Methodologies dominate across all levels, with notably high percentages in kindergarten (100%), primary (85.6%), secondary high (79.3%), and other (87.5%), leading to an overall representation of 82.4%. In contrast, tools are much less prevalent, appearing only in primary (14.4%), secondary low (30.2%), secondary high (20.7%), and other (12.5%), and making up just 17.6% of the total.

This distribution suggests a strong preference for methodological approaches across educational levels, particularly in early childhood education, such as kindergarten, where tools are entirely absent. as students progress through higher educational levels, especially in secondary low, the presence of tools gradually increases, indicating a greater integration of practical resources or technology as students advance. This pattern underscores a predominant focus on methodologies over tools in early education, while tools gain relevance in higher levels, likely reflecting the adaptation of learning strategies and technology use according to age and developmental stage.

7. Discussion

For the needs analysis of their students, the most indicated keyword by the participating teachers is “observation” in general. Aptitude and competence tests were mostly used by the participants. For the process of personalization, teachers describe not only approaches such as reinforcement and remedial activities, but also methodologies and tools that refer to cooperative methodologies and the development of broader aptitudes.

However, 4,4 % of the teachers referred to the use of “standardized tests” to identify needs.

The types of tests they refer to must be further investigated qualitatively. When responding to the questions about the activities carried out to personalize teaching, the most frequently used labels were differentiation, group work, and the use of “new technologies in the classroom”.

In the RAV, these aspects are explored more in the “learning environment”: especially in the “organizational dimension” (e.g.: How does the school use the technological equipment available?) and in the “methodological dimension” (e.g: What teaching methods are used by the teachers?). This aspect can lead us to think about how the evaluation rubric for this area can also be enriched by links with other guiding questions from other areas of the RAV.

The RAV and the national tests, which are designed to describe the school level of the organization, could also provide teachers with useful informa-

tion for a class level analysis, particularly for personalizing instruction (Pinto and Gasperat, 2021). They emphasize the importance of interpreting the data from these assessments in a meaningful way for individual students and using it to guide personalized teaching strategies The questionnaire carried out online by the INDIRE researchers on the tools and methods useful for the personalization of pathways does not explicitly refer to the assessment/self-assessment tools provided by the national assessment system (SNV). This also opens the possibility to reflect on how the synergy with other aspects dealt with in the RAV in areas of assessment other than "differentiation" can constitute useful elements for reflection. The terms "personalization" and "differentiation" have different definitions in different studies (Shemshack and Spector, 2020) and it may therefore be interesting to deepen the definition of these processes and the tools used in this area, which is so important for bringing the educational process closer to the unique reality of each student.

8. Conclusion and future developments

The Personalization of learning implies in-depth knowledge of students' learning preferences, skills, and interests.

This knowledge is instrumental in identifying the specific learning objectives and content, the method and pace of work that define the essential elements of the design (De Freitas and Yapp, 2005).

Previous projects conducted in cooperation between INDIRE and INVALSI (i.e.: PQM, that is Quality and Merit Plan) have demonstrated the possibility of considering standardized tests as diagnostic tools for students' potential. Overall, INVALSI tests can provide teachers with valuable information about their students' strengths and weaknesses in specific subjects and this information can be used to adapt teaching methods and lesson plans to better meet the individual needs of each student and address their areas of weakness.

However, in this study, the participants did not mention this possibility. It would be desirable to explore this aspect with further projects on how invalidity tests can contribute to the recognition of students' peculiarities: tests could be a tool to identify students who are struggling and provide additional support where needed, such as tutoring or special education services. INVALSI tests could be used with other tools aimed to explore students' interests and learning preferences to develop personalized learning plans.

With respect to using the RAV to describe the complexity of personalization, the analysis of teachers' words leads us to integrate what emerges in

the "Inclusion and differentiation" rubric with elements from the "Learning environments". Including questions on individualized interventions for educational needs, aspects such as learning environments and teaching methodologies for differentiation to the evaluation rubrics on differentiation and personalization in the "Remedial and reinforcement" sub-area of the RAV could provide an insight of their use in the class activities. This approach could lead to a deeper analysis of useful elements for reflecting on the implementation of personalization and differentiation in learning and teaching pathways. Even if the sample used in this study is not representative and thus cannot be generalized a co-design of free and useful tools for teachers to identify students' needs and aptitudes for the personalization of teaching could be useful in guiding teachers' self-reflection on their teaching practices. Since the elements considered are complex, it is also important to identify strategies capable of detecting peculiarities and competences in the long term and in a vertical perspective.

References

Alam M.M., Mohiuddin K., Das A.K., Islam M.K., Kaonain M.S., Ali M.H. (2018), "A Reduced feature based neural network approach to classify the category of students", *Proceedings of the 2nd International Conference on Innovation in Artificial Intelligence*, Association for Computing Machinery, ICIAI '18, pp. 28-32, https://doi.org/10.1145/3194206.31942.

Armstrong T. (2012), *Neurodiversity in the Classroom: Strength-Based Strategies to Help Students with Special Needs Succeed in School and Life*, ASCD, Alexandria.

Bray B., McClaskey K. (2012), *Personalization vs differentiation vs individualization*, available at: http://education.ky.gov/school/innov/Documents/BB-KM-Personalizedlearningchart-2012.pdf.

Creswell J.W., Plano Clark V.L. (2011), *Designing and Conducting Mixed Methods Research*, SAGE, Los Angeles.

De Freitas S., Yapp C. (*eds.*) (2005), *Personalzing Learning in the 21st Century*, Network Educational Press, Stafford.

Higgins S. *et al.* (2008), *Personalising learning: the learner perspective and their influence on demand*, BECTA, Coventry.

Hsieh C.W., Chen S.Y. (2016), "A cognitive style perspective to handheld devices: customization vs. personalization", *International Review of Research in Open and Distance Learning*, 17 (1), pp. 1-22, retrieved on April 6, 2023 from: https://doi.org/10.19173/irrodl.v17i1.2168.

INVALSI (2017), *Guida al Rapporto di Autovalutazione*, retrieved on April 6, 2023 from: https://www.invalsi.it/snv/index.php?action=ais.

Järvelä S. (2006), "Personalised Learning? New Insights into Fostering Learning Capacity", in OECD, *Personalising Education*, OECD Publishing, Paris.

Jones A., Scanlon E., Gaved M., Blake C., Collins T., Clough G., Kerawalla L., Littleton K., Mulholland P., Petrou M., Twiner A. (2013), *Challenges in personalisation: supporting mobile science inquiry learning across contexts. Research and Practice in Technology Enhanced Learning*, 8 (1) pp. 21-42, retrieved on April 6, 2023 from: http://www.apsce.net/RPTEL/RPTEL2013MarIssue/RPTEL.

Kuckartz U. (2014), *Qualitative text analysis: A guide to methods, practice & using software*, SAGE Publications, New York.

Lin C.F., Yeh Y.C., Hung Y.H., Chang R.I. (2013), "Data mining for providing a personalized learning path in creativity: an application of decision trees", *Computers & Education*, 68, pp. 199-210, retrieved on April 6, 2023 from: https://doi.org/10.1016/j.compedu.2013.05.009.

Leadbeater C. (2004a), *Personalisation through Participation: A New Script for Public Services*, DfES, Nottingham.

Leadbeater C. (2004b), *Learning about Personalisation: How can we put the learner at the heart of the education system?*, DfES, Nottingham.

Leadbeater C. (2005), *The Shape of Things to Come: Personalised Learning Through Collaboration*, DfES, Nottingham.

Miliband D. (2006), "Choice and voice in personalised learning", in OECD (*ed.*), *Schooling for tomorrow: personalising education*, OECD Publishing, Paris, pp. 21-30.

Meroni E. C., Abbiati G. (2016), "How do students react to longer instruction time? Evidence from Italy", *Education Economics*, 24 (6), pp. 592-611, doi: 10.1080/09645292.2015.1122742.

Mori S., Abbiati G., Vidoni D. (2012), "Collaborazione nella scuola per l'efficacia nella scuola. Evidenze dal progetto PQM", *Induzioni*, 44 (1), pp. 13-28.

Pinto C.L.P., Gasperat A.L. (2021), "I dati INVALSI come strumento per sostenere l'innovazione e il miglioramento nei progetti per le aree a rischio", in P. Falzetti (a cura di), *I dati INVALSI come strumento per l'innovazione e il miglioramento scolastico*, FrancoAngeli, Milano, pp. 38-56.

Pintrich P. (2003), "A motivational science perspective on the role of student motivation in teaching and learning contexts", *Journal of Educational Psychology*, 95 (4), pp. 667-686.

Pintrich P. (2003), "A Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts", *Journal of Educational Psychology*, 95, pp. 667-686, doi:10.1037/0022-0663.95.4.667.

Shemshack A., Spector J.M. (2020), "A systematic literature review of personalized learning terms", *Smart Learning Environments*, 7 (33), https://doi.org/10.1186/s40561-020-00140-9.

Schmid R., Petko D. (2003), "Does the use of educational technology in personalized learning environments correlate with self-reported digital skills and beliefs of secondary-school students?", *Computers & Education*, 136 (2019), pp. 75-86, retrieved on April 6, 2023 from: https://doi.org/10.1016/j.compedu.2019.03.006.

4. Gen Z, Personality Traits and Sustainability Awareness: An Econometric Investigation

by Luciano Canova, Giovanna Paladino

Since 2018, the awareness of climate sustainability change has increased significantly, especially among the younger generation. The COVID-19 pandemic and the related shutdown of many economic activities contributed to raising concerns about the conservation of biodiversity, the environment, and personal economic well-being.

In this study, we examine how members of Generation Z deal with issues related to environmental sustainability and personal finance.

By using the technique of the principal component analysis, two synthetic indexes were computed from a set of variables associated with the answers to a questionnaire that investigates the approach to environmental and economic sustainability by a representative sample of 400 Italian youngsters aged between 13 and 18 years. The Green Index is the result of the aggregation of environmental practices while the Money Index represents habits in personal money management. They are used as dependent variables of linear, ordered probit, and bivariate probit regressions to detect how socio-demographic factors and personality characteristics are associated with sustainability awareness.

Our results show the overall importance of character traits – such as curiosity and scrupulousness – in improving the level of awareness and the strong statistical association between attention to money management and a sense of responsibility toward the environment. These finding hint that working on one dimension may produce a positive spillover effect on the other, setting in motion a virtuous circle for policy implementation.

Dal 2018, la consapevolezza della rilevanza dei problemi di sostenibilità e dei cambiamenti climatici è aumentata in modo significativo soprattutto tra le giovani generazioni. La pandemia da COVID-19 e il relativo arresto di molte attività economiche hanno suscitato preoccupazioni sulla conser-

vazione della biodiversità e dell'ambiente, nonchè sullo stato del benessere economico personale.

In questo lavoro studiamo come la Gen Z affronta i temi della sostenibilità e della gestione del denaro. The Green Index is the result of the aggregation of environmental practices while the Money Index represents habits in personal money management. They are used as dependent variables of linear, ordered probit, and bivariate probit regressions to detect how socio-demographic factors and personality characteristics are associated with sustainability awareness.

La tecnica delle componenti principali ha consentito di derivare due indici sintetici basati un set di informazioni multivariate provenienti da un questionario che indaga l'approccio al tema della sostenibilità da parte di un campione rappresentativo di 400 ragazze e ragazze tra i 13 e i 18 anni. Il Green Index rappresenta le pratiche ambientali e il Money Index le abitudini nella gestione del denaro. Essi sono utilizzati come variabili dipendenti di regressioni lineari, probit, ordered probit e bivariate probit per comprendere in che modo fattori sociodemografici e elementi della personalità influenzano il grado di consapevolezza.

I risultati evidenziano l'importanza dei tratti caratteriali – in particolare curiosità e scrupolosità- nell'aumentare il grado di consapevolezza per entrambi gli indici e la forte associazione statistica tra attenzione alla gestione del denaro e senso di responsabilità nei confronti dell'ambiente. Questi risultati supportano l'idea che lavorando su una dimensione si possa produrre un effetto positivo anche sull'altra, mettendo così in moto un circolo virtuoso a sostegno di politiche implementative specifiche.

1. Introduction

In this study, we examine how members of Generation Z deal with issues related to environmental sustainability and personal finance.

Gen Z is an expression that commonly refers to people born between 1996 and 2010. They're the second-youngest generation, between millennials and Generation Alpha (born after 2010) (Michael Dimock for Pew Research Center, 2019).

On a broad scale, environmental sustainability can be defined as the responsible management and conservation of natural resources to meet current human needs without compromising the ability of future generations to meet theirs (Brundtland, 1987). In the context of a social science paper focused on Gen Z's attitudes toward natural resources, environmental sustainability can

be explored through the lens of how this generation perceives the long-term impact of human activity on ecosystems, their support for sustainable practices, and their behavioral shifts towards reducing their ecological footprint.

With regards to personal finance (Remund, 2010), for an economic paper investigating Gen Z's attitudes, the concept can be explored by examining how this generation perceives the use of money, their engagement with saving and investing (often through modern fintech platforms), and their preference for financial education. Personal finance for Gen Z also encompasses the tension between short-term spending on experiences and material goods, and long-term financial planning.

The importance of education in this context is demonstrated by the fact that sustainability is increasingly at the center of policymakers' agenda, as demonstrated by the Sustainable Development Goals (SDG) initiative – approved in 2015 by the United Nations (UN) and defined as a framework of 17 objectives in a roadmap toward achieving a more sustainable economy in 2030 – and by the centrality of Environmental, Social and Governance (ESG) issues in business strategies. The issue of sustainability has become the subject of investment programs, which have significant monetary impacts, as shown by the European Union's commitment to allocate 30% of the 1,800 billion allocated for the 2021-2027 budget for the transition toward a decarbonized economy.

The 360° vision of sustainability also includes specific attention to economic/financial education: humankind and the environment have no sustainable interaction without careful and rational management of economic resources both at personal and community levels. Unfortunately, less than 40% of individuals in Italy can be defined as financially literate, with significant inequalities in terms of gender, educational qualification, and at least partially, geographical distribution. Financial education is also very low among students of all grade levels. According to the latest OECD PISA survey[1], the percentage of Italian students that can solve the most complex tasks (top 5-level performer) is less than half of the OECD average (4.5% vs 10.5%), while about one out of five students lack the minimum skills necessary to make responsible and well-informed personal financial decisions.

An analysis by the Bank of Italy in 2020 (D'Alessio *et al.*, 2020) confirmed the deficiency of financial skills of Italians[2]. The survey uses the

[1] *PISA 2022 Results* (Volume IV), *How Financially Smart Are Students?* (https://www.oecd.org/en/publications/pisa-2022-results-volume-iv_5a849c2a-en.html).

[2] The survey was conducted among a representative sample of 2,000 adults aged between 18 and 79.

OECD methodology, which derives an overall indicator of skills from the scores calculated for three subdimensions: knowledge, behavior, and attitudes. The study, through an econometric analysis, shows only a small improvement in the knowledge component, while behaviors and attitudes are slightly worse off than in previous surveys.

Our research intends to investigate a segment of the population that is not included in surveys aimed at understanding the sensitivity to environmental issues and the link the latter has with money management at the personal level. Our purpose is to identify any spillover between the two dimensions, which could be useful from the perspective of public policy design and implementation. Therefore, we focus on the attitude component and potential effectiveness of the educational interventions that exploit the overlap between the environmental and economic dimensions (the latter meant personal money management).

The question, though of great relevance, has not yet been the subject of careful analysis in the literature. Only a few contributions have investigated the attitudes and behaviors of Generation Z in terms of money management and sustainability. For example, Li and Leonas (2022) present the results of an analysis of the purchase of swimwear by a sample of 257 young women. The authors found that price is the key factor but is accompanied by elements such as the materials used in the production of the apparel and information concerning the sustainability of the production process. Another study (Bollani *et al.*, 2019) based on data from 267 university students of the millennial generation focuses on the relationship with food, highlighting how the information and actions aimed at reducing food waste and waste generation are increasingly important for young people.

Other studies have delved into aspects more related to lifestyles and behaviors. An analysis of 362 young American consumers (McCoy *et al.*, 2021) reveals how much the consumption pattern inspired by the so-called sharing economy is gaining ground compared with that based on ownership of tangible assets or electronic devices. The study in question highlights how, among the reasons that explain the phenomenon, sustainability is perceived more and more as a "status". By using American data, Cho *et al.* (2018) study the trend of reduced participation in voluntary activities and focus on a sample of 360 Generation Z individuals to investigate the motives behind this trend, highlighting sustainability as a strong motivational lever.

Beyond surveys or questionnaires that contain specific sections dedicated to environmental sensitivity, such as the European Social Survey of 2018[3],

[3] Here is the link to the European Social Survey website: https://www.europeansocialsurvey.org.

which has a thematic focus on global warming, not many studies have focused on the younger population so far and examined the relationship between environmental sustainability and attitudes in the use of personal money.

Some literature reviews have focused on the role of sustainability education starting from childhood. For example, Breßler *et al.* (2017) established a taxonomy of the characteristics and learning objectives of Education to sustainable development paths, more importantly identifying a framework that can guide research, including empirical studies, to examine their impacts.

In a meta-analysis, Somerville and Williams (2015) highlight how the number of publications dedicated to the topic of sustainability education has increased significantly over the last few years by identifying three lines along which the research has been developing: connection with nature, children's rights, and theoretical frameworks.

A review by Samuelsson and Park (2017) details the results of an analysis of the relationship between sustainability and childhood by focusing on objective 4 of the SDGs, which guarantees access to quality education services for all. The article addresses the importance of introducing the theme of sustainable education from the primary school and focused on the quality of school services as an intrinsic element of sustainability if viewed from the perspective of lifelong education.

An interesting study is that of Bamberg and Moser (2007), who conducted a meta-analysis of psychological articles, published over the last few decades. It reports the main results regarding the characteristics/behavioral components that correlate with environmental sensitivity. This is an article of interest for our research, as the following analysis delves into the dimension of character traits.

Another relevant factor is the role played by the family in kids' education. Our behavioral indexes catch personality traits in which the family has a part for personal examples and transmitted values. A recent retrospective analysis (Le Baron *et al.*, 2020) – conducted on a sample of 437 adults from the Flourishing Families Project – finds that overt financial education from parents during childhood is associated with a greater frequency of healthy financial management behaviors in emerging adulthood.

On parental socialization, a paper by Olszewski-Kubilius (2018) points out that family has relevant effects on childhood development by impacting attitudes, beliefs, opportunities, habits, and personality traits. The paper discusses why different outcomes and patterns occur with a focus on the psychological aspects of family functioning as it affects the talent development of kids. With specific reference to environmental sensitivity, Thomas *et al.* (2018) indirectly enlighten the complex role of the family by investigating the impact of having children on the development of environmental aware-

ness. Willingness to engage in sustainable actions may be limited by the psychological distance of climate change. In this study, the authors test the legacy hypothesis. Using the Understanding society dataset, a longitudinal survey representative of the UK population (N = 18,176), the researchers assess how having children may change environmental attitudes and behavior. We contribute to this literature by focusing on Gen Z and the possibility of growing a mature awareness of environmental and economic sustainability.

Another interesting study that investigates the relationship between environmental sensitivity and other specific skills is that of List *et al.* (2020), which use data from the OECD PISA survey and compare the linguistic, mathematical, and scientific skills of 15-year-old students at an international level. The research highlights the correlation between scientific skills and the development of awareness of sustainability importance.

Finally, the only study we found that considers environmental and economic sustainability for young adults between 23 and 26 years old, is that of White *et al.* (2018). Based on a series of self-reported intentions and answers to some questions about the knowledge of financial concepts, saving behaviors, and the perception of trust and self-control the authors find a positive association between attitudes to environmental sustainability and healthy money management practices.

Our analysis focuses on two research hypothesis:

– 1: Gen Z members' awareness of environmental sustainability and personal finance management is affected by socio-cultural variables, family, school, and personal characteristics;
– 2: environmental sustainability and personal finance management are strongly statistically associated in this cohort.

This chapter is structured as follows: section 2 describes the data and the main sociodemographic characteristics of the sample; section 3 illustrates the empirical methodology and the identification strategy of the econometric model; section 4 presents the results, and section 5 concludes.

2. Research tool and data description

An online survey was conducted to collect sociodemographic data and information on a sample of 400 Italian adolescents aged between 13 and 18 years[4]. To our knowledge, it is the first attempt to investigate directly pref-

[4] In the field, the investigation was conducted by Episteme. Thirteen-year-olds completed the questionnaire in the presence/with the help of their parents under the provisions of the

erences and attitudes of this specific cohort and we believe the originality of the dataset contributes to a, still limited, literature in social sciences. The questionnaire is divided into five sections and includes 10 questions on sociodemographic information, 6 questions on environmental sustainability, 17 questions on personality traits (related to emotional stability, openness to experience, agreeableness, conscientiousness, social trust, sociability, happiness), 6 questions on money management, and 3 questions on parents as role models and the use of social media.The degree of coherence among the answers highlights the ability of the youngsters to provide sensible and reasonable feedback and gives reassurance on the degree of reliability of the survey[5].

The questionnairewas distributed between February 5 and 17, 2021 and administered with a CAWI procedure. The sample was selected to be representative of the population of young Italians at many levels: age, gender, and geographical residence.

The socio-demographic information provides some indications of family background and is summarized in Tables 2 and 3. From these, we can deduce that the sample shows a slight prevalence of males over females and – given the age range- a majority of high school students (80.6% of the sample), with a preponderance of technical institutes rather than lyceums (71.4% vs 28.6%). The age groups are equally distributed. The geographical distribution indicates a greater number of adolescents interviewed in the South and the Islands than in other macro-areas, in line with a relatively greater presence of youths in these regions[6]. From the point of view of the family cultural background, the presence of at least a parent with a BA degree (19%) is consistent with national data and decreases to 14% in the case of both parents with a BA degree. The most frequent qualification for both parents is a high school diploma. Approximately 31% of the students interviewed have between 25 and 100 books at home. This percentage gradually decreases to 10% for ranges between 201 and 500 and to 5% for more than 500 volumes.

Among boys and girls, a split in "preferences" with respect to school track chosen is apparent between scientific subjects (55% of males versus 32% of females) and humanities (43% of females versus 20% of males). This is a likely indication of the cultural conditioning of social origin, which has its roots in the family and is not questioned even by schools.

law enforced. The sample was representative of the Italian population, stratified by age and geographical residence.

[5] The sample size is equivalent to those in other published studies on Generation Z.

[6] See Demo.istat.it – Popolazione residente per sesso, età e stato civile al 1° gennaio 2021.

In addition to demographic information, the survey consists of four sections. The first one aims at identifying the personality traits prevalent among youngsters and is inspired by the questions included in the European Social Survey[7]. It is composed of a collection of items that aim to measure: emotional stability, open-mindedness, conscientiousness, social confidence, happiness, and sociability. These characteristics may be related, directly or indirectly, to the well-known Five-Factor Model (Matthews and Whiteman (2003), among others). They are openness, conscientiousness, extroversion, agreeableness, and neuroticism. Despite this model being well-established in psychological literature, there are a few limitations. These are associated with its inability to address core constructs of personality functioning beyond the level of traits, limitations concerning the prediction of specific behavior and the adequate description of persons' lives, and with failure to provide compelling causal explanations for human behavior (McAdams, 1992).

Thus, we prefer to test 17 personality features in our questionnaire and leave the task of regrouping them to the cluster analysis as described below.

Table 4[8] presents the grouping of personality traits obtained through a principal component analysis[9] that reveals three main character groups based on specific answers exemplified by scores from 0 to 10. The "restless" are the angriest children, who show lower self-esteem and find it difficult to contain anger and respond to evil rationally. The "curious" are the most open, are eager to learn, interested in the stories and opinions of others – even when they do not share them – and are very sociable. Finally, the "conscientious"

[7] European Social Survey (European Research Infrastructure Consortium), Round 11 Survey Specification for ESS ERIC Member, Observer and Guest Countries, https://europeansocialsurvey.org/sites/default/files/2023-06/ESS11_survey_specification.pdf.

[8] Table 4 does not report two questions included in the section on personality traits because they are not relevant to the cluster analysis that divided children into three main groups. Specifically, the questions are: "How often do you meet your friends outside of school?" (possible answers go from daily to less than once or twice a week) and "How many people do you talk to about your most intimate and private things?" (answers go from none to more than 7). The answers show a distinct lack of sociability, with a third (34%) of the sample not dating friends outside their school even once a week, more than a third (36%) of the sample having at most two close friends, and 37% having no friends or with only one close friend. These responses seem to be affected by the confinement imposed by the pandemic situation associated with the spread of coronavirus disease.

[9] The cluster analysis is a common way to group data in the field of sociological/ psychological data analysis. It allows to start from a large number of characteristics and detect statistical similarities to form groups. The names were chosen in Italian and then translated in English by a native speaker. The word restless was preferred to angry as probably more suitable to describe the mental condition of youngsters.

Table 4 shows the logical association of the scores with the labels.

are precise, scrupulous, and eager to learn too; they show a good deal of trust in others because they believe that others behave like them.

At the level of distribution among the character groups, approximately 50% of the sample falls into the "restless" group. Males are more represented among the "curious"; and females, among the "conscientious." High school students appear more frequently among the "restless" and "curious" (as might be expected at their age), whereas middle school students are overrepresented among the "conscientious". Openness and restlessness are growing functions with the age of the respondents.

The second section of the questionnaire contains four questions aimed at capturing knowledge and sensitivity to environmental issues. It includes questions on knowledge of the problem of global warming, personal commitment to the fight against waste and global warming, value judgments about the responsibilities of adults, and the effectiveness of the role of governments in addressing environmental sustainability. From Table 5, section A, approximately 80% of the sample attributes values higher than 6 (on a scale from 0 to 10) to the possibility of making a difference to reduce global warming and waste. As many as 64% of the sample believes that the responsibility of adults in the current situation of environmental degradation is high, assigning a score between 8 and 10. The opinion on the possibility that governments can act to encourage a change of habits is more evenly distributed, with 70% of the sample assigning it a value higher than 6.

The third section of the questionnaire contains questions on practices associated with the use of money. Six questions (Table 5, section B) assess access to money to be managed independently and the habits associated with the use of the available sums: from planning to the ability to renounce what is superfluous and to the generosity of spirit. The availability of periodic money (i.e., allowances) concerns only a quarter of the sample, with a higher prevalence of males than females. If money is available, the students declare that they pay attention to its use and that they save by thinking about the realization of a project (77%). Furthermore, approximately 77% declares to know how to keep unnecessary expenses under control (giving themselves a grade higher than 6, with 47% of them being convinced that they deserve a grade between 8 and 10) and 63% declares to be particularly satisfied when, after having set aside some money, they manage to buy what they want. In the sample, a certain degree of individualism prevails, which leads to use money above all for personal happiness, even if the girls are slightly more altruistic and attribute an average value of 5 for themselves rather than 6, which is the males' average score.

The fourth section of the survey contains information on the parenting model and the use of digital media such as social media and video games.

It allows us to understand the ability of parents to be role models for their children and the extent to which the use of social media and video games can influence the decisions of children regarding issues such as the use of scarce resources. Parents remain a point of reference for children. Of the young people, 82% answered "a lot" or "enough" to the question of how much of a role model in the management of environmental resources their parents are, while the percentage increases to 91% when the question is related to the management of economic resources. At the character level, the conscientious feel closest to the parental model (89.7% and 96.6%). Regarding the use of social media, 91% of girls spend time on social media every day versus 79% of males and 80% of boys (against 31% of the s) are more attracted to video games and play them daily10. It is important to recall that the use of social platforms is important for the youngest to access the world and collect information, as shown in a recent survey by the PEW Research Center, which refers to the case of the United States (Shearer and Mitchell, 2021).

3. Methodology

To assess the degree of awareness of the use of scarce resources, two synthetic indicators were constructed, which we use as a proxy for environmental awareness (Green Index) and economic awareness (Money Index), in the form of weighted averages of the answers to the questions in the second and third sections of the questionnaire (presented briefly in Table 5). The weight values were obtained using the principal components analysis (PCA). The use of multivariate statistical techniques for the construction of synthetic indicators is widespread in the literature (Vyas and Kumaranayake, 2006; Filmer and Pritchett, 2001). Of course, the choice of questions to be aggregated lends itself to different interpretations. However, the flexibility of this methodology in the case of items with categorical variables for answers to subjective questions makes the analysis particularly useful and has hundreds of applications in the social science literature (Poirier, Grepin and Grignon, 2020).

The Green Index indicator is based on retaining the first main component (i.e., vectors 0.27, 0.28, 0.22, and 0.23) of the principal component analysis which is a proxy of the answers to questions (1), (2), (3), and (4), related to environmental sustainability and reported in section A of Table 5. The four

[10] Data on the use of social media and video games are available upon request https://www.museodelrisparmio.it/wp-content/uploads/2021/04/R.21.101-Museo-del-Risparmio_Report_REV05-002.pdf.

questions relate only to environmental awareness. For the Money Index, given the nature of the questions, the synthetic index is constructed in two steps. The answers to questions (8), (9), and (10), reported in Table 5 and related to the availability of personal money, are grouped with the arithmetic mean[11]. This value is then used to estimate the principal components of a vector, which include the answers to questions (5), (6), and (7) also related to money decisions. The synthetic value was obtained using the coefficients of the first principal component (0.21, 0.33, 0.11, and 0.35) to weigh the vector formed by the average value of the answers to questions (8), (9), and (10) and the single answers to questions (5), (6), and (7).

The distributions of the two synthetic indexes are shown in Table 6[12]. The highest values of both indexes are achieved by the so-called "conscientious" children; and the lowest values, by the "restless". These differences are statistically significant according to the adjusted Wald test. On the contrary, the differences between males and females based on the same test are not statistically relevant. The "conscientious" type is associated with the highest average values of the two indexes, followed by the "curious". The lowest average values (albeit higher than 6) are associated with the "restless".

Table 7 shows an interesting fact: 48.5% of the youngsters who fall into the fourth quartile of the economic awareness index also belong to the top quartile of the environmental awareness index. This evidence also characterizes the lowest quartile and shows a close relationship between the two types of awareness at both extremes. The average correlation between the indexes is 41%.

Money Index and Green Index represent the two proxy dependent variables of the degree of awareness of the use of scarce environmental and economic resources, which can be considered the latent variable. H1 investigates how sociocultural status, family, school, and personal characteristics affect the degree of awareness.

We started by first estimating linear sample regressions for each synthetic indicator. Four models, from the simplest to the most complex, were computed. Many coded answers in the questionnaire, together with some fixed regionally based variables, were used as independent variables in different specifications (see Tab. A1, "Variables' list").

[11] The PCA on the first answers (8), (9), and (10) gives results equivalent to the arithmetic mean.

[12] In our case given the relevance of the topics in kids' daily life and use of a scale 0-10 ("sufficient" score is 6) it may well be that this structure has eventually pushed the distribution to the right, thus creating the asymmetry.

Linear equations can be briefly represented as follows:

$$Y_{1i} = \alpha_{1i} X_{i1+} \gamma_1 Y_{2i} + \varepsilon_{1i} \quad (1)$$

$$Y_{2i} = \alpha_{2i} X_{i2+} \gamma_2 Y_{1i} + \varepsilon_{2i} \quad (2)$$

where $Y_{1i} Y_{2i}$ are the latent variables of economic and environmental awareness approximated to the synthetic indicator Money Index and Green Index, X is the explanatory variable, α is the vector of the coefficients, and finally, ε is the error term assumed to be distributed as a normal standard ε N (0,1), with n = 1,2.

The next step was to estimate the probability of falling into the different quartiles of the indexes through two distinct ordered probits. In this case, each dependent variable Y, where n = 1,2, can take on four values from 1 to 4, as coded below:

$$\begin{cases} 1se 0 \leq Y_n \leq \mu_1 \\ 2se \mu_1 < Y_n \leq \mu_2 \\ 3se \mu_2 < Y_n \leq \mu_3 \\ 4se \mu_3 < Y_n \leq \mu_4, \end{cases} \quad (3)$$

where $\mu_1, \mu_2, \mu_3, \mu_4$ are the threshold values identified respectively as the first, second, third, and fourth quartiles of the distribution of Y.

It is therefore possible to estimate the values of the coefficients α in the three cutoff terms k_n and, consequently, the probability that Y assumes the values 1, 2, 3, and 4 using the standard formulas of the ordered probit.

We are aware of endogeneity concerns regarding the correlation between the dependent variables and the residuals, which makes it impossible to isolate a causal relationship between the explanatory variables and the dependent variables. Nonetheless, since the analysis is functional to the following identification strategy and we base policy suggestion on the statistical association between environmental sensitivity and money management, we do not proceed further with the 2-step procedure to correct for the bias.

The final step was to focus attention on the top performers and consider that a correlation may exist between the error terms, given the importance of what we defined as cross-awareness, the composition between economic and environmental awareness. For this reason, we decided to adopt a bivariate probit estimate (biprobit) that allows simultaneous modeling of two dependent variables related to each other. In this case, cross-awareness was not included in the exogenous list.

The bivariate probit model follows Greene's (2017) formula:

$$Y_{1i} = \alpha_{1i} X_{i1} + \varepsilon_{1i},\ Y_{1i} = 1 se Y_{1i} > \mu 13,\ 0 \text{ otherwise} \quad (4)$$

$$Y_{2i} = \alpha_{2i} X_{i2} + \varepsilon_{2i},\ Y_{2i} = 1 se Y_{2i} > \mu 23,\ 0 \text{ otherwise} \quad (5)$$

where the dependent variables assume a value of 1 if the value of the indicator composed of economic and environmental awareness falls in the fourth quartile and the error terms are expressed by the following formula:

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \rho \\ \rho 1 \end{pmatrix} \right] \quad (6)$$

where ρ is coefficient of the cross-equation correlation of errors. The econometric estimates were obtained with STATA, which allows the estimation of a two-equation model in the seemingly unrelated probit version using the maximum likelihood method. In this case, the margins are to be interpreted as the impact of each independent variable on both dependent variables or pseudo-elasticity calculated on the average values.

4. Empirical results and discussion

Table 8 in sections A and B reproduces the estimate of four explanatory models (Eqs. 1 and 2) for the two synthetic indexes, respectively Money Index and Green Index. The models use a similar set of explanatory variables for the two indexes.

Model 1 (M1) considers only the impact of demographic variables, the sociocultural context, the type and (a proxy of) quality of the school, and the use of social media by children, making it a useful tool for collecting information from sources outside the family and school itself. The model shows a greater explanatory power of the variance for the Money Index (R^2 = 10%), where the SES variables[13], SOCIAL, ISTEC, and QSCHOOL, are statistically significant and of the expected sign, compared with the

[13] The SES variable (socio-economic status) is constructed as a weighted average using the weights of the first component (0.6216, 0.6173, and −0.4783) of the following responses on maternal education level (1 = undergraduate or postgraduate, …, 6 = no formal education), paternal education level (1 = graduate or post-university, …, 6 = no formal education), and the number of books available at home (1 = 0–10, …, 6 = more than 500). In line with the construction of the variable, low values indicate a high sociocultural context.

Green Index (R^2 = 6%), where only the SES variable is significant and of the expected sign.

M2 and M2.A incorporate, among the explanatory variables, the character traits grouped into the three clusters described in paragraph 2. For both indexes, the Personality variable is significant and can increase the variance explained by the model by more than 4 and by approximately 9 percentage points.

By opening the clusters and considering some specific character traits, the M3 model highlights the key role of aspects such as Curiosity, Scrupulousness, and Confidence for both awareness indexes. The increases in R^2 range from 14 to 18 percentage points, reaching 28% and 34%, respectively. The Relationship With Others or the ability to get along with even those who do not have the same ideas is significant only in the case of the Money Index.

The M4 model adds to the explanatory variables both the parental model represented by the Role Model variable (Eco/Green) and the synthetic index of cross-awareness to account for the evidence provided in Table 7. From the variance inflation factor and link tests, the introduction of this variable as an explanatory variable does not involve significant distortions but increases the explanatory power of the model by approximately 4% in the case of the Money Index and by approximately 3% in the case of the Green Index. This confirms the evidence that attention to environmental causes and money management are linked, which is our main empirical scope[14]. If the models are estimated by distinguishing by gender, see model 4.A and 4.B the explanatory power increases, highlighting differences in the significance of some single variable. Taken together, these differences do not affect the overall results. If anything, they offer interesting nuances. However, it should be remembered that the SEX variable is never significant in regressions, highlighting substantial uniformity between genders in terms of economic and environmental awareness.

Overall, while economic awareness (Money Index) is influenced by external contextual factors connected to family, school, use of social media, and personality, the structure of statistical relationships that affect environmental awareness is quite different (Green Index). In this case, the influence of context variables is residual, and that of the sociocultural condition (SES) of the family is not stable and disappears in the more complex specification (Table 8, Section B M3 and M4), where the use of social media appears to play a significant role as a probable source of information more for males

[14] Excluding the cross-awareness index, the results of the regression did not change from a qualitative point of view. However, the percentage of explained variance was lower.

than for females (see M4.B). Also relevant is the age structure of the opposite sign for girls and boys. The variable relating to the air quality of the region where the respondent lives (QUALAIR) does not appear to have any relevance. All the regressions reported in Table 8 show a correct specification and a substantial absence of collinearity among the regressors.

The ordinal structure of the proxy indexes of awareness offers the possibility of analyzing the impact of exogenous variables on the probability of belonging to the different quartiles and therefore on the level/degree of awareness. Table 9 presents the ordered probit estimates for the Money Index and Green Index variables transformed according to formula (3) into ordinal variables (Money Quartiles and Green Quartiles). They allow us to highlight clearer similarities and differences between the explanatory factors of economic and environmental awareness. The incidence of context variables on economic awareness is higher than that on environmental awareness. Some personality traits such as *curiosity*, *scrupulousness*, and *confidence* exert a significant impact. The role of parents stands (role model) on the value of both economic and environmental awareness, in line with the response rates reported in paragraph 2[15], and the relevance of cross-awareness is beyond doubt. We emphasize that the structure of significance remains unaffected if cross-awareness is not included in the regression.

Based on the results of the analysis of the marginal effects, as reported in Table 10, we can state that the probability of falling into the four quartiles is a function of the exogenous variables identified as relevant. For example, a family's good sociocultural status increases the probability of falling into the fourth quartile by 3% (i.e., in the class where the awareness is highest). Symmetrically, a low social status increases the probability of being in the first quartile by 3% (among those with the lowest degree of awareness). School quality also increases the likelihood of being in the top quartile by 6%. Among the personality traits, *curiosity* has the highest impact (3.2% for environmental awareness and 2.5% for economic awareness), followed by *scrupulousness* (2.4% and 1.6%, respectively). The difference in the marginal effect of *confidence* is similarly measured (2.6% and 1.6%, respectively). The level of self-esteem also affects economic awareness in probabilistic terms; that is, higher self-esteem increases the probability of being among

[15] If we replace the variable with a dummy [0,1] in correspondence with the answer "a lot" in the case of economic awareness, the positive role of parents will stand out clearly, with a significance of 5% in the linear regression and 10% in forms probit with a significant marginal effect of 5 percentage points. It was not relevant in the linear regression for green awareness but became relevant in the oprobit and biprobit estimates, where the coefficients must be read as a variation from the baseline value corresponding to 1 = "a lot".

the most aware by 1.1%. Cross-awareness is relevant in both specifications, with a symmetrical impact in terms of probability of approximately 7%.

Finally, to analyze further the impact of the exogenous variables on the last quartile (i,e. on the highest level of awareness), the Green Index and the Money Index variables were estimated using a bivariate probit, represented by equation (4). The results are reported in Table 11, where the dependent variables assume values [1, 0], depending on whether the value of the synthetic indexes falls into the highest quartile of the distribution. They are called Money Top and Green Top. The estimation of the coefficient of equation (5) confirms the advisability of jointly estimating the two regressions. It takes on a value of 0.40 and is highly significant. The significance of the variables was largely confirmed for both specifications.

Finally, Table 12 presents a joint estimate of the marginal impact of the explanatory variables when a bivariate probit is implemented. A non-secondary role of age is evident in the 14- and 18-year-olds being more sensitive to both issues (with a marginal effect of approximately 8 percentage points). The impact of the character aspects, such as *curiosity*, *scrupulousness*, *confidence*, and *relationship with others*, brings about a 2% increase in the probability of being among the top performers.

5. Conclusions

This study examines the statistical relationship between environmental awareness and the attitude of Generation Z toward the use of money in Italy through a representative sample of the population composed of 400 youngsters, which is an original dataset and a relevant contribution to a literature that lacks many direct investigations among this specific segment of the population.

Even within the context in which sustainability is an increasingly central issue in the agenda of both the policymakers and the corporate world, we did not find previous studies that investigate the relationship between these two dimensions. We believe that this research is promising because of the possible spillover effects that can translate into specific policy orientations.

The empirical analysis was conducted starting from the construction of two synthetic indicators using the principal components analysis. Money Index and Green Index capture awareness about the use of money and the exploitation of environmental resources, respectively.

Both indicators have been used in different econometric specifications as dependent variables to explain which factors correlate with them and to determine whether the two variables have some elements in common.

A simple correlation between the Green Index and Money Index – equal to 0.41 and significant at 1% – shows a clear statistical association between the two dimensions, later confirmed in the econometric analysis. The probability of belonging to the last quartile of the environmental sustainability indicator is also strictly connected to that of belonging to the last quartile of respondents that showed to be the most responsible in the use of money (H2).

As for the variables that correlate with the two indexes, some common features and differences emerged.

Character traits have a significant impact on both dimensions, with the propensity to show trust in others and curiosity playing the most important roles. Young people who show greater confidence and curiosity are also more likely to be more aware of environmental issues and to use their money more responsibly. Scrupulousness and sociability have also positive impacts on both indexes. Fostering these personality features during childhood development can create a fertile ground for adopting behaviors coherent with the 2030 goals of future generations.

The main differences between the two indexes are in the role of external context variables. If the cultural background of the family has a significant impact, and with the expected sign, on the propensity to use money responsibly (higher levels of cultural background correspond to higher levels of money awareness), such an effect is not significant for environmental awareness.

The role of the parents as role-models and the school track chosen by students[16] is more relevant for the use of money, while they have a residual impact on environmental sensitivity. A possible explanation could lie in the fact that the use of money, for many young people, represents a practical skill that they already exercise managing small personal sums, or that they observe closely in the family (Moreno Herrero, Salas Velasco and Sanchez Campillo, 2018). Conversely, environmental issues are trend topics on social media and information channels (Elroy, Komendantova and Yosipof, 2024), thus youngsters may well have external role models to follow, and this may mitigate the role of the family in this regard. Nonetheless, our analysis confirms in depth the relevance of research hypothesis H1.

Due caution must be exercised when interpreting the data. The present survey provides a significant statistical association, but we believe that more insights can be gained from further research aimed at more clearly isolating the causal link between the two dimensions. Our data consent to measure a

[16] As for our proxy QSCHOOL the econometric analysis shows little significance of the variable. This may be due to the fact that financial education has been introduced only recently in the school curricula in a mandatory way.

correlation coefficient with a certain degree of precision. Nonetheless, the spillover between the two dimensions is a useful indication for educators and policymakers, and it can translate into the efficient use of public monetary funds to achieve a dual purpose.

If training on the responsible use of money is also associated with greater environmental awareness, it can be a formidable tool for sustainable education according to the 360° definition of sustainability that also emerges from frameworks such as the UN 2030 Agenda.

References

Bamber S., Moser G. (2007), "Twenty Years after Hines, Hungerford, and Tomera: A New Meta-Analysis of Psycho-Social Determinants of Pro-environmental Behaviour", *Journal of Environmental Psychology*, 27, pp. 14-25.

Bollani L., Bonadonna A., Peira G. (2019), "The Millennials' Concept of Sustainability in the Food Sector", *Sustainability*, 11, 2984, retrieved on April 7, 2021 from: https://doi.org/10.3390/su11102984.

Breßler J., Kappler S. (2017), "A Systematic Review of Education for Sustainable Development", *Chemnitz Economic Papers*, No. 007, Chemnitz University of Technology, Faculty of Economics and Business Administration, Chemnitz.

Brundtland G.H. (1987), *Our Common Future: Report of the World Commission on Environment and Development*, Geneva, UN-Dokument A/42/427, http://www.un-documents.net/ocf-ov.htm.

Cho M., Bonn M.A., Han S.J. (2018), "Generation Z's Sustainable Volunteering: Motivations, Attitudes and Job Performance", *Sustainability*, 10, 1400, retrieved on April 7, 2021 from: https://doi.org/10.3390/su10051400.

D'Alessio G., De Bonis R., Neri A., Rampazzi C. (2020), "L'alfabetizzazione finanziaria degli italiani: i risultati dell'indagine della Banca d'Italia del 2020", *Questioni di Economia e Finanza (Occasional Papers)*, Banca d'Italia, #588.

Dimock M. (2019), *Defining generations: where Millennials end and Generation Z begins*, PEW Research Center, retrieved on March 3, 2025 from: https://www.pewresearch.org/short-reads/2019/01/17/where-millennials-end-and-generation-z-begins/.

Elroy O., Komendantova N., Yosipof A. (2024), "Cyber-echoes of climate crisis: Unraveling anthropogenic climate change narratives on social media", *Current Research in Environmental Sustainability*, https://doi.org/10.1016/j.crsust.2024.100256, https://pure.iiasa.ac.at/19811.

Filmer D., Pritchett L.H. (2001), "Estimating Wealth Effect Without Expenditure Data – or Tears: An Application to Educational Enrollments in States of India", *Demography*, 38, pp. 115-132.

Greene H.W. (2017), *Econometric Analysis*, Eighth Edition Pearson Educational International, Upper Saddle River.

LeBaron B.A., Holmes E.K., Lorgensen B.L., Bean R.A. (2020), "Parental Financial Education During Childhood and Financial Behaviors of Emerging Adult", *Journal of Financial Counseling and Planning*, 31, pp. 42-54.

Li J., Leonas K.K. (2022), "Generation Z Female Consumers' Preferences for Swimwear Products with Sustainability-Relevant Attributes", *Journal of Global Fashion Marketing*, 13, pp. 44-60, retrieved on April 7, 2023 from: https://doi.org/10.1080/20932685.2021.1960580.

List M.K., Schmidt F.T.C., Mundt D., Foste-Eggers D. (2020), "Still Green at Fifteen? Investigating Environmental Awareness of the PISA 2015 Population: Cross-National Differences and Correlates, Sustainability", *MDPI, Open Access Journal*, 12, pp. 1-20.

Matthews G., Deary I. J., Whiteman M. C. (2003), *Personality traits*, Cambridge University Press, Cambridge, 2nd ed., retrieved on April 7, 2023 from: https://doi.org/10.1017/CBO9780511812736.

McAdams D.P. (1992), "The Five-Factor Model in Personality: A Critical Appraisal", *Journal of Personality*, 60, pp. 329-61.

McCoy L., Wang Y.-T., Chi T. (2021), "Why Is Collaborative Apparel Consumption Gaining Popularity? An Empirical Study of US Gen Z Consumers", *Sustainability*, 13, 8360, retrieved on April 7, 2023 from: https://doi.org/10.3390/su13158360.

Moreno-Herrero D., Salas-Velasco M., Sánchez-Campillo J. (2018), "Factors that influence the level of financial literacy among young people: The role of parental engagement and students' experiences with money matters", *Children and Youth Services Review*, 95, pp. 334-351, https://doi.org/10.1016/j.childyouth.2018.10.042.

Olszewski-Kubilius P., Worrell F. C., Subotnik R. F. (2018), "The Role of the Family in Talent Development", in S.I. Pfeiffer, E. Shaunessy-Dedrick, M. Foley-Nicpon (*eds.*), *APA Handbook of Giftedness and Talent*, American Psychological Association, pp. 465-477, retrieved on April 7, 2023 from: https://doi.org/10.1037/0000038-030.

Poirier J.M., Grepin K.A., Grignon M. (2020), "Approaches and Alternatives to the Wealth Index to Measure Socioeconomic Status Using Survey Data: A Critical Interpretive Synthesis", *Social Indicators Research*, 148, pp. 1-46.

Remund D. (2010), "Financial Literacy Explicated: The Case for a Clearer Definition in an Increasingly Complex Economy", *The Journal of Consumer Affairs*, 44 (2), pp. 276-295.

Samuelsson I.P., Park E. (2017), "How to Educate Children for Sustainable Learning and for a Sustainable World", *International Journal of Early Childhood*, 49, pp. 273-285.

Shearer E., Mitchell A. (2021), *News Use Across Social Platforms in 2020*, https://www.pewresearch.org/journalism/2021/01/12/news-use-across-social-media-platforms-in-2020/.

Sommerville M., Williams C. (2015), "Sustainability Education in Early Childhood: An Updated Review of Research in the Field", *Contemporary Issues in Early Childhood*, 16, pp. 102-117.

Thomas G.O., Fisher R., Whitmarsh L., Milfont T.L., Poortinga W. (2018), "The Impact of Parenthood on Environmental Attitudes and Behaviour: A Longitudinal Investigation of the Legacy Hypothesis", *Population and Environment*, 39, pp. 3261-276, retrieved on April 7, 2023 from: doi: 10.1007/s11111-017-0291-1, epub 2017 Dec 18. PMID: 29568145; PMCID: PMC5846977.

Vyas S., Kumaranayake L. (2006), "Constructing Socio-Economic Status Indices: How to Use Principal Components Analysis", *Health Policy and Planning*, 21, pp. 459-468.

White J., McCoy M., Watkins K. (2018), "Exploring the Relationship between Sustainability and Personal Finance", *The Forum Journal*, 22, pp. 1-12.

Tab. A1 – Variables' list

First name	*Description*	*Average*	*Source*	*Note*
Money Index	Synthetic economic awareness indicator	6.90	In-house calculation	See paragraph 3
Green Index	Synthetic indicator of environmental awareness	6.88	In-house calculation	See paragraph 3
Sex	Gender of the interviewee	1.48	Survey	Variable [1, 2]
Age	Age of the interviewee	15 and 1/2	Survey	Variables [1–6] 1 = 13; …; 6 = 18
Ses	Family cultural background	1.42	In house calculation	The weighted average of each parent's educational qualification and the number of books in the home. See note 7
Social	Use of social channels	0.85	Survey	Variable [0.1]
Istec	School type, technical institute	0.42	Survey	Variable [0.1]
Qschool	Proxy of the educational standard of students in the Region	11	ISTAT (May 2021)	Average of the median grades of the high school diploma of I and II degrees by region in 2019 – regional data
Qualair	Air quality	37.78	National System for Environmental Protection	Number of stations with a 50-µg/m3 limit exceedance for PM10 in 2020 – regional data
Personality	Character cluster	1.75	In house calculation	The variable that takes a value of 1 for the restless, 2 for the curious, and 3 for the conscientious
Self-esteem	Assertion evaluation: "Sometimes I feel like I'm not worth much"	5.40	Survey	Variable [0.10]
Curiosity	Assessment of the claim: "I like lessons where I learn something new that I didn't know before"	7.68	Survey	Variable [0.10]
Relationship with others	Assessment of the claim: "I try to get along with people even when they don't have my ideas"	7.11	Survey	Variable [0.10]
Scrupulousness	Evaluation of the claim: "I am very attentive and thorough when I do my school homework"	6.68	Survey	Variable [0.10]
Humor	Assessment of the claim: "My mood depends a lot on the judgment of others"	5.05	Survey	Variable [0.10]
Confidence	How much do you think others can be trusted?	5.74	Survey	Variable [0.10]
Friendship	How many people are you able to talk to about your intimate and private matters?	2.85	Survey	Variable [1,6] [none, 1, 2, 3, 4–6, 7+]
Role Model/Eco	How inspired are you by your parents? How much of a role model for money management are they to you?	1.69	Survey	Variable [1.4] [much, enough, little, not at all]
Role Model/Green	How inspired are you by your parents? How much of a role model for the management of environmental resources are they to you?	1.89	Survey	Variable [1.4] [much, enough, little, not at all]

Tab. A2 – Demographic features (sample of 400 young Italians aged between 13-18)

	Percentage values							
	Sex	*School*		*Type of high school*		*Preferred subjects*		
		Middle	*High*	*Lyceum*	*Technical*	*Humanities*	*Scientific*	*None*
Male	51.6	19.4	80.6	28.6	71.4	20.3	54.7	25.0
Female	48.4	15.0	85.0	45.7	54.3	43.5	32.3	24.2

Age		*Area*		*Number of inhabitants*	
13	16.7	N-W	25.8	up to 10k	26.0
14	16.6	N-E	19.1	10-30k	23.5
15	16.8	CE	18.9	30-100k	24.5
16	16.6	S-I	36.1	100-500k	14.6
17	16.6			>500K	11.5
18	16.8				

Tab. A3 – Family cultural background

Percentage values			
Qualification	*Parents*	*Father*	*Mother*
Bachelor's degree or more (2)	13.9		
Bachelor's degree or more (1)	18.7	22.5	23.6
High School Diploma (2)	26.5		
High School Diploma (1)	20.3	40.0	47.7
Lower Licences	20.7	37.5	28.7
How many books do have you at home?			
0-10	12.28		
11-25	18.73		
26-100	30.76		
101-200	22.64		
201-500	10.45		
More	5.14		

Tab. A4 – Personality Traits

Section A. Average value [range 0-10]	*Total*	*Restless*	*Curious*	*Conscientious*
I often lose control, and I am quick to anger	4.7	6.2	3.2	3.2
There are events that I do not like and that make me angry	6.7	6.9	6.5	6.4
Sometimes I feel like I'm not worth much	5.4	6.5	4.6	4.1
If someone says something bad about me I reply immediately by saying something nasty	5.2	6.4	3.7	4.2
I like lessons where I learn something new that I didn't know before	7.7	6.8	8.7	8.5
I'm interested in stories of people living in other countries of the world	6.4	5.9	7.6	6.5
I try to get along with people, even when they don't have my ideas	7.1	6.3	8.2	7.6
It's easy to socialize with me	7.2	6.3	8.5	7.8
I like reading different kinds of books (essays, novels, comics)	5.5	4.9	4.9	7.4
I always finish everything I start	6.6	5.7	6.8	8.1
I'm very attentive and thorough when I do my school homework	6.7	5.9	6.6	8.4
I tidy up everything I use as soon as I finish using them	6.0	5.4	5.4	7.7
In general, how much do you think you can trust people?	5.7	5.5	5.4	6.5
Section B. Percentage values	*Total*	*Restless*	*Curious*	*Conscientious*
Male	51.6	51.3	54.2	49.7
Female	48.4	48.7	45.8	50.3
High School	17.3	14.5	14.2	25.8
Middle School	82.7	85.5	85.8	74.2
North-West	25.8	24.2	25.8	29.1
North-East	19.1	23.7	16.9	12.5
Centre	18.9	19.3	20.5	16.6
South and Islands	36.1	32.8	36.9	41.8

Tab. A5 – Components of the synthetic indexes (row percentage)

Section A. Vote [range 0-10]	*Components Green Index*		
Score	[0-5]	[6-7]	[8-10]
Make the difference (1)	22.0	37.9	39.8
Personal commitment (2)	24.1	40.0	35.7
Adults' responsibility (3)	11.2	25.4	63.9
Government action (4)	28.8	39.1	32.0
Section B. Vote [range 0-10]	*Components Money Index*		
	[0-5]	[6-7]	[8-10]
Expenses under control (5)	23.0	30.3	47.1
For my own happiness (6)	37.6	32.8	20.4
Satisfaction (7)	11.4	26.1	62.8
	Periodically	*Occasionally*	*None*
Money at disposal (8)	25.1	42.6	32.0
	Yes		No
Think about how to use money (9)	84.9		15.1
	Project	*Fear*	*By chance*
Reasons to save (10)	77.0	10.3	13.4

Questions:
1) How much do you think that your daily attitude can make a difference to reduce global warming?
2) How much do you feel personally committed to reducing the problem of waste of natural resources?
3) How severe do you believe the impact of the lack of attention to sustainability themes by adults will be?
4) How likely do you believe that Governments might act to change the world development model so that it becomes more sustainable over time?
5) I keep my expenses under control and, if something is not necessary, I avoid buying it
6) My money is used only and exclusively for my happiness, without thinking about the impact it can generate on others
7) If you want something you like, and you cannot afford it immediately, how much satisfaction do you feel in being able to buy it when you reach the needed amount?
8) Do you usually have money at your disposal?
9) If you have money at your disposal, do you have the habit of thinking about how to use it?
10) Why do you save money?

Tab. A6 – Respondents' evaluation of the relevance of environmental and economic sustainability

Percentage values	*x <=2*	*>2 x <=3*	*>3 x <=4*	*>4 x <=5*	*>5 x <=6*	*>6 x <=7*	*>7 x <=8*	*>8 x <=9*	*>9 x <=10*
Green Index	0.75	1.25	2.00	8.25	13.00	27.70	24.25	17.75	5.00
Money Index		1.75	3.50	7.75	15.50	20.50	23.75	18.25	9.00
Average value [range 0-10]	*Mean*	*Std.Dev.*	*Min*	*Max*	*Male*	*Female*	*Restless*	*Curious*	*Consc.*
Green Index	6.88	1.53	0.00	10.00	6.91 (a)	6.87 (a)	6.49(b)	7.17(c)	7.38(d)
Money Index	6.90	1.61	2.43	10.00	6.82 (e)	6.91(e)	6.37(f)	7.01(g)	7.65(h)

Average value [range 0-10]	*Mean*				*Std. dev.*			
	1° quartile	*2° quartile*	*3° quartile*	*4° quartile*	*1° quartile*	*2° quartile*	*3° quartile*	*4° quartile*
Green Index	4.88	6.52	7.44	8.67	0.12	0.03	0.03	0.05
Money Index	4.75	6.43	7.56	8.86	0.09	0.04	0.03	0.05

Adjusted Wald test:
(a) Male=Female adj. Wald test 0.07 (p=0.7880)
(b) Restless=Curious adj. Wald test 11.22 (p=0.0009)
(c) Restless=Conscientious adj. Wald test 23.06 (p=0.0000)
(d) Curious=Conscientious adj. Wald test 0.93 (p=0.3347)
(e) Male=Female adj. Wald test 0.31 (p=0.5751)
(f) Restless=Curious adj. Wald test 11.99 (p=0.0006)
(g) Restless=Conscientious adj. Wald test 56.66 (p=0.0000)
(h) Curious=Conscientious adj. Wald Test 10.56 (p=0.0013)

Tab. A7 – Joint presence by quartile (percentage values)

	Green Index			
Money Index	*4° quartile*	*3° quartile*	*2° quartile*	*1° quartile*
4° quartile	48.5	26.6	11,7	13.4
3° quartile	21.1	38.2	25.2	15.0
2° quartile	20.7	23.6	25.6	28.0
1° quartile	9.6	11.6	37.6	43.6

Tab. A8a – Survey Linear Regressions – Dependent variable Money Index

	M1		M2		M2.A		M3		M4		M4.A		M4.B	
Exogenous Variables											*(femmine)*		*(maschi)*	
	Coeff	St.Error	Coeff	St.Error	Coeff	St.Error	Coeff	St.Error	Coeff	St.Error	Coeff	St.Error	Coeff	St.Error
SEX	0.0563	0.1688	0.0650	0.1646	0.0711	0.1643	-0.0193	0.1518	-0.0344	0.1476				
AGE	0.0743	0.0479	0.0723	0.0475	0.0680	0.0480	0.0943	0.0486 **	0.0801	0.0472 *	**0.1067953**	0.0660	0.0497	0.0670
SES	-0.2755	0.0547 ***	-0.2417	0.0549 ***	-0.2420	0.0550 ***	-0.1459	0.0518 ***	-0.1341	0.0466 ***	-0.1728	0.0610 ***	-0.0990	0.0681
SOCIAL	0.4962	0.2464 **	0.4897	0.2388 **	0.4723	0.2373 **	0.3891	0.2171 *	0.4559	0.2190 **	-0.0883	0.3940	0.7117	0.2524 ***
ISTEC	0.3671	0.1777 **	0.4738	0.1746 ***	0.4740	0.1746 ***	0.4369	0.1649 ***	0.4611	0.1547 ***	0.3717	0.1959 *	0.6022	0.2220 ***
QSCHOOL	0.5214	0.1951 ***	0.4833	0.1912 **	0.4794	0.1913 **	0.3162	0.1770 *	0.3296	0.1720 *	0.1414	0.2213	0.3000	0.2585
PERSONALITY			0.4216	0.0895 ***										
2					0.5533	0.1929 ***								
3					0.8168	0.1800 ***								
SELF-ESTEEM							0.0353	0.0300	0.0406	0.0285	0.0790	0.0425 *	0.0061	0.0410
CURIOSITY							0.1616	0.0461 ***	0.1067	0.0470 **	0.0519	0.0637	0.1161	0.0710
RELATIONSHIP WITH OTHERS							0.1167	0.0459 **	0.1014	0.0442 **	0.1126	0.0557 **	0.1318	0.0629 **
SCRUPULOUSNESS							0.1094	0.0369 ***	0.0686	0.0366 *	0.0637	0.0488	0.0841	0.0527
HUMOR							-0.0323	0.0355	-0.0416	0.0343	-0.1115	0.0482 **	0.0208	0.0445
CONFIDENCE							0.1121	0.0410 ***	0.0765	0.0409 *	0.0597	0.0528	0.1034	0.0520 **
FRIENDSHIP							0.0464	0.0642	0.0266	0.0605	0.0234	0.0873	0.0164	0.0818
ROLEMODEL/Eco							-0.1599	0.1166	-0.1551	0.1120	0.0894	0.1429	-0.3377	0.1645 **
GREEN INDEX									0.2507	0.0601 ***	0.3821	0.0811 ***	0.1336	0.0764 *
CONS	0.6204	2.1559	0.2055	2.1325	0.6649	2.1253	-0.5358	2.0320	-1.3385	1.9896	0.4450	2.5540	-0.8073	3.0117
R2	0.0993		0.1436		0.1451		0.2804		0.3184		0.3741		0.3398	
F(ve,dof)	6.6 ***		8.68 ***		7.94 ***		11.00 ***		13.56 ***		12.27 ***		8.51 ***	
LINKTEST	NO		NO		NO		NO		NO		NO		NO	
VIF									1.61					

LINKTEST NO means that the single-equation model is specified correctly; ve= number of exogenous variables, dof=degrees of freedom; VIF = Variable inflation factors test for the presence of multicollinearity; *** p<0.01; **p<0.05; * p<0.10

Tab. A8b – Survey Linear Regressions – Dependent Variable Green Index

	M1		M2		M2.A		M3		M4		M4.A		M4.B	
Exogenous Variables											*(femmine)*		*(maschi)*	
	Coeff	St.Error	Coeff	St.Error	Coeff	St.Error	Coeff	St.Error	Coeff	St.Error	Coeff	St.Error	Coeff	St.Error
SEX	**0.0980**	**0.1667**	**0.1098**	**0.1582**	**0.1103**	**0.1584**	0.0591	**0.1360**	0.0582	**0.1317**				
AGE	0.0337	0.0469	0.0311	0.0445	0.0307	0.0444	0.0553	0.0405	0.0398	0.0396	-0.0968	0.0490 **	0.1403	0.0565 **
SES	-0.1991	0.0583 ***	-0.1534	0.0567 **	-0.1535	0.0568 ***	-0.0480	0.0521	-0.0145	0.0475	-0.0095	0.0546	0.0165	0.0686
SOCIAL	-0.0674	0.2211	-0.0761	0.2091	-0.0776	0.2088	-0.2814	0.1597 *	-0.3521	0.1628 **	-0.1544	0.2283	-0.5203	0.2219 **
ISTEC	-0.2520	0.1767	-0.1080	0.1707	-0.1080	0.1707	-0.0947	0.1530	-0.1993	0.1458	-0.2247	0.2188	-0.1902	0.1995
QSCHOOL	0.1766	0.1899	0.1252	0.1816	0.1249	0.1811	-0.0068	0.1965	-0.0886	0.1960	0.2388	0.2470	-0.3186	0.2898
PERSONALITY			0.5688	0.0823 ***										
2					0.5803	0.1779 ***								
3					1.1353	0.1647 ***								
SELF-ESTEEM							-0.0212	0.0301	-0.0266	0.0284	0.0120	0.0450	-0.0668	0.0341 *
CURIOSITY							0.2198	0.0465 ***	0.1825	0.0462 ***	0.3014	0.0536 ***	0.0605	0.0603
RELATIONSHIP WITH OTHERS							0.0598	0.0418	0.0348	0.0403	-0.0495	0.0492	0.1179	0.0671 *
SCRUPULOUSNESS							0.1631	0.0391 ***	0.1372	0.0380 ***	0.0804	0.0461 *	0.2168	0.0537 ***
HUMOR							0.0369	0.0311	0.0441	0.0297	0.0427	0.0462	0.0436	0.0364
CONFIDENCE							0.1429	0.0378 ***	0.1166	0.0372 ***	0.0511	0.0420	0.1662	0.0545 ***
FRIENDSHIP							0.0794	0.0616	0.0654	0.0580	0.0756	0.0822	0.0830	0.0748
QUALAIR							0.0021	0.0042	0.0015	0.0041	0.0020	0.0061	0.0005	0.0053
ROLEMODEL/green							-0.0035	0.1042	-0.0359	0.1006	0.0375	0.1174	-0.1454	0.1536
MONEY INDEX									0.2110	0.0488 ***	0.2928	0.0671 ***	0.1134	0.0629 *
CONS	5.0958	2.1243	4.5360	2.0296	5.1081	2.0216	2.6153	2.3534	3.0782	2.3354	-0.6270	2.8184	6.0181	3.5409 *
R2	0.0619		0.1502		0.1502		0.3390		0.3741		0.4570		0.4086	
F(ve,dof)	2.05 ***		8.75 ***		7.71 ***		8.92 ***		12.43 ***		9.67 ***		10.55 ***	
LINKTEST	NO		NO		NO		NO		NO		NO		NO	
VIF									1.58					

LINKTEST NO means that the single-equation model is specified correctly; ve= number of exogenous variables, dof=degrees of freedom; VIF = Variable inflation factors test for the presence of multicollinearity; *** p<0.01; **p<0.05; * p<0.10

Tab. A9 – Ordered Probit

Dependent variable	*Money quartiles*		*Green quartiles*	
	M4		*M4*	
Exogenous variables	*Coeff.*	*Std. error*	*Coeff.*	*Std. error*
Sex	-0.0509	0.1210	0.0327	0.1197
Age				
14	0.3787	0.2075*	0.0347	0.2000
15	0.3929	0.1917**	-0.1008	0.2061
16	0.4922	0.1996**	-0.1046	0.1936
17	0.2931	0.2059	0.0413	0.2174
18	0.3987	0.2181*	0.1339	0.1984
Ses	-0.1066	0.0390***	-0.0087	0.0403
Social	0.1823	0.1880	-0.1830	0.1627
Istec	0.3401	0.1317**	-0.1344	0.1385
Qschool	0.2420	0.1441*	-0.0053	0.1630
Self-eteem	0.0427	0.0250*	-0.0310	0.0265***
Curiosity	0.0948	0.0368**	0.1244	0.0416
Relationship with others	0.0837	0.0361**	0.0487	0.0333
Scrupulousness	0.0625	0.0309**	0.0931	0.0331***
Humor	-0.0417	0.0286	0.0330	0.0279
Confidence	0.0596	0.0339*	0.1007	0.0332***
Friendship	0.0434	0.0501	0.0156	0.0528
Qualair			-0.0002	0.0036
Rolemodel/Eco				
2	-0.2425	0.1261*		
3	-0.0613	0.2503		
4	0.1448	0.3800		
Rolemodel/Green				
2			-0.2202	0.1309*
3			0.0198	0.2240
4			-0.1828	0.4124
Green quartiles	0.2830	0.0621***		
Money quartiles			0.2694	0.0615***
cut/1	4.7821		1.7795	
cut/2	5.6010		2.5869	
cut/3	6.4321		3.4620	
F(ve, dof)	ve = 21 dof = 379	6.36***	ve = 22 dof = 378	5.69***

ve = number exogenous variables, dof = degrees of freedom; *** p<0.01; **p<0.05; * p<0.10

Tab. A10 – Margins ordered probit (delta method)

Dependent variable	*Money quartiles*		*Green quartiles*	
	Outcome 4° quartile		*Outcome 4° quartile*	
	Coeff.	*Std. error*	*Coeff.*	*Std. error*
Sex (Female)	-0.0132	0.0314	0.0084	0.0307
Age				
14	0.0907	0.0508*	0.0090	0.0517
15	0.0945	0.0456**	-0.0251	0.0512
16	0.1219	0.0493**	-0.0260	0.0482
17	0.0683	0.0485	0.0107	0.0563
18	0.0961	0.0533*	0.0355	0.0526
Ses	-0.0277	0.0101***	-0.0022	0.0103
Social	0.0473	0.0487	-0.0469	0.0415
Istec	0.0883	0.0340***	-0.0344	0.0358
Qschool	0.0628	0.0373*	-0.0014	0.0417
Self-esteem	0.0111	0.0065*	-0.0079	0.0067
Curiosity	0.0246	0.0096**	0.0318	0.0106***
Relationship with others	0.0217	0.0093**	0.0125	0.0085
Scrupulousness	0.0162	0.0080**	0.0238	0.0083***
Humor	-0.0108	0.0074	0.0084	0.0071
Confidence	0.0155	0.0087*	0.0258	0.0085***
Friendship	0.0113	0.0130	0.0040	0.0135
Qualair			-0.0001	0.0009
Rolemodel/Eco				
2	-0.0634	0.0334**		
3	-0.0168	0.0679		
4	0.0415	0.1115		
Rolemodel/Green				
2			-0.0571	0.0348*
3			0.0054	0.0619
4			-0.0479	0.1035
Green quartiles	0.0735	0.0156***		
Money quartiles			0.0690	0.0154***

*** p < 0.01; **p < 0.05; * p < 0.10

Tab. A11 – Bivariate probit

Dependent variable	*Money top*		*Green top*	
	M4		*M4*	
Exogenous variable				
	Coeff.	*Std. error*	*Coeff.*	*Std. error*
Sex	-0.0688	0.1645	0.0656	0.1602
Age				
14	0.8090	0.2936***	0.1178	0.2551
15	0.5130	0.3087*	-0.1108	0.2860
16	0.5503	0.3071*	-0.1844	0.2603
17	0.6982	0.2996**	0.0380	0.2699
18	0.7725	0.3029**	0.1452	0.2644
Ses	-0.0870	0.0519*	-0.0078	0.0601
Social	0.0025	0.2236	-0.0916	0.2086
Istec	0.3925	0.1745**	-0.1957	0.1847
Qschool	0.4299	0.1954**	-0.1152	0.2098
Self-esteem	0.0379	0.0318	-0.0011	0.0322
Curiosity	0.1333	0.0573**	0.1475	0.0584**
Relationship with others	0.1259	0.0503**	0.1137	0.0517**
Scrupulousness	0.1160	0.0451**	0.1018	0.0468**
Humor	-0.0266	0.0356	0.0000	0.0324
Confidence	0.0443	0.0430	0.0839	0.0440*
Friendship	0.1209	0.0654*	0.0171	0.0732
Qualair			-0.0015	0.0049
Rolemodel/Eco				
2	-0.2826	0.1585*		
3	-0.1108	0.3952		
4	0.3679	0.4294		
Rolemodel/Green				
2			-0.4592	0.1689***
3			-0.0697	0.2601
4			-0.2832	0.6251
CONS	-9.3978	2.2851***	-2.2765	2.4509
F(40,360)	3.10	***		
athrho	0.4246	0.1116***		
rho	0.4008	0.0946***		

rho = error terms cross-equation correlation cfr. equation (6); *** p < 0.01; ** p < 0.05; * p < 0.10

Tab. A12 – Margins bivariate probit (delta method)

	M4		
	dy/dx	*Std. error*	*z*
Sex (Female)	0.0001	0.0228	0.00
Age			
14	0.0841	0.0389	2.16**
15	0.0341	0.0373	0.91
16	0.0305	0.0317	0.96
17	0.0645	0.0365	1.77*
18	0.0835	0.0397	2.10**
Ses	-0.0083	0.0082	-1.01
Social (Si)	-0.0083	0.0308	-0.27
Istec	0.0150	0.0273	0.55
Qschool	0.0269	0.0282	0.95
Self-esteem	0.0032	0.0047	0.68
Curiosity	0.0254	0.0090	2.81***
Relationship with others	0.0216	0.0078	2.77***
Scrupuluosness	0.0196	0.0070	2.82***
Humor	-0.0023	0.0050	-0.46
Confidence	0.0117	0.0063	1.85*
Friendship	0.0122	0.0101	1.21
Rolemodel/Eco			
2	-0.0249	0.0138	-1.81*
3	-0.0100	0.0352	-0.28
4	0.0316	0.0350	0.90
Qualair	-0.0001	0.0004	-0.32
Rolemodel/Green			
2	-0.0435	0.0158	-2.76***
3	-0.0066	0.0248	-0.27
4	-0.0270	0.0597	-0.45

*** p < 0.01; ** p < 0.05; * p < 0.10

5. School, distance learning and competence levels in Italy: differences by territory

by Barbara Baldazzi, Claudia Busetti, Silvia Montecolle

In the 2020/2021 school year, Italian upper secondary schools implemented measures to mitigate the impact of school closures and facilitate the continuity of education for all students through remote learning.

Face-to-face teaching periods have been supplemented with periods of distance teaching. The hours in distance teaching exceeded those in presence with some differences among Italian regions. In addition, INVALSI tests conducted in the 2020/2021 school year show a generalized loss of learning in upper secondary school, particularly in literacy and numeracy and the loss of competence is evident in all Italian regions (Campodifiori *et al.*, 2021; INVALSI, 2021).

In addition to distance learning, several other factors may have influenced students' competencies, among others: the situation of the school in terms of connectivity/electronic devices/overall quality of the school environment; the situation in the household in terms of connection and/or electronic devices; loss of social skills; students' previous school grades.

This study explores these aspects and their relationship with the competence levels of students in grade 13 in reading (literacy) and mathematics (numeracy) during the pandemic.

Exploratory analysis of some statistical indicators from different sources available at Italian regional level for upper secondary schools could help in finding possible associations between the analyzed aspects. In addition to INVALSI data on the 2021 competence levels of students from grade 13 and their variation from the 2018/2019 school year, the analysis focused on critical issues faced by the school system as a result of the pandemic: distance learning and related IT problems, student perceptions, increased teacher turnover.

To synthesize information from the different sources a principal component analysis (PCA) was conducted at the regional level and a subsequent

composite indicator was created. The difficulties of the health emergency did not substantially change the geography of learning inequalities compared to previous school year, but an increase of difficulties is observed for some regions in the South and Islands, partly due to the problems that schools and households have had in adjusting to the required changes.

In Italia, nell'a.s. 2020/2021, il protrarsi dell'emergenza sanitaria legata al COVID-19 ha comportato l'utilizzo della didattica a distanza nelle scuole come strumento flessibile per gestire i contagi, soprattutto nelle scuole secondarie di secondo grado, dove si è previsto un alternarsi tra lezioni a distanza e in presenza. Sebbene con alcune differenze tra le regioni italiane, le ore di insegnamento a distanza nelle scuole secondarie superiori hanno superato quelle in presenza.

Nell'a.s. 2020/2021, le prove INVALSI hanno mostrato una perdita generalizzata di competenza tra gli studenti dell'ultimo anno della scuola secondaria superiore, in particolare nelle materie Italiano e Matematica, registrata in tutte le regioni italiane (Campodifiori et al.*, 2021; INVALSI, 2021).*

La didattica svolta a distanza non è stato l'unico elemento che può aver influenzato le competenze degli studenti, ma ci sono stati diversi aspetti che hanno influito: la mancanza di dispositivi elettronici in famiglia o a scuola, i frequenti problemi di connessione, la perdita di abilità sociali, la situazione scolastica precedente e altri ancora. Il presente lavoro esplora questi aspetti e la loro relazione con i livelli di competenza in italiano e matematica degli studenti del quinto anno della scuola secondaria di secondo grado durante la pandemia.

L'analisi esplorativa di alcuni indicatori statistici provenienti da fonti diverse e disponibili a livello regionale per le scuole secondarie di secondo grado potrebbe aiutare a trovare possibili associazioni tra gli aspetti analizzati. Oltre ai dati INVALSI sui livelli di competenza 2020/2021 degli studenti dell'ultimo anno delle scuole secondarie superiori e alla loro variazione rispetto all'anno scolastico 2018/2019, rappresentativa del periodo prepandemico, l'analisi si è concentrata sulle criticità affrontate dal sistema scolastico a seguito della pandemia: l'apprendimento a distanza e i relativi problemi informatici a esso connessi, la percezione degli studenti, l'aumento del turnover degli insegnanti a seguito delle restrizioni nella mobilità territoriale e delle misure di quarantena e auto-sorveglianza.

Per sintetizzare l'informazione proveniente da diversi fonti è stata condotta una analisi in componenti principali (ACP) a livello regionale e un successivo indicatore composito è stato calcolato sulle componenti significative. Le difficoltà dell'emergenza sanitaria non hanno sostanzialmente cam-

biato la geografia delle disuguaglianze di apprendimento rispetto all'anno scolastico precedente, ma si osserva un aumento delle difficoltà per alcune regioni del Sud e delle Isole, in parte a causa dei problemi che le scuole e le famiglie hanno avuto nell'adattarsi ai cambiamenti richiesti.

1. Introduction

In the 2020/2021 school year, in Italian upper secondary schools, distance teaching has been planned as an alternative to face-to-face teaching. Face-to-face teaching periods were supplemented with periods of distance teaching, as a flexible tool to manage peaks of COVID-19 contagion (UNESCO *et al.*, 2020).

An analysis was conducted not only to study possible associations between distance learning and competencies, but also with other aspects that may have had an impact on the level of competency at the regional level.

This study examines reading and mathematics outcomes in Italy, in order to evaluate the effectiveness of distance learning mechanisms utilised during school closures.

Objective of this paper is to use different sources and indicators available at the regional level and to jointly analyze indicators that relate to the relevant thematic areas: 1) the INVALSI proficiency levels of students in the last year of upper secondary school, 2) critical issues in school organization, 3) critical issues in household organization and 4) student's point of view. Finally, to provide a summary map of the situation in Italy.

Distance learning was an organizational form introduced to cope with the health emergency. Its application was proposed to households and schools that had more or less pronounced organizational problems at the territorial level (Bagnarol *et al.*, 2021; ISTAT, 2019, 2020), especially from the point of view of the availability of computer devices and the ability to use them. In addition, there was the discomfort that the young people may have experienced during the pandemic, e.g. the social distancing that has limited the possibility of sharing experiences and skills. For this reason, we thought to introduce some indicators that would give the reader the students' point of view, their evaluation of distance learning and their level of competence compared to the past.

2. Distance teaching, distance learning and learning loss

Following the literature and for the purpose of this study, we define distance teaching as teaching via the internet, where students are not physically present in a classroom. In the 2020/2021 school year, a sizable share of schools, 81%, activated distance teaching at least once during the year, and only 3% of schools conducted exclusively face-to-face classes (ISTAT, 2022a). Observing the percentage of hours conducted at a distance in the different school levels, in upper secondary schools hours in distance teaching exceeded those in-presence (57%; Figure 1). The use of distance teaching has occurred with different frequency and intensity depending on the regional territory[1]. Moreover, in southern Italian schools, the share of hours carried out at a distance exceeded the national average value by 20 percentage points. In contrast, distance learning was less frequent in the North-East regions.

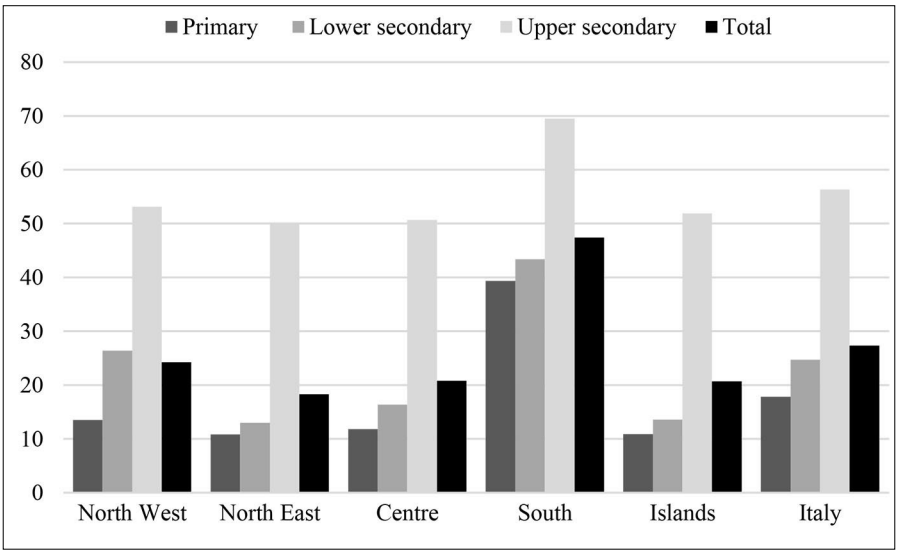


Fig. 1 – Distance teaching hours by geographical area and school level (percentage). School year 2020/2021

Source: ISTAT, *Survey on integration of students with disability*

[1] The COVID-19 pandemic has led to school closures in most countries around the world and has interrupted the school attendance of at least 1.5 billion students in 2020 and 2021 (Vincent *et al.*, 2022).

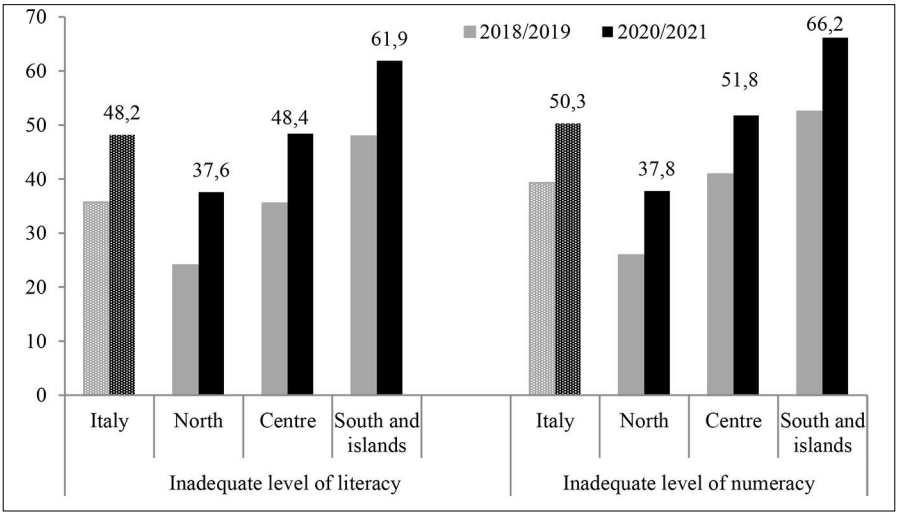


Fig. 2 – Inadequate level of Italian and Mathematics in upper secondary school (percentage). School years 2018/2019 and 2020/2021

Source: INVALSI, Proficiency levels (students in grade 13)

Although a large share of students attending grade 13 showed already in 2019 inadequate levels of proficiency[2] in Italian and Mathematics, in 2021 this share has increased (ISTAT, 2022b; INVALSI, 2021). In 2021, 48 students out of 100 did not achieve a sufficient knowledge in Italian (+12 percent points compared to 2019) and 50 out of 100 did not achieve sufficient skills in Mathematics (+11 percent points compared to 2019; Figure 2). The loss of competence is evident in all Italian regions[3]. In many regions of the Centre, South and Islands, a high proportion of students did not reach the national average of 50% in Italian: low performers in national language were 69.3% in Calabria, 65.6% in Campania, 61.6% in Sicilia, 60.7% in Puglia,

[2] The levels for Italian and Mathematics are established on an empirical basis and correspond to score bands on a quantitative scale (Rasch), on which students' responses and their level of proficiency are ordered on the same continuum. They thus refer not to a criterion but to a statistical norm directly constructed on student performance. There are five levels for Italian and Math, in ascending order from level 1 to level 5.

[3] The learning loss is confirmed at a global level, as well. In fact «between 2018 and 2022, and on average across 35 OECD countries, mean performance dropped by almost 15 score points in mathematics and 10 score points in reading» (OECD, 2023, p. 153). «In mathematics and reading, most countries and economies that can compare PISA 2022 results to PISA 2018 dropped in mean performance (41 countries/economies in mathematics, 35 in reading)» (*ibid.*).

58.3% in Sardegna, 55% in Basilicata and 50.5% in Lazio. In Mathematics the percentages of students below the minimum level of proficiency were even higher, involving the following regions: Calabria 72.8%, Sardegna 69.7%, Campania 68.3%, Sicilia 66.5%, Puglia 64.3%, Basilicata 57.8%, Lazio 57%, Abruzzo 55.1% and Umbria 50%.

3. Data and methods

3.1. Sources of data

A principal component analysis (PCA) is conducted at the regional level and a subsequent composite indicator is computed to synthesize information from different sources available for upper secondary schools. Several indicators, listed in Table 1 and related to different thematic areas, are identified and then jointly analyzed in the PCA analysis.

For proficiency levels achieved by the students, the analysis focuses on the shares of students from upper secondary school (grade 13) which don't reach adequate levels of Italian or Mathematic (*Italian, Math*) skills during the 2020/2021 school year, according to INVALSI national tests. Then, for both Italian and Mathematics tests, the percentage changes of 2020/2021 with the 2018/2019 school year are computed, comparing the 2021 percentage values with those of the pre-pandemic period (*Var_Ita, Var_Math*).

For critical issues faced by the schools concerning organization, some data from the ISTAT Survey on the integration of students with disabilities are selected. In particular, the analysis considered: i) the percentage of distance learning hours (excluding quarantines) on total planned teaching hours in upper secondary schools (*DAD*); ii) the percentage of students who received PCs or tablets from their upper secondary schools for use during the pandemic (*Devices*). These indicators can indirectly measure the frequency on the use of distance learning by the schools. Another indicator taken into consideration, computed from Ministry of Education data, is the variation in the number of substitute teachers in upper secondary schools during the 2020/2021 school year compared to the relative values observed in the 2018/2019 school year (*Sub_teach*). This indicator can provide information about the minor or major resort to substitute teachers with respect to the pre-pandemic period. The differences may be due to restrictions on territorial mobility and quarantine and health monitoring measures.

Tab. 1 – Indicators and sources: name, description, unit of measure, source and year

Name	*Description*	*Unit of measure*	*Source*	*Year*
Proficiency level				
Italian	Inadequate level of literacy (students in grade 13)	Percentage values	INVALSI	2020/2021
Var_Ita	Inadequate level of literacy (students in grade 13)	Percentage change, 2020/2021 on 2018/2019	INVALSI	2020/2021
Math	Inadequate level of numeracy (students in grade 13)	Percentage values	INVALSI	2020/2021
Var_Math	Inadequate level of numeracy (students in grade 13)	Percentage change, 2020/2021 on 2018/2019	INVALSI	2020/2021
Critical issues in school organization				
DAD	Hours of distance learning (excluding quarantines) in upper secondary schools on total hours of learning activities	Percentage values	ISTAT, *Survey on integration of students with disabilities*	2020/2021
Devices	Upper secondary school students who received devices from their schools for use during the pandemic (PCs, tablets)	Percentage values	ISTAT, *Survey on integration of students with disabilities*	2020/2021
Sub_teach	Substitute teachers in upper secondary schools	Percentage change, 2020/2021 on 2018/2019.	ISTAT, elaborations on Ministry of Education data	2020/2021
Critical issues in household organization				
Prob_Conn	Upper secondary school students who report internet connection problems at home	Percentage values	ISTAT, *Survey on the behaviour, attitude and future projects of children and adolescents*	2021
No_Internet	Households reporting to have at least one computer and Internet connection	Percentage values (polarity reversed to conform with other indicators)	ISTAT, *Survey on aspects of daily life*	2021
Students point of view				
Pref_Pres	Upper secondary school students who report that they prefer teaching in presence	Percentage values	ISTAT, Survey on the behaviour, attitude and future projects of children and adolescents	2021
DAD_ko	Upper secondary school students that have taken distance learning and who report that distance learning is more tiring	Percentage values	ISTAT, *Survey on the behaviour, attitude and future projects of children and adolescents*	2021
Scores_ko	Upper secondary school students that have taken distance learning and who report that school marks are negatively affected by distance learning	Percentage values	ISTAT, *Survey on the behaviour, attitude and future projects of children and adolescents*	2021

For critical issues faced by households, two indicators from ISTAT surveys were chosen to assess the computer availability present in students' households. The presence of a personal computer and a smooth connection (*Prob_conn; No_internet*) are elements that ensure better distance learning.

Finally, students' opinion on distance learning and its impact on their school performance are subjective summary indicators that help define the picture of organizational difficulties (*Pref_pres; DAD_ko, Scores_ko*).

3.2. Composite PCA Index

Through a principal components analysis, the composite PCA Index was calculated, as the weighted arithmetic mean with the contributions that the variables made to the formation of the 3 components according to the formula:

$$\lambda_I \frac{\sum_{v=1}^{V} z_{vi} a_{iv} + \lambda_{II} \sum_{v=1}^{V} z_{vii} a_{iiv} + \lambda_{III} \sum_{v=1}^{V} z_{viii} a_{iiiv}}{\lambda_I + \lambda_{II} + \lambda_{III}}$$

where *z* are the standardised variables, *a* represents the three principal components (I, II and III) and λ expresses the variance of the principal components.

The method adopted has the advantage of being based on uncorrelated factors (Gismondi and Russo, 2004). In addition, the weight for the variables in the computation of the indicator is given by the corresponding shares of variance explained by the factors: variables with higher factorial coordinates are thus given more weight.

4. Main results

4.1. PCA results

The first three components extracted with the principal components analysis account for 79% of the variability of all indicators. The first component (49%) is represented on the negative x-axis by the variables indicating greater difficulty toward distance learning reported by students (Figure 3). On the other hand, the variables concerning poor skills achieved in both Italian and Mathematics, more frequent use of distance learning, greater use of substitute teachers than in the 2018/2019 school year, and less internet presence in the household are located on the positive x-axis. On the negative semi-axis

mainly the northern regions are projected (Figure 4), while on the positive x-axis there are six southern regions (Sicilia, Abruzzo, Basilicata, Calabria, Campania and Puglia).

The second component (19%) is characterized on the negative y-axis by the variable related to *devices* (PCs, tablets) and on the positive y-axis by the variables indicating greater negative variation in Mathematics and Italian skills. On the positive semi-axis Veneto, the Autonomous Province of Bolzano, Friuli-Venezia Giulia, Marche, Liguria, Abruzzo and Puglia are projected, regions where the decline in learning was highest even though, for students in the northern regions, average results still remain higher than for students in the South.

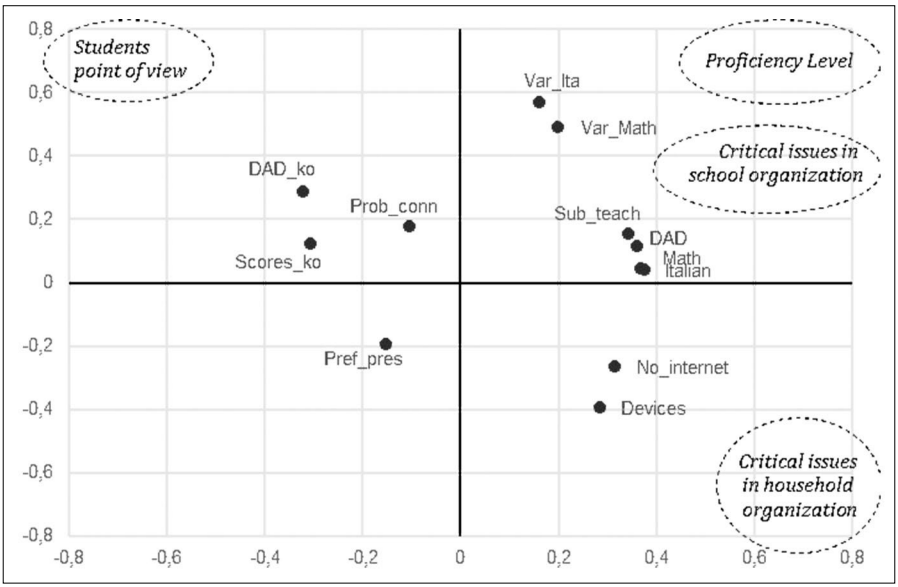


Fig. 3 – Principal component analysis (axes 1 e 2) – Variables

The third component (11%) is characterized on the positive semi-axis by variables derived from students' reports on the preference of face-to-face teaching to distance learning, the negative influence of distance learning on school marks and internet connection problems; the regions projecting on this semi-axis are Marche, Sardegna and Campania.

4.2. PCA plots of variables and regions

In the top right quadrant of the plane, identified by the first two components, the regions with the highest share of low performers and the most marked deteriorations in students' performance on national student assessments are located (Figure 4): Abruzzo and especially Puglia and Campania. In these regions, distance learning hours were among the highest in Italy, and there was also the highest use of substitute teachers to carry out annual substitutions.

In the bottom right quadrant of the plane, the regions with a high percentage of low-performing students are located, but in these regions the worsening in students' performance compared to the 2018/2019 school year is lower. Calabria, Sicilia and Basilicata are in this quadrant, with Umbria, closer to the average values. Basilicata is also characterized by a high demand for devices from households.

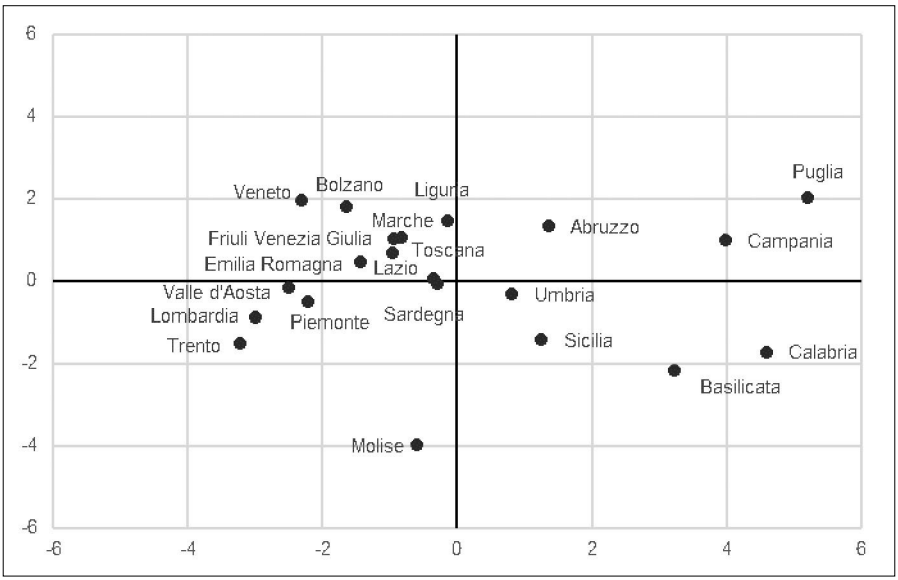


Fig. 4 – Principal component analysis (axes 1 e 2) – Italian regions

In the top left quadrant are regions with marked declines in proficiency levels compared to the 2018/2019 school year, but they are still above the national average. These regions are Veneto, the Autonomous Province of Bolzano, Liguria, Toscana, Marche and Friuli-Venezia Giulia, Emilia Romagna and Lazio (the latter two regions are very close to the national aver-

age). These also have the highest share of students who perceive a worsening in their school performance, who found distance learning tiring and have also experienced internet connection problems.

Finally, in the bottom left quadrant are regions where the share of low performers and the negative change in skills compared to the pre-pandemic school year are low. These are Valle d'Aosta, Lombardia, the Autonomous Province of Trento, Piemonte and Molise (in addition to Sardegna, which ranks very close to the average). In particular, the autonomous province of Trento is confirmed as the best performing region on both skills in the 2020/2021 school year, followed by Lombardia, Valle d'Aosta and Piemonte. Molise stands out for being the only region with no deterioration.

4.3. Composite PCA Index results

The composite PCA Index is calculated on the three axes (Fig. 5). The variables that weigh most heavily are students' inadequate level of literacy and numeracy, the teaching hours allocated to distance (for the first axis), the percentage change for literacy and numeracy skills (for the second axis) and the connection problems and variable of students who report that they prefer teaching in presence. The composite index calculated shows a North-South divide with some exceptions.

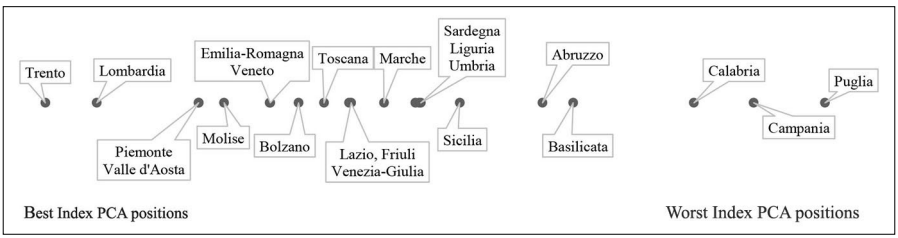


Fig. 5 – PCA Index by Italian regions

The northern regions lead the ranking of regions that have coped better with the pandemic and the consequences on the school system with the exception of Friuli-Venezia Giulia and Liguria, which are in the middle of the ranking. Instead, Molise and Sardegna diverge from the southern regions (positioned at the bottom of the ranking) due to better performance in learning and internet presence at home. In contrast, students in Puglia suffered a higher loss in learning.

5. Next steps and conclusions

The current study explored aspects of the Italian education system with possible impact on the competence levels of students in grade 13 during the COVID-19 pandemic. In particular, we examined reading and mathematics outcomes of Italian students, in order to evaluate the effectiveness of distance learning mechanisms utilised during school closures.

Overall results show that during the pandemic the loss of skills recorded in the 2020/2021 school year is evident throughout the territory. Using different data sources, we computed and analyzed indicators that provide an integrated picture of the situation and of the regions that have had greater difficulty than others in ensuring skills acquisition during the COVID-19 pandemic. Distance teaching mixed results among different regions are also evident and may be explained by various factors such as disparities in access to technology, differences in infrastructure, and varying levels of support provided by schools and families. Some regions, in particular internal territory, had limited or no access to high-speed internet. Additionally, some schools were better equipped to transition to distance learning, while others faced significant obstacles in adapting to the new mode of education. Socio-economic factors also played a crucial role, as students from disadvantaged backgrounds often lacked the necessary resources and conducive learning environments at home (especially in Calabria and Basilicata).

The analysis could be further enriched by also considering the latest data on students in the 2021/2022 school year (INVALSI, 2022; ISTAT, 2022c), since the results of the national tests conducted in INVALSI seem to confirm the decline in students' skills and the territorial differences (Fig. 6).

The difficulties of the health emergency did not substantially change the geography of learning inequalities compared to the 2018/2019 school year, but an increase of difficulties is observed for some regions in the South and Islands, partly due to the problems that schools and households have had in adjusting to the required changes. The data does not allow us to understand the underlying reasons for the difficulties that have emerged, but further investigations[4] and in-depth analysis may help uncover these reasons.

Our study may be useful for both policy and practice. For policy, it could serve as a basis for organizing coping strategies and cocrete actions to counteract to new emergencies in the future. For school practice, it highlights

[4] For example see the SIRD study here: https://www.sird.it/ricerca-nazionale-sird-2020/"; or Istat Statistical Report "I ragazzi e la pandemia: vita quotidiana a distanza", https://www.istat.it/it/files//2022/05/REPORT_ALUNNI-SCUOLE-SECONDARIE_2021_2.pdf.

the need for investment in digitalization and teacher training, by providing schools with the necessary technological infrastructure and resources. This could help ensure that all students have equal access to quality education, regardless of their region or socioeconomic background.

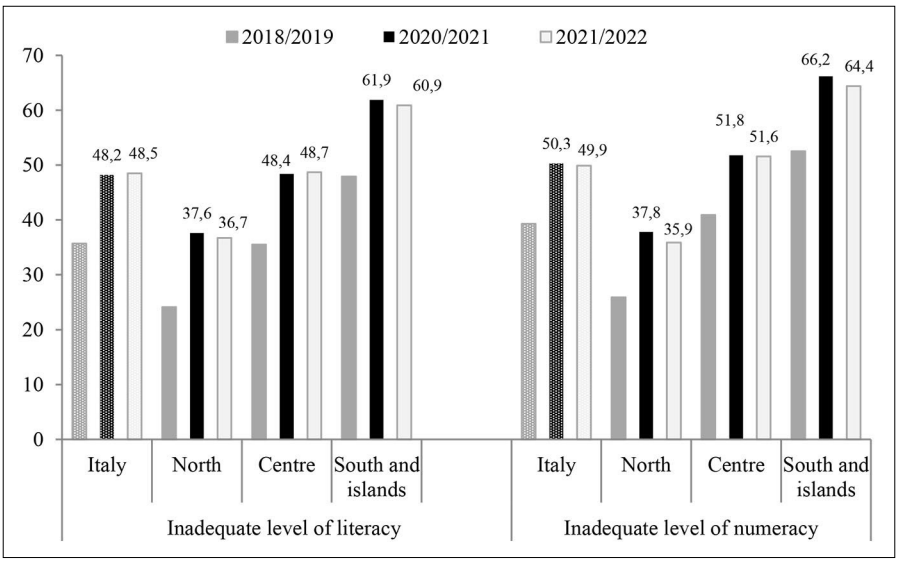


Fig. 6 – Inadequate level of Italian and Mathematics in upper secondary school (percentage). School years 2018/2019, 2020/2021 and 2021/2022

Source: INVALSI, Proficiency levels (students in grade 13)

References

Bagnarol C., Donno S., Riccardi V. (2021), "Doing school but not at school: territorial peculiarities and socio-economic inequalities in access to distance learning" (2021), in P. Falzetti (*ed.*), *Agenda 2030: Survey on Sustainable Development Goals through INVALSI Data*, FrancoAngeli, Milano.

Campodifiori E., Falzetti P., Marsili M. (2021), "Territorial inequalities in education: some aspects related to the territory", in P. Falzetti (*ed.*), *Agenda 2030: Survey on Sustainable Development Goals through INVALSI Data*, FrancoAngeli, Milano.

Gismondi R., Russo M.A. (2004), "Definizione e calcolo di un indice territoriale di turisticità: un approccio statistico multivariato", *Statistica*, 3, pp. 545-571.

INVALSI (2022), *Rapporto INVALSI 2022*, retrieved on April 7, 2023 from: https://invalsi-areaprove.cineca.it/docs/2022/Rilevazioni_Nazionali/Rapporto/Rapporto_Prove_INVALSI_2022.pdf.

INVALSI (2021), *I risultati delle Prove INVALSI 2021*, retrieved on April 7, 2023 from: https://www.invalsiopen.it/risultati/risultati-prove-invalsi-2021/.
ISTAT (2019), *Cittadini e ICT, Statistica Report*, retrieved on March 7, 2025, from: https://www.ISTAT.it/wp-content/uploads/2019/12/Cittadini-e-ICT-2019.pdf.
ISTAT (2020), *L'inclusione a distanza degli alunni con disabilità, Statistica Report*, https://www.ISTAT.it/wp-content/uploads/2020/12/Report-alunni-con-disabil-it%C3%A0.pdf.
ISTAT (2022a), *Rapporto Annuale 2022 – La situazione del Paese*, retrieved on April 7, 2023 from: https://www.ISTAT.it/it/archivio/271806.
ISTAT (2022b), *2021 SDGs Report. Statistical Information for 2030 Agenda in Italy*, retrieved on April 7, 2023 from: https://www.ISTAT.it/it/files//2022/02/2021-SDGS-Report_Inglese.pdf.
ISTAT (2022c), *Rapporto SDGs 2022. Informazioni statistiche per l'Agenda 2030 in Italia*, retrieved on April 7, 2023 from: https://www.ISTAT.it/it/archivio/275718.
OECD (2023), *PISA 2022 Results*, vol. I: *The State of Learning and Equity in Education*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/53f23881-en.
UNESCO, UNICEF, World Bank (2020), *What have we learnt? Overview of findings from a survey of ministries of education on national responses to COVID-19*, https://data.unicef.org/resources/national-education-responses-to-covid19/.
Vincent-Lancrin S., Cobo Romaní C., Reimers F. (*eds.*) (2022), *How Learning Continued during the COVID-19 Pandemic: Global Lessons from Initiatives to Support Learners and Teachers*, OECD Publishing, Paris, https://doi.org/10.1787/bbeca162-en.

6. Exploring the predictive power of standardized assessments and non-standardized evaluations on Italian university freshpeople's performance

by Gabriele Lombardi, Roberta Cipriano, Giulio Ghellini

The present chapter aims to compare the ability of standardized assessments and non-standardized evaluations at the secondary education level in explaining and predicting students' performance in tertiaty education. In particular, standardized assessments are identified through INVALSI scores, while non-standardized evaluations through high schools' final marks. Thanks to data drawn from the Italian national registry of tertiary education (Anagrafe nazionale della formazione superiore) merged with INVALSI data for the cohort subsequently enrolled in the Italian tertiary education system in the academic year 2019/2020, we perform different mixed-effect logit models with random intercepts, clustering students by high schools' regions, in order to control for the opposite geographical results which emerge in Italy between INVALSI scores and final exams' marks. Moreover, predictive ability of these two different kinds of evaluations is tested through a Naïve Bayes algorithm. Our analysis shows that non-standardized evaluations are slightly preferable both in terms of goodness-of-fit and predictive power for individual assessment. Nonetheless, our estimation also highlights the invaluable role that standardized assessments play for evaluating the overall Italian education system.

Il presente capitolo vuole confrontare la capacità delle valutazioni standardizzate e non standardizzate a livello di istruzione secondaria nello spiegare e predire la performance universitaria degli studenti. In particolare, le valutazioni standardizzate sono identificate tramite i punteggi INVALSI, mentre le non-standardizzate tramite il voto finale degli esami di maturità. Grazie ai dati estratti dall'Anagrafe nazionale studenti italiana, collegati ai dati INVALSI per la coorte di studenti immatricolati nelle università italiane nell'anno accademico 2019/2020, andiamo a stimare diversi modelli logit ad effetti misti, con intercette random che raggruppano gli studenti in

base alle regioni in cui si sono diplomati, così da controllare anche per i risultati contrastanti che emergono a livello geografico in Italia tra punteggi INVALSI e voti di diploma. Inoltre, la capacità predittiva di queste due distinte tipologie di valutazione saranno controllate tramite un algoritmo Naive Bayes. Le nostre analisi mostreranno come le valutazioni non-standardizzate siano moderatamente preferibili sia in termini di bontà di adattamento che di capacità predittiva per la valutazione individuale. Ciononostante, le nostre analisi mettono anche in evidenza l'impagabile ruolo compiuto dal sistema di valutazione standardizzato nel raggiungere l'obiettivo di misurare lo stato di salute complessivo del sistema educativo italiano.

1. Introduction

Since 2004, the INVALSI Institute has had the role of implementing periodic tests on learning outcomes in the Italian educational system. At the secondary education level, this goal is pursued through computer-based national INVALSI tests, which are carried out yearly by each Italian student enrolled at the second and fifth year, in Italian, Mathematics and English competences. The main aim of these tests is to check if Italian students have obtained the key competences required by Italian curricular guidelines. Thus, results are computed in order to evaluate the conditions of the Italian educational system at a class, school and overall country level. Currently, the Institute does not have among its tasks that of evaluating students individually. Nonetheless, despite INVALSI's national assessments are embedded into the Italian curricular frameworks, many debates arouse in the past around the possibility of extensive use of these tests.

On the one side, scholars, such as Boeri (2014), attribute the aversion towards INVALSI tests to a strategy employed by teachers to avoid being evaluated. Among the replies, the Italian teacher Pascuzzi (2014) criticizes this hypothesis highlighting how INVALSI tests suffer for a lack of objectivity, aiming at measuring a quality (learning) through a quantity (number of correct answers). Nonetheless, the main response to this objection is that a quantitative "objective" indicator is useful at least in order to estimate the overall condition of the educational system, through the certification of a sort of "credentials reliability", controlling for the attainment of *minimum standards* without interfering with professors' freedom of teaching (Augias, 2014; Pedrizzi, 2011).

On the other side, Baccini and Latempa (2021), considering standardized assessments mainly as a political choice, have warned against the danger that INVALSI test scores could be used in the future not only for profiling

schools, but students themselves. Indeed, a document by Ichino *et al.* (2008) suggests to employ INVALSI results to allocate funds among schools depending on goals reached, to incentivize teachers, and even to inform universities about the adequacy of individual student preparation. The risk is that INVALSI tests could become a sort of pre-admission university tests, but also that southern schools, already penalized by poorer territorial conditions, would be even more stressed rather than incentivized, since their results in terms of INVALSI test scores are consistently lower than their northern counterparts (INVALSI, 2022). Nonetheless, Paglieri (2021) proposes to consider a clear distinction between *evaluation* and *measurement*. From this point of view, evaluation implies a judgment, sensitive to context and individual differences, and it has to be delegated to teachers, who have the chance of knowing their students, day by day. *Measurement* is needed for obtaining information about the effective competencies acquired, and from this point of view INVALSI assessments can play an invaluable role.

Worth noticing, the description of the debate around INVALSI's capability in assessing students had to rely mainly on opinion articles, while no scientific study was found about a comparison among them and teachers' evaluations. Nonetheless, since decades academic literature has been debating the dual role of standardized and non-standardized assessments. Indeed, if Sharpley and Edgar (1986) highlight the importance of the time spent between students and teachers engaging in teaching-learning process for the assessment of achievement levels, so Herman and Golan (1993) consider standardized assessment an important policy tool, focusing schools on important standards and valued skills. Nonetheless, the time spent in preparation of specific test administration activities is subtracted to meaningful educational activities. Accordingly, Airasian *et al.* (1977) find that standardized assessments cause teachers to alter their own perception, tending to increase students' marks. Moreover, also students' brains change and adapt responding to the new assessment environment, so that employing these kinds of measurement as a way to shortlist applicants can cause the exclusion of those with wider mindsets in favor with those more "adapted" to confront themselves with standardized tests (Nalliah and Reddy, 2022).

We would try to fill this gap in the literature by comparing fitting and predictive capabilities of different models including alternatively standardized assessments (e.g. INVALSI test scores) and non-standardized evaluations (e.g. final exams' marks) in predicting students' performance at the Tertiary Education level. The chapter is structured as follows: the second section describes data merging on Italian students in secondary education with data on the same students enrolled in tertiary education during academic year

2019-2020. In the same section we describe the empirical analytical strategy of such merged data, involving a mixed-effects logit model for controlling goodness-of-fit and a naive bayes algorithm for testing predictive capacity. Finally, a fourth section will conclude that, since non-standardized evaluations shows both better goodness-of-fit and better predictive capability, our educational system should not get rid of the irreplaceable role of teachers for students' evaluation processes. Nonetheless, INVALSI measures offer an independent view on the overall Italian education system, which should be constantly taken into account by policy makers, especially for pursuing the purpose of helping those – schools or students – who are lagging behind.

2. Data and model

This study is conducted on data drawn from the Italian national registry of tertiary education (Anagrafe nazionale della formazione superiore), and processed according to the research protocol "From high school to the job placement: analysis of university careers and university mobility from southern to northern Italy"[1] carried out by the University of Palermo (head of the research program), the Italian Ministry of Education and Merit, and INVALSI. The data preparation included the exclusion of subjects due to the presence of an enormous number of missing values throughout students' characteristics in INVALSI data, with the exception of those regarding Economic, Social and Cultural Status Index (ESCS), which were imputed through a Conditional Tree Regression, validated with a 10-fold Cross Validation (Kuipers *et al.*, 2023).

2.1. Data

The final sample includes 112,267 students who attended the last year of Secondary School in Italy during the school year 2018/2019, then deciding to enroll in an Italian university for tertiary education for academic year 2019/2020. From this point of view, selection bias emerges as a first limitation of this study, since we cannot control the high school performance of those students who decide not to enroll in university, or to enroll in a foreign academic institution.

[1] The data used in this study have been processed in accordance with the Research protocol for the study "From high school to the job placement: analysis of university careers and university mobility from Southern to Northern Italy" between the Ministry of University and Research, the Ministry of Education and Merit, the University of Palermo as the lead institution, and the INVALSI Institute. The reference researcher is Bruno Bertaccini (UNIFI).

Tab. 1 – Descriptive statistics for the full sample and by number of credits earned

Variables	*All sample*	*<=5 ECTF*	*>5 ECTF*	*<=30 ECTF*	*>30 ECTF*
Males	0.432	0.504	0.421	0.489	0.405
Mother Ed.: < HS	0.247	0.281	0.237	0.272	0.228
Mother Ed.: HS	0.421	0.423	0.421	0.421	0.421
Mother Ed.: > HS	0.336	0.296	0.343	0.308	0.350
Father Ed.: < HS	0.339	0.366	0.335	0.361	0.329
Father Ed.: HS	0.348	0.356	0.346	0.353	0.345
Father Ed.: > than HS	0.313	0.278	0.319	0.286	0.326
Individual ESCS	0.351	0.234	0.370	0.260	0.396
Non-italian resident	0.012	0.016	0.011	0.015	0.010
Scientific lyceum	0.414	0.309	0.431	0.357	0.442
Classic lyceum	0.117	0.091	0.121	0.102	0.124
Other lyceum	0.123	0.119	0.123	0.111	0.129
Not lyceum	0.346	0.480	0.325	0.429	0.306
Private HS	0.024	0.031	0.023	0.026	0.024
HS ESCS	0.238	0.139	0.254	0.169	0.271
Geo-rip. Uni: North-West	0.313	0.292	0.317	0.277	0.331
Geo-rip. Uni: North-East	0.247	0.205	0.254	0.201	0.270
Geo-rip. Uni: Centre	0.211	0.226	0.208	0.242	0.196
Geo-rip. Uni: South	0.179	0.220	0.173	0.219	0.160
Geo-rip. Uni: Islands	0.049	0.057	0.048	0.062	0.043
Private university	0.078	0.066	0.080	0.051	0.091
Mover	0.258	0.239	0.262	0.241	0.267
Field of study	N = 10	N = 10	N = 10	N = 10	N = 10
HS final mark	81.56	76.72	82.34	77.52	83.54
INVALSI Math	222.84	210.34	223.3	213.23	227.56
INVALSI Ita	221.2	208.44	224.86	211.2	226.2
HS regions	N = 19	N = 19	N = 19	N = 19	N = 19
Students	112,267	15,574	96,693	36,968	75,299

Table 1 collects the main descriptive statistics. Among them, 13.9% gains less than five credits (according to the European Credit Transfer and Accumulation System, ECTF) during the first year of university, while 67.1% earn more than thirty credits. Females are the majority (56.8%), also regarding the share of good performing freshpeople (59.9% of those obtaining more than 30 ECTF credits). Students are substantially well-balanced among parents' educational level but, as an example, the share of students with a mother without a secondary education degree increases from the 24.7% to the 28.1% when data are restricted on those freshpeople gaining less than 5 ECTF cred-

its. On the other side, the share of those with a mother holding a Tertiary Education degree decreases from the 33.6% to the 29.6%.

The same happens looking at the Individual ESCS, which defines the social, economic and cultural status of students' families participating to INVALSI tests, whose mean for those students obtaining less than 5 ECTF credits is more than 30% lower than the overall mean, while the one for those with more than 30 is the 13% higher. This result becomes even more impressive when looking at the average *ESCS* at school level. Indeed, in this case the mean for students with less than 5 credits is almost a half of the overall one, while the mean for those with more than 30 credits is 14% higher.

The 53.1% of our sample is composed of students from scientific (41.1%) and classic (11.7%) lyceums, which represent the more similar institutes to an *academic track* in the Italian secondary education system. Also in this case, the share of these students becomes higher when considering only the top performing subset. The majority of students comes from northern regions (56%) and their share is always higher looking at top performing students. Students who decide to study outside the region where they attended secondary education studies (*movers*) are the 25.8%, but they increase up to the 26.7% when considering freshpeople able to gain more than 30 ECTF credits during the first year.

Moreover, the analysis is also controlled for the field of studies chosen by each student, namely: Agriculture, forestry, fisheries and veterinary; Arts and Humanities; Business, Administration and Law; Education; Engineering, manufacturing and construction; Health and welfare; Information and Communication Technologies (ICTs); Natural sciences, mathematics and statistics; Services; Social sciences, journalism and information.

In our study, non-standardized evaluations are represented by the high school final mark, which is the result of three written exams and an oral examination, evaluated by a commission composed by 50% of professors within the school, and 50% of external professors. On average, our sample exhibits a final mark of 81.56/100, which increases up to 83.54 when considering top performing students (> 30 ECTF credits). On the other side, standardized assessments are represented by INVALSI test results in Italian language and mathematics, with an average respectively of 221.2 (226.2 for > 30 ECTF credits), and 222.84 (227.56 for >30 ECTF credits).

Other information available include: having attended a private high school (2.4%), being enrolled in a private university (7.8%), the degree courses macro-area, and the high school region.

2.2. Empirical strategy

As the main aim of this study is to investigate differences in explanatory and predictive power of standardized assessments and non-standardized evaluations, the analysis will follow three steps.

The first step will consist in running a logit model with random intercepts (Bates *et al.*, 2015). Following Hartzel *et al.* (2001), we estimate the probability that each student *i* obtains at least 30 ECTF credits during their first year at university through the function:

$$\mu_{ij} = E(y_{ij}|u_i) = h(\eta_{ij}) = \frac{exp(\eta_{ij})}{1 + exp(\eta_{ij})}, \eta_{ij} = Z_{ij}\beta + u_i, \quad (1)$$

Where u_i is the vector of the *j* random effects representing the Italian regions where each student was enrolled during Secondary Education, β are fixed effects, Z_{ij} is the matrix containing all the fixed effects covariates, and $h(\eta_{ij})$ is a vector of inverse link function. The dependent variable is represented by y_{ij}, where:

$$y_{ij} = \begin{cases} 1 \text{ if I gained at least 30 ECTF credits during the first academic year} \\ 0 \text{ otherwise} \end{cases} \quad (2)$$

The decision of employing Italian regions as random effects is needed in order to control for heterogeneity (Wagner and Duller, 2011), so admitting the possibility that having attended the high school in a specific part of the country modifies the average outcome of an individual with the same characteristics of a similar peer enrolled in a different geographic area (Agresti, 2018).

Indeed, looking at Figure 1 we can notice how INVALSI test scores are generally higher for northerners, while final marks at the Secondary Education level are higher for southerners. It is important to notice that, for the sake of brevity, Figure 1 only includes scores in Mathematics for the INVALSI tests. Nonetheless, the picture resulting from scores in Italian language is quite similar. Of course, it is also important to recall that INVALSI scores refer to a very specific subject, while graduation mark is the result of an extensive exam (both written and oral) including in one score the assessment of a wide range of subjects. Interestingly, INVALSI tests have a similar distribution to that of high performing students in their first year of university (> 30 credits), while graduation marks are more similar to the distribution of students that do not perform very well in their first year at university (< 5 credits).

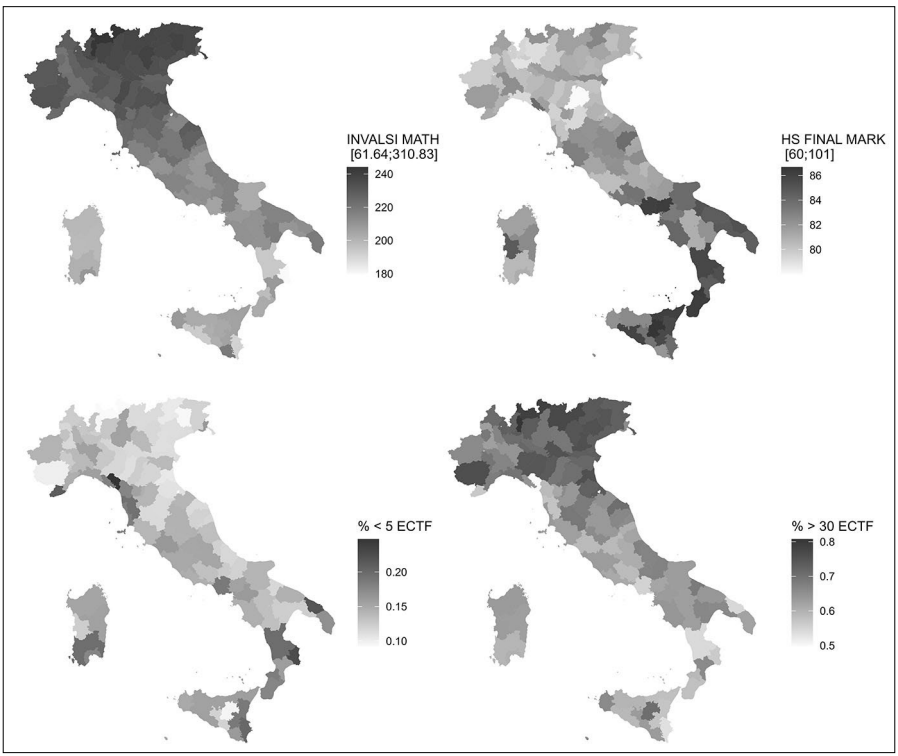


Fig. 1 – Sample distribution by region of INVALSI test scores in mathematics, high school final mark, share of students gaining less than 5 ECTFs or more than 30

The second step of our empirical strategy regards the evaluation of the fitting of different models. The three main regressions to be evaluated will concern: i) a model which does not include any kind of students' assessments as fixed effects; ii) a model including only standardized assessments; iii) a model including only non-standardized evaluations. In particular, these models will be compared in terms of: information criteria (AIC and BIC), which evaluate each model taking into account both its fitting and complexity (Liddle, 2007); Marginal and Conditional R^2, where the first attempts to measure variance explained by fixed effects, while the second by both fixed and random effects (Nakagawa and Schielzeth, 2013); finally, the *percentage correctly predicted*, which considers as correctly predicted an observation if for a high performer a probability is estimated of at least 0.5, or less than 0.5 for a low performer (Herron, 1999). The same indicator will be tried changing the threshold at 0.66 for high and 0.33 for low performers.

Finally, as a third step, predictive capability of the two types of grading (INVALSI assessments and teachers' evaluations) will be checked through a Naïve Bayes estimation (Meyer *et al.*, 2019), the simplest way of applying Bayes' theorem on classification problems, even if relying on the very strong assumption that all the included characteristics are independent and equally important (Lantz, 2019). For each model classification, we will measure accuracy (i.e. % of correctly predicted observations) and we will validate it by means of a 10-fold cross validation.

As a robustness check, all these steps will be replicated for a different dependent variable, namely:

$$y_{ij} = \begin{cases} 1 \text{ if I gained at least 5 ECTF credits during the first academic year} \\ 0 \text{ otherwise} \end{cases} \quad (3)$$

3. Results

3.1. Mixed-effects logit estimates

Figure 2 shows random intercepts by region for four different models. In the first, no fixed effect is included. As it can be seen, six over seven among those regions exhibiting a statistically significant positive intercept (Calabria, Campania, Sardinia, Sicily, Tuscany, and Lazio) are from the Centre-South part of the country, while five over six ones (Piedmont, Lombardy, Friuli Venezia Giulia, Trentino Alto Adige, Emilia Romagna, and Veneto) with a statistically significant negative intercept are based in the North. From this point of view, it should be argued that high performers at the Tertiary Education level come in general from the South of Italy. If this were true, it should be also concluded that in some way INVALSI tests underestimate southerners' performance, while final exam's commissions are abler to genuinely evaluate them. The inclusion of fixed effects somehow overturns regions' ranking especially at its bottom, where now only regions based at the Centre-South appear. The interpretation is that a student from Liguria who shares the same characteristics of a student from Molise has a higher probability of gaining more than 30 ECTF credits during its first university year (Agresti, 2018). Nonetheless, six over eight intercepts not significantly different from zero pertain to northern regions, making the suspect arise that differences are almost present only among regions of Centre/South. The order of the random intercepts changes again dramatically when final high school mark is included among fixed effects. Six over eight positively signif-

icant intercepts pertain to northern regions (Lombardy, Trentino Alto Adige, Veneto, Friuli Venezia Giulia, Emilia Romagna, and Piedmont), while all the negatively significant are in the Centre-South (Lazio, Apulia, Campania, Sicily, and Calabria). This latter result increases to six regions (Campania, Abruzzo, Marche, Apulia, Molise and Basilicata) when including INVALSI results, while seven over nine positively significant intercepts belong to northern regions (Liguria, Trentino Alto Adige, Lombardy, Friuli Venezia Giulia, Veneto, Piedmont, and Emilia Romagna). In a nutshell, as the inclusion of both standardized assessments and non-standardized evaluations predicts a higher probability of performing better during the first year of enrolment for the northerners, so this result is more evident when controlling for the former kind of assessment.

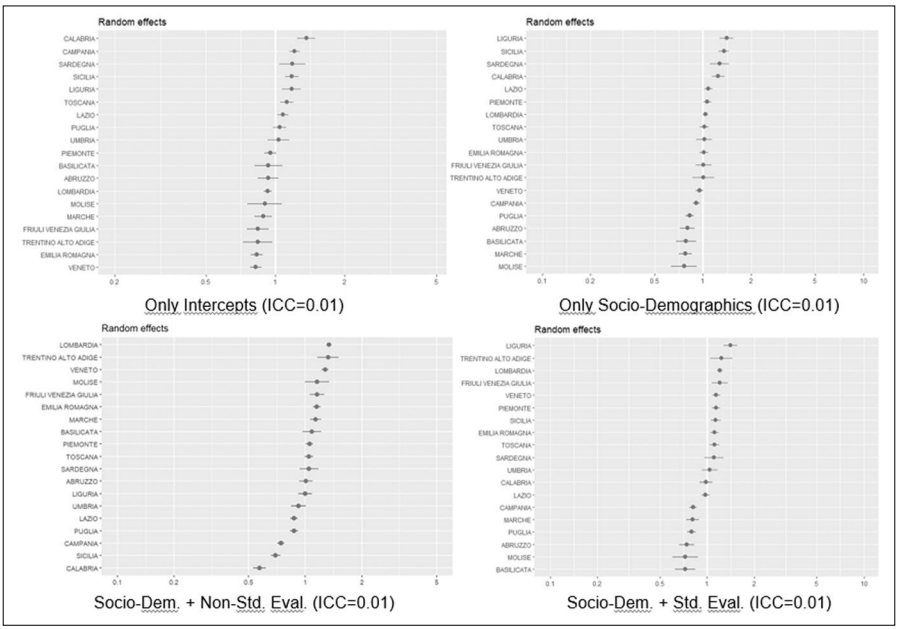


Fig. 2 – Plot of the random intercepts (HS regions) by four different model specifications for the probability of obtaining at least 30 ECTF credits during the first university year

Tab. 2 – Mixed-effect logit models for the probability of obtaining at least 30 ECTF credits during the first university year

	Socio-dem.		*SD + Non-std. eval.*		*SD + Std. eval.*	
Dep. Var.: >30 ECTF	*OR*	*SE*	*OR*	*SE*	*OR*	*SE*
(Intercept)	0.85	(0.06)	0.88(-)	(0.07)	1.28	(0.05)
Gender: Male	0.76	(0.01)	0.88	(0.02)	0.67	(0.02)
Parent ed. ref: non-completed HS						
Mother: HS	1.10	(0.04)	1.05(-)	(0.04)	1.05(-)	(0.04)
Mother: > HS	1.09	(0.02)	1.09	(0.03)	1.10	(0.03)
Father: HS	0.97(-)	(0.04)	0.98(-)	(0.03)	1.00(-)	(0.04)
Father: > HS	0.99(-)	(0.02)	1.02(-)	(0.04)	1.04(-)	(0.02)
Individual ESCS	1.03	(0.01)	1.01(-)	(0.01)	1.02(-)	(0.01)
Non-Italian Res.	0.64	(0.06)	0.69	(0.06)	0.68	(0.06)
HS ESCS	1.13	(0.01)	1.16	(0.01)	1.08	(0.01)
Private HS	0.58	(0.04)	1.24	(0.05)	0.72	(0.05)
HS Ref.: Non lyceum						
Scientific lyceum	1.93	(0.02)	2.28	(0.02)	1.34	(0.02)
Classic lyceum	1.46	(0.03)	1.34	(0.03)	1.19	(0.03)
Other lyceum	1.29	(0.02)	1.28	(0.03)	1.22	(0.02)
Private university	1.26	(0.03)	1.24	(0.03)	1.31	(0.03)
Mover	1.10	(0.02)	1.02(-)	(0.02)	1.08	(0.02)
Geo-rip. uni. ref: Centre						
North-East	1.65	(0.04)	1.62	(0.04)	1.49	(0.04)
North-West	1.45	(0.04)	1.33	(0.04)	1.25	(0.04)
South	0.99(-)	(0.04)	1.00(-)	(0.04)	1.07	(0.03)
Islands	0.98(-)	(0.06)	0.93(-)	(0.06)	1.05(-)	(0.05)
Field of study	Yes		Yes		Yes	
HS final mark			2.04	(0.01)		
INVALSI Mat					1.59	(0.01)
INVALSI Ita					1.20	(0.01)
Random effects						
HS Regions	19		19		19	
Observations						
Students	112,267		112,267		112,267	

(-): p-value > 0.1

Table 2 shows detailed estimations for the three models which include fixed-effects. Results are substantially consistent among estimations. Females are more likely of being good performers than males, but when not

standardized evaluations are added females exhibit only one fifth chance more of success, against the one third estimated employing standardized assessments. Fathers' education level shows no effect, while those who have mothers with tertiary education level are consistently more likely to be high performers. Surprisingly, individual ESCS shows a very small association with success for the model without evaluations, while no effect arises for the other two. Non-Italian residents have about one third chances less than natives of being high performers. The ESCS at school level seems to have a very robust positive and significant association with high performance at university level. The highest association emerges for the model including non-standardized evaluations, while the lowest for the model with INVALSI scores. Apparently, not only the fact of having a high socio-economic background, but also being surrounded by a large number of peers with high ESCS scores positively influences the likelihood of becoming a high performer at tertiary education level.

As expected, students from lyceums have a higher probability of becoming high performers. In particular, the biggest difference emerges regarding scientific lyceums between the two models: in the model with high school final marks, students have twice the chances of success compared to their non-lyceum peers, while the model with INVALSI scores estimates only one-third more chances. Surprisingly, as the models with and without standardized assessments show a negative association (OR < 1) between private high schools and university performance, so the model with non-standardized evaluations estimates a one-fourth greater chance of being high performers for this category of freshpeople. Another difference emerges regarding movers, namely those students who enroll for a university based in a different region than their high school one. Indeed, the model with final exam's grade is the only one showing no association, while the other two estimate about one-tenth greater chances of success compared to stayers. Consistently among the three models, freshpeople who decide to enroll in a northern university have more chances to be high performers. Consistent results also emerge when looking at different types of competence measurement: in both cases, students with higher marks in their final high school examination or with higher test scores are more likely to obtain more than 30 ECTF credits during their first year in higher education.

Tab. 3 – Summary of diagnostics for different mixed-effect logit models for the probability of obtaining at least 30 ECTF credits during the first university year

	AIC	*BIC*	*Marg. R2*	*Cond. R2*	*% Corr. Pred. (Th. = 0.5)*	*% Corr. Pred. (Th. < 0.33, > 0.66)*
Only intercept	141,029.29	141,048.55	0.000	0.016	67.07	39.63
Only socio-demographics	133,748.47	134,027.70	0.097	0.101	68.70	45.39
Socio-demog. + Non- std. eval.	124,243.17	124,532.03	0.205	0.217	72.07	50.86
Socio-demog. + Std. eval.	129,018.78	1293,17.27	0.165	0.167	70.55	47.73
All variables	122,881.14	123,208.51	0.228	0.232	72.49	51.50

Table 3 compares the goodness-of-fit of five different models: i) a model with only random intercepts; ii) a model with only socio-demographic covariates; iii) a model with sociodemographic covariates and non-standardized evaluations; iv) a model with sociodemographic covariates and standardized assessments; v) a model with sociodemographic covariates and both non-standardized evaluations and standardized assessments. As it can be seen, the two models including non-standardized evaluations are the best-performing both in terms of information criteria, which exhibit the lowest scores, and in terms of marginal and conditional R^2, which shows the highest values. Moreover, both the percentage correctly predicted at the thresholds of one half and one third is higher when considering non-standardized evaluations. Namely, the two models including final exam's mark are the only ones in which the correctly predicted observations are more than the 72% in the first case, and more than the 50% in the second case.

3.2. Naïve Bayes predictions

Tab. 4 – Confusion Matrix summarizing Naïve Bayes algorithm's predictions for different subsets of characteristics for the probability of obtaining at least 30 ECTF credits during the first university year (%)

	TLP	*FLP*	*THP*	*FHP*	*10-Fold CV (Accuracy)*
Only Non-std. evaluation	4.1	3.2	63.9	28.9	67.97
Only Std. assessments	7.3	7.4	59.7	25.7	66.94
Only Socio-demographics	7.8	7.4	59.7	25.1	67.54
Socio-demographics + Std. eval.	11.6	10.7	56.4	21.3	68.04
Socio-demographics + Non-std. eval.	10.1	7.0	60.0	22.9	70.09

THP = "True High Performers"; FHP = "False High Performers"; TLP = "True Low Performers"; FLP = "False Low Performers"; CV = "Cross Validation".

Finally, we can control the predictive capability of five different models, estimated through a Naïve Bayes algorithm. The confusion matrix synthetized in Table 4 highlights that when a set of characteristics also includes non-standardized evaluations, it exhibits an accuarcy at least 1% higher than the models with standardized assessments. Interestingly, standardized assessments are better in estimating True Low Performers (TLP), while non-standardized evaluations better predict True High Performers (THP). From this point of view, the decision about which kind of evaluation tool should be preferred implies to take another fundamental choice about what is the *Type*

I error in this framework. In other words, is it more serious to label a good student as a bad one, or a bad one as a good one? A deeper discussion of this issue is postponed to section 4.

3.3. Robustness check

As a robustness check, the independent variable was modified in order to estimate the probability of being a strongly bad performing student, namely, of not being able to obtain at least 5 ECTF credits during the first university academic year. For the sake of simplicity, results are described here and they can be checked in Tables 5, 6, and 7 in Appendix A. Results are strongly robust both in terms of fixed effects (Table 5) and goodness-of-fit measures (Table 6), where models with non-standardized evaluations are always better than the other, even if marginally, as in the case of *percentage correctly predicted*. Also Naïve Bayes predictions (Table 7) are consistent, even if not so much reliable: indeed, choosing a so extreme threshold in the dependent variable have caused the algorithm to label almost always our students as low performers. Nonetheless, some interesting differences emerge when looking at Figure 3. In fact, it is possible to observe a change in the order of the random intercepts much stronger than the one highlighted from the main analysis. Indeed, as the model with INVALSI scores consider students from northern regions more likely to be high performers, something different happens in the model with final exam's marks. Namely, northern regions are highly concentrated in the middle part of the graph, where zero effect intercepts are estimated. On the other side, southern regions are more or less separated between the top and the bottom of the plot. Our interpretation of this subgraph suggests that the model distinguishes between southern regions that are less likely to provide low-performing students to the higher education system (e.g., Molise, Basilicata, Abruzzo, and Marche), which exhibit negative intercepts, and those that may inflate final exam marks, as suggested by critics of non-standardized evaluations. In the latter case, the model estimates a positive intercept to mitigate the overall negative effect of high final marks on the probability of being a low-performing student. Among the other southern regions, Puglia and Campania show non-significant intercepts, indicating non statistical differences from their central-northern counterparts.

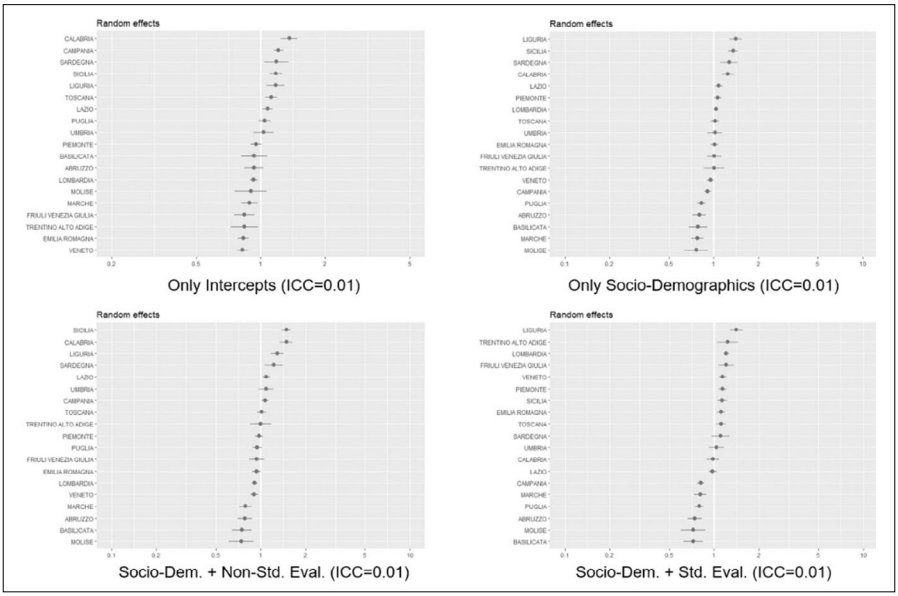


Fig. 3 – Plot of the random intercepts (HS regions) by four different model specifications for the probability of obtaining at most 5 ECTF credits during the first university year

4. Discussion and conclusions

The present chapter tried to compare standardized assessments with INVALSI test scores and non-standardized evaluations based on high school final examination marks, from the point of view both of data adaptability and predictive power.

The main result is that the final mark at the secondary education level is consistently better than INVALSI scores in predicting who will be a good student at the tertiary education level. Coherently with other findings in the academic literature (see, for example: Sharpley and Edgar, 1986; Hoge and Coladarci, 1989) this result can be attributed to the fact that teachers have the opportunity of evaluating students' competences not only on the basis of performance, but also thanks to the knowledge of the deep individual mindset learned through a continuous process of classroom activities. Indeed, according to ministerial decree 62/2017[2], the final exam mark also considers

[2] Retrieved on 9 february, 2025, from https://www.gazzettaufficiale.it/eli/id/2017/05/16/17G00070/sq.

teachers' judgments about the last three years of a student's secondary school career. On the other side, INVALSI tests can only take a static picture aimed at verifying the acquisition of basic skills required by students in the Italian education system. Nonetheless, this evidence does not make INVALSI tests a useless nor dangerous tool. On the contrary, our analysis clearly shows what a fundamental role INVALSI tests can play in shedding light on inequalities and stratification problems in the Italian secondary education system. As an example, the only fact that high ESCS at school level are so consistently associated with a good individual performance makes us suppose that Italy suffers for the presence of schools *for the rich*, which guarantee better future opportunities to their own students, and schools *for the poor*, whose students remain locked in their current status quo.

Coherently with the opinion by Paglieri (2021), from this point of view standardized assessments and non-standardized evaluations emerge as two different and indispensable instruments, which play two different roles, anyway. Therefore, while evaluation should remain a teacher's prerogative, INVALSI should continue to assess the overall state of the system without encroaching on individual evaluations. One possible consequence of this potential situation can already be seen when examining random intercepts. A significant disparity in INVALSI scores emerges between the North and South, with the North consistently achieving higher results. This leads to an underestimation of southern students' performance when standardized assessments are used, particularly in regions like Molise, Marche, and Basilicata, which may weaken the predictive power of these assessments. Conversely, when using non-standardized evaluations, we can identify certain southern regions, such as Sicily, Calabria, and Campania, that appear to have lower chances of producing high-performing students, assuming all other factors are equal.

Moreover, it is important to notice the trade-off which emerges looking at predictions obtained through standardized assessments and non-standardized evaluations: in the first case, the algorithm was more accurate in predicting low performers, while in the second case in predicting high performers. This issue raises a non-trivial question: is it worse to label a high performer as low, or the opposite? From this perspective, individual standardized assessments can be valuable for universities in identifying students who need remedial support. However, they can also be problematic if misused as a barrier to accessing higher education. Indeed, if the Tertiary Education still aims at being *universal* and inclusive, our results suggest to keep trusting teachers and their day-by-day knowledge of students' mindsets, leaving to standardized assessments the measurement role of the overall education system con-

ditions. Accordingly, our analysis reveals that the majority of the variability in our outcome is attributable to contextual factors, such as regional and socio-economic or cultural backgrounds. From this perspective, the role of INVALSI measurements serves as a valuable tool for identifying regions or categories that require policy interventions. Furthermore, incorporating both standardized assessments and non-standardized results enhances the accuracy of our predictions concerning students attaining at least 30 ECTF credits in their first year of university, based on socio-demographics, with improvements of only 0.5% and 2.55%, respectively.

Further research is needed to determine whether these estimates can be made more precise. It can be argued that actions should be taken to refine student assessment and avaluation through both standardized and non-standardized tools. However, it's important to remember that a student's profile encompasses a complex combination of personal, social, and contextual factors that can be difficult to fully capture. For instance, D'Agostino *et al.* (2021, 2022) observed that non-local students experience a decline in performance during their first year of university due to simultaneous changes in institutional, geographical, and social contexts, but they tend to recover in the second year. Consequently, the precision of estimating students' abilities will always be influenced by various exogenous factors. Nonetheless, this does not imply that evaluation and assessment should be neglected; rather, they should be employed thoughtfully and in combination.

Thus, scope of policy makers should be to provide teachers all the instruments they need in order to adequately do their job and consequently their evaluations, at the same time looking at INVALSI results in order to identify the main crucial areas where to intervene. As an example, schools who turn out to be clusters of students divided by their socio-economic background, or operating in the poorest territories of the country are much more easily detectable thanks to INVALSI data. At the same time, the INVALSI Institute needs to go on with its invaluable measurement role, with the aim of constantly improving the quality of its data collection, and the reduction of the missing information especially in those geographical clusters which exhibit the greatest difficulties in satisfying the data quality required by this kind of investigation.

Appendix: 5 ECTF credits

Tab. 5 – Mixed-effect logit models for the probability of obtaining at most 5 ECTF credits during the first university year

	Socio-dem.		*S.D + Non-std. eval.*		*S.D + Std. eval.*	
Dep. var.: < ECTF	*OR*	*SE*	*OR*	*SE*	*OR*	*SE*
(Intercept)	0.26	(0.08)	0.23	(0.08)	0.17	(0.08)
Gender: Male	1.35	(0.02)	1.19	(0.02)	1.50	(0.02)
Parent ed. ref: non-completed HS						
Mother: high school	0.93(-)	(0.05)	0.97(-)	(0.05)	0.97(-)	(0.05)
Mother: university	0.93	(0.03)	0.93	(0.03)	0.92	(0.03)
Father: high school	1.03(-)	(0.05)	1.01(-)	(0.05)	0.99(-)	(0.05)
Father: university	1.03(-)	(0.05)	1.00(-)	(0.03)	0.99(-)	(0.03)
Individual ESCS	0.97	(0.01)	0.98(-)	(0.01)	0.98(-)	(0.01)
Non-italian resident	1.34	(0.07)	1.24	(0.08)	1.27	(0.07)
High School ESCS	0.87	(0.01)	0.85	(0.01)	0.91	(0.01)
Private HS	1.89	(0.06)	1.95	(0.06)	1.56	(0.06)
HS ref.: Non-lyceum						
Scientific lyceum	0.48	(0.02)	0.43	(0.03)	0.66	(0.03)
Classic lyceum	0.62	(0.04)	0.68	(0.04)	0.74	(0.04)
Other lyceum	0.81	(0.03)	0.83	(0.03)	0.86	(0.03)
Private university	1.07(-)	(0.04)	1.10	(0.04)	1.04(-)	(0.04)
Mover	0.95	(0.02)	1.01(-)	(0.02)	0.97(-)	(0.02)
Geo-rip. Uni. Ref: Centre						
North-East	0.70	(0.05)	0.73	(0.05)	0.76	(0.05)
North-West	0.77	(0.05)	0.82	(0.05)	0.86	(0.05)
South	1.19	(0.05)	1.21	(0.05)	1.14	(0.05)
Islands	0.81	(0.08)	0.85	(0.08)	0.77	(0.07)
Field of study	Yes		Yes		Yes	
HS Final Mark			0.55	(0.01)		
INVALSI Math					0.67	(0.01)
INVALSI Lit					0.86	(0.01)
Random effects						
HS regions	19		19		19	
Observations						
Students	112,267		112,267		112,267	

(-): p-value > 0.1

Tab. 6 – Summary of diagnostics for different mixed-effect logit models for the probability of obtaining at most 5 ECTFs during the first university year

	AIC	*BIC*	*Marg. R2*	*Cond. R2*	*% Corr. pred. (Th. = 0.5)*	*% Corr. pred. (Th. < 0.33, > 0.66)*
Only Intercept	90,179.10	90,198.36	0.000	0.007	86.13	86.13
Only socio-demographics	86,362.63	86,641.86	0.087	0.097	86.12	85.13
Socio-demog. + Non-std. eval.	82,493.01	82,781.87	0.164	0.175	86.18	83.10
Socio-demog. + Std. eval.	84,253.05	84,551.54	0.140	0.151	86.13	84.02
All variables	81,845.42	82,172.79	0.182	0.189	86.20	82.87

Tab. 7 – Confusion Matrix summarizing Naïve Bayes algorithm's predictions for different subsets of characteristics for the probability of obtaining at least 30 ECTF credits during the first university year

	THP	*FHP*	*TLP*	*FLP*	*10-Fold CV (Accuracy)*
Only Non-std. eval.	86.1%	13.9%	0.0%	0.0%	86.12%
Only Std. eval.	85.7%	13.7%	0.2%	0.5%	85.87%
Only Socio-demographics	82.8%	13.1%	0.8%	3.3%	83.63%
Socio-demographics + Std. eval.	80.8%	12.0%	1.9%	5.4%	82.62%
Socio-demographics + Non-std. eval.	82.7%	12.9%	1.0%	3.4%	83.72%

THP = "True High Performers"; FHP = "False High Performers"; TLP = "True Low Performers"; FLP = "False Low Performers"; CV = "Cross Validation".

Acknowledgments

We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGenerationEU– Project Title From high school to university: Assessing peers' influence in educational inequalities and performances – CUP F53D23006150006 – Grant Assignment Decree No. 1060 adopted on 07/17/2023 by the Italian Ministry of Ministry of University and Research (MUR).

In addition, we acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2/2/2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGenerationEU– Project Title Stem in Higher Education & Women INequalitieS [SHE WINS], CUP I53D23004810006, Grant Assignment Decree No. 1060 adopted on 7/17/2023 by the Italian Ministry of Ministry of University and Research (MUR).

References

Agresti A. (2018), *Statistical methods for the social sciences*, Pearson Education, Harlow.

Airasian P.W., Kellaghan T., Madaus G.F., Pedulla J.J. (1977), "Proportion and direction of teacher rating changes of pupils' progress attributable to standardized test information", *Journal of Educational Psychology*, 69 (6), pp. 702-709.

Augias C. (2014), *La risposta di Corrado Augias. Gilda degli insegnanti Federazione Gilda-Unams*, retrieved on April 7, 2023 from: https://win.gildavenezia.it/docs/Archivio/2014/giu20014/augias-boeri.htm.

Baccini A., Latempa R. (2021), "A che cosa servono davvero i test INVALSI", *Il Mulino*, retrieved on April 7, 2023 from: https://www.rivistailmulino.it/a/a-che-cosa-servono-davvero-i-test-invalsi.

Bates D., Maechler M., Bolker B., Walker S. (2015), "Fitting Linear Mixed-Effects Models Using lme4", *Journal of Statistical Software*, 67 (1), pp. 1-48.

Boeri T. (2014), "Scuola, la rivolta contro i quiz", *la Repubblica*, 15/5, retrieved on April 7, 2023 from: https://ricerca.repubblica.it/repubblica/archivio/repubblica/2014/05/15/scuola-la-rivolta-contro-i-quiz31.html?ref=search.

Checchi D., Ichino A., Vittadini G. (2008), *Un sistema di misurazione degli apprendimenti per la valutazione delle scuole: finalità e aspetti metodologici. Proposta preparata per l'INVALSI. 04/12/2008*, retrieved on April 7, 2023 from: https://www.pietroichino.it/wp-content/uploads/2008/12/invalsi_pro_11.pdf.

D'Agostino A., Ghellini G., Lombardi G. (2021), "Movers and stayers in STEM enrollment in Italy: who performs better?", *Genus*, 77, pp. 1-19.
D'Agostino A., Ghellini G., Lombardi G. (2022), "Gender effect at the beginning of higher education careers in STEM studies: Does female recover better than male?", in *Methodology for Multilevel Modeling in Educational Research: Concepts and Applications*, Springer Singapore, Singapore, pp. 363-382.
Hartzel J., Agresti A., Caffo B. (2001), "Multinomial logit random effects models", *Statistical Modelling*, 1 (2), pp. 81-102.
Herman J.L., Golan S. (1993), "The effects of standardized testing on teaching and schools", *Educational Measurement: Issues and Practice*, 12 (4), pp. 20-25.
Herron M. (1999), "Postestimation uncertainty in limited dependent variable models", *Political Analysis*, 8, pp. 83-98.
Hoge R.D., Coladarci T. (1989), "Teacher-based judgments of academic achievement: A review of literature", *Review of Educational Research*, 59 (3), pp. 297-313.
INVALSI (2022), *Rapporto INVALSI 2022*, retrieved on April 7, 2023 from: https://invalsi-areaprove.cineca.it/docs/2022/Rilevazioni_Nazionali/Rapporto/Rapporto_Prove_INVALSI_2022.pdf.
Kuipers M.A., Kang K., Dragomir A.D., Monshouwer K., Benedetti E., Lombardi G., Luta G., Kunst A.E. (2023), "A novel methodological approach to measure linear trends in health inequalities: proof-of-concept for adolescent smoking in Europe", *American Journal of Epidemiology*, 192 (6), pp. 963-971.
Lantz B. (2019), *Machine learning with R: expert techniques for predictive modeling*, Packt publishing ltd., Birmingham.
Liddle A.R. (2007), "Information criteria for astrophysical model selection", *Monthly Notices of the Royal Astronomical Society: Letters*, 377 (1), L74-L78.
Meyer D., Dimitriadou E., Hornik K., Weingessel A., Leisch F. (2024), *Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-16*, https://CRAN.R-project.org/package=e1071.
Nakagawa S., Schielzeth H. (2013), "A general and simple method for obtaining R2 from generalized linear mixed-effects models", *Methods in Ecology and Evolution*, 4 (2), pp. 133-142.
Nalliah R.P., Reddy M. (2022), "When standardized testing is not standardized", *Journal of Dental Education*, 86 (7), pp. 823-829.
Paglieri F. (2021), "La polemica sui test INVALSI è vecchia quanto i test stessi, e non accenna a placarsi. La decisione di misurare le competenze degli studenti è una scelta politica, ma sensata ed encomiabile", *Il Mulino*, retrieved on April 7, 2023 from: https://www.rivistailmulino.it/a/chi-ha-paura-della-standardizzazione-ancora-sui-test-invalsi.
Pascuzzi V. (2014), *Ancora sulla non-oggettività dei test INVALSI. Da Corrado Augias a Tito Boeri. Gilda degli insegnanti Federazione Gilda-Unams*, 1/6/2014, retrieved on April 7, 2023 from: https://win.gildavenezia.it/docs/Archivio/2014/giu20014/augias-boeri.htm.

Pedrizzi T. (2011), "Ora 'usiamo' INVALSI e PISA per certificare le competenze", *IlSussidiario.net*, 12/5, retrieved on April 7, 2023 from: https://www.ilsussidiario.net/news/educazione/2011/5/12/scuola-ora-usiamo-invalsi-e-pisa-per-certificare-le-competenze/176294/.
Sharpley C.F., Edgar E. (1986), "Teachers' ratings vs standardized tests: An empirical investigation of agreement between two indices of achievement", *Psychology in the Schools*, 23 (1), pp. 106-111.
Wagner H., Duller C. (2012), "Bayesian model selection for logistic regression models with random intercept", *Computational Statistics & Data Analysis*, 56 (5), pp. 1256-1274.

7. Measuring early school leaving at a sub-regional territorial detail: working hypotheses

by Massimo Armenise, Barbara Baldazzi[1]

Early school leaving is an obstacle to economic growth and employment. It lowers productivity and competitiveness and increases poverty and social exclusion. According to ISTAT, the percentage of young people aged between 18 and 24 who have left education and training prematurely (with only secondary education) is 12.7% in Italy in 2021. Furthermore, one of the usual characteristics of our country is to have large internal disparities. The strong disparities in drop-out rates that appear in certain geographical areas of the country could indicate specific structural problems. If this evidence is combined with the need to implement policies aimed at reducing territorial disparities, it explains the importance of refining the level of territorial analysis as much as possible, in order to identify and implement specific measures for the territories and schools most affected by this phenomenon. The aim of this work is to develop an indicator of early school leaving that is capable of going down to an increasingly detailed territorial level in order to show the extent to which the Italian territories are differently affected by this phenomenon and which are the most exposed.

L'abbandono scolastico rappresenta di fatto un ostacolo per la crescita economica e l'occupazione, frenando la produttività e la competitività e alimentando povertà ed esclusione sociale. Secondo l'ISTAT, la percentuale di giovani fra i 18 e i 24 anni che hanno abbandonato prematuramente gli studi e la formazione (conseguendo solo la licenza della scuola secondaria di primo grado) risulta essere in Italia pari al 12,7% nel 2021. Inoltre, una

[1] This chapter is the output of a joint work of the authors. Nonetheless, M. Armenise contributes substantially to section 2 and 3 and B. Baldazzi contributes substantially to section 1 and 4.

delle consuete caratteristiche del nostro paese risulta essere quella di avere ampi divari al proprio interno. Le forti disparità nei tassi di abbandono scolastico, che emergono in determinate aree geografiche del Paese, potrebbero indicare problemi strutturali specifici e se tale evidenza la si coniuga con la necessità di attuare politiche volte all'attenuazione delle differenze territoriali, si spiega dunque quanto sia rilevante l'esigenza di raffinare il più possibile il livello di analisi territoriale, così da poter distinguere e avviare misure specifiche per quei territori e per quelle scuole maggiormente colpiti da tale fenomeno. Obiettivo di questo lavoro è quello di elaborare un indicatore dell'abbandono scolastico, capace di scendere ad un dettaglio territoriale sempre più spinto così da mostrare quanto i territori italiani siano differentemente danneggiati da questo fenomeno e quali siano con precisione quelli maggiormente esposti.

1. Introduction

School dropout, is an obstacle to economic growth and employment, holding back productivity and economic competitiveness and fueling poverty and social exclusion. In a country like Italy, which is characterized by a substantial demographic decrease (ISTAT, 2022a) and slow labor productivity growth (Bugamelli and Lotti, 2018), even more pronounced in comparison with other European countries, the loss of human capital due to early school-leaving is even more likely to increase social inequalities and erode territorial economic cohesion. Moreover, the particular geographical concentration of the phenomenon, that is evident through data on early school-leaving in some regions, risks further undermining territorial cohesion, jeopardizing any possibility of future economic convergence (ISTAT, 2022b). According to the Italian national Institute of statistics (ISTAT), in Italy, the percentage of young people between the ages of eighteen and twenty-four who are early leavers from education and training turns out to be 12.7 percent (year 2021). Despite the considerable progress the country has made on this phenomenon, this share remains among the highest of European Union countries (Figure 1), lower only than that of Spain and Romania (with percentages of 13.3 and 15.3, respectively), well above the targets set in the Community framework (Council Resolution on a strategic framework for European cooperation in education and training towards a European area of education and beyond (2021-2030) (2021/C 66/01).

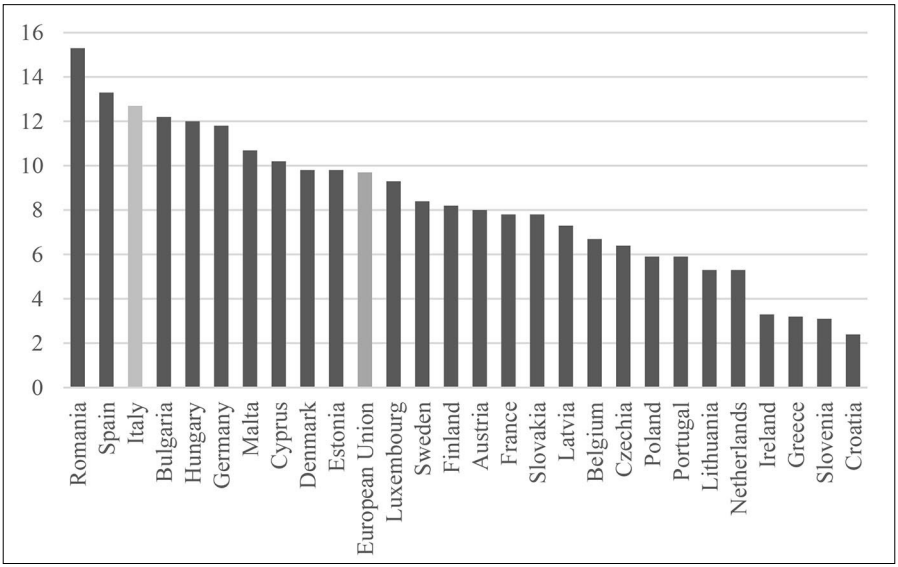


Fig. 1 – Early leavers from education and training by EU nations, percentage value, year 2021

Source: Eurostat, *Labour Force Survey*

In addition, one of the usual characteristics of our country is that it has wide gaps within it. Looking at the ISTAT data at regional level (NUTS2) on early school leavers, we can see an imbalance between the South and the North – Centre regions (Figure 2). In fact, the top five regions in the ranking are the five largest regions in the South: in first place is Sicily with an Eary School Leaving (ESL) rate of 21.2%, followed by Puglia and Campania (with 17.6% and 16.4% respectively). These regions are above the national average (12.7 percent). On the other hand, regions like Molise, Marche, Abruzzo, Friuli-Venezia Giulia, Basilicata, Lazio, Veneto and Emilia Romagna are below the European Union target of 10 percent.

The strong disparities in school dropout rates that emerge in certain geographical 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 aim of this chapter is to elaborate a statistical indicator of early school leavers, capable of measuring this phenomen at an increasingly detailed and flexible spatial statistical detail.

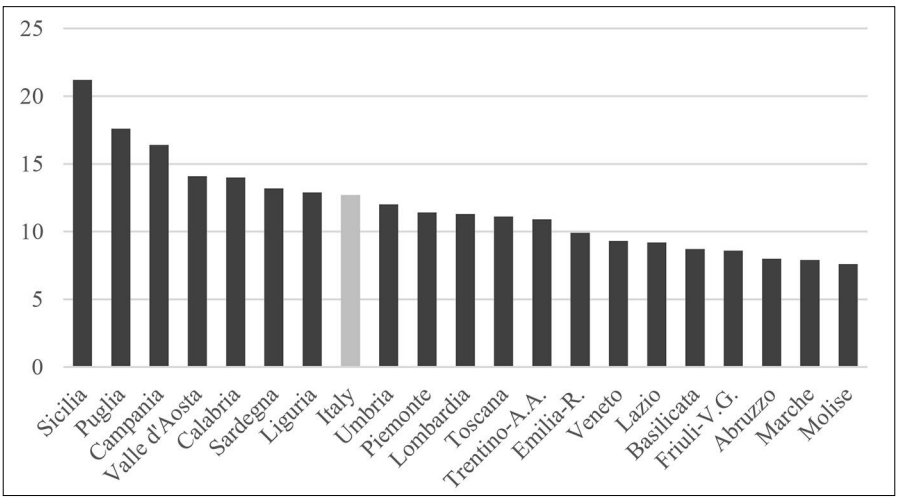


Fig. 2 – Early leavers from education and training by regions, percentage value, year 2021

Source: ISTAT, *Labour Force Survey*

2. Data and methods

To try to quantify the phenomenon of early school leavers at the territorial level, open data from the national student registry (source: Italian Ministry of Education – MI) were used. The information from the open data registry of the Italian Ministry of Education is currently available for the school years 2015/2016, 2016/2017, 2017/2018, 2018/2019, 2019/2020, and 2020/2021, 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. Therefore, by tracking the school pathway of students in a given territory (e.g. the province) for the period at the end of compulsory schooling (i.e. between 15 and 18 years of age), it is possible to calculate how many drop-outs there are in that territory. Through these data, it is possible to track fifthteen-year-old students present in Italian schools, up to the age of eighteen, in order to quantify how many of them, in the pre-COVID (school year 2018/2019) and

post-COVID period (school year 2021/2022), stopped attending schools in a given territory. In this specific exercise, the early school leavers rate has been calculated at the provincial level and this territorial level allows important spatial differences to emerge. But an important aspect of this indicator is that it could also be calculated at the municipal level or at 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 and so on.

In practice, we follow the schooling pathway of the cohort of 15-year-olds enrolled in all schools[2] 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 calculate 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):

$$OUT\text{-}E_{i,\,t} = (ISCR18_{i,\,t} - ISCR15_{i,\,t-4}) / (ISCR15_{i,\,t-4})*100$$

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 not to be due to actual school drop-outs, but to other possible factors which are difficult to identify statistically and have a small 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 can change the value of the drop-out rate, overestimating or underestimating it, even if:

– internal or external migratory processes are more likely for young people, and having a child of at least 15 years old to change school in another country or region (or another province, municipality) means not being so young (at least over 35): for this reason, it seems to us that this factor may have little effect on 15-18 year olds.
– there might be the cancellation of a student in a certain school due to traumatic events (death); but, mortality rates are very low in that age group;
– there might be a student's decision to change school, choosing one outside the territory taken into consideration. In Italy it is not common for

[2] In this analysis, only state public schools were considered. Other types of public schools and new regional training institutions (ESL) were not taken into account.

minors to go far from their families, so even this factor seems likely to have little effect for fairly large territories[3].

Second, the presence of "anticipatari" students (students with anticipated enrolment and so 15-year-olds attending the third year of secondary school,): these students could graduate and therefore leave school before the age of 18. This factor can be neutralised by separating from the share of 15-year-olds those attending grade 11.

Third: in Italy, if you do not pass the school exams, there is the possibility that you will have to repeat the exams and therefore will not be able to graduate at the age of 18.

Fourth: data are not available for Trento and Bolzano provinces and currently not available for the Sardegna region.

Fifth: students who choose different schooling paths, such as IFP (Istruzione e Formazione Professionale) – VET (Vocational Education and Training) becouse this type of scholastic offer is regulated and controlled by regions, therefore no national registry exists concerning students enrolled in this type of offer.

3. Main results

3.1. ELET and OUT-E

Comparison of the two different sources for the regional level shows good convergence of the indicators OUT-E and ELET (Figure 3). The OUT-E indicator is higher than the ELET indicator: 15 percent for Italy compared to 12.7 percent.

OUT-E is lower than ELET for Liguria (-1.9 p.p. lower), Umbria (-1.3), Piemonte (-3.1), Lombardia (-1.5) Veneto (-2.6) and Friuli Venezia Giulia (-2.9), while for regions in the south, islands and center it is higher (in Campania OUT-E is 9.2 percent points higher than ELET).

[3] Students changing school as a result of a family separation event are not considered in this case, but rather as part of the internal migration flows mentioned above. The only difference is that the motivation to move is not work-related but separation-related.

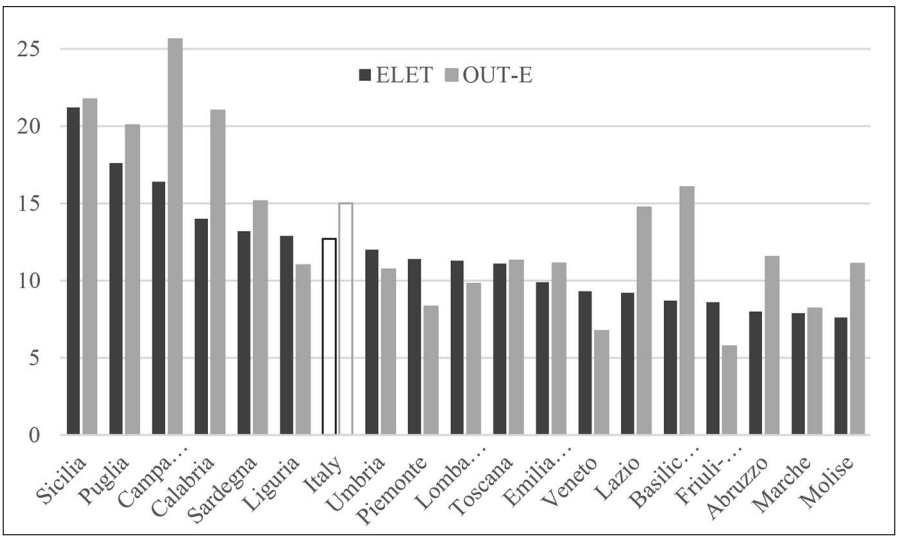


Fig. 3 – Early leavers from education and training (ELET) in year 2021 and Out of Education (OUT-E) in school year 2020/2021 by regions, percentage value

Source: ISTAT, *Labour Force Survey and processing on MI data*

3.2. OUT-E provinces

An overall territorial consistency seems to emerge from the preliminary results even at provincial level (Figure 4).

The southern Italian provinces are characterized by higher early school leavers rates: twenty provinces have one student in five that drops out of school early; this rate increases in the Naples and Caserta provinces where the OUT-E rate is close to thirty percent; but among the worst provinces, some central and northern territories (such as Prato, Piacenza, Florence, Livorno, Reggio nell'Emilia, and Imperia) also stand out.

This last result also highlights how there is 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 the OUT-E 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.

From these results, two pieces of evidence emerge that are worthy of further analysis on the determinants of school dropout rates:

– precisely in the territories with more job opportunities, dropout rates are lower, so seeking a job position in the labor market is not a cause of high dropout rates;
– the pandemic, resulting in a "freeze" on "failures" in schools, has limited school dropouts, but this does not mean that skills have improved (INVALSI, 2022). Making a student repeat the school year would thus seem to be the main tool for pushing him or her out of school. Are we sure this is what is wanted?

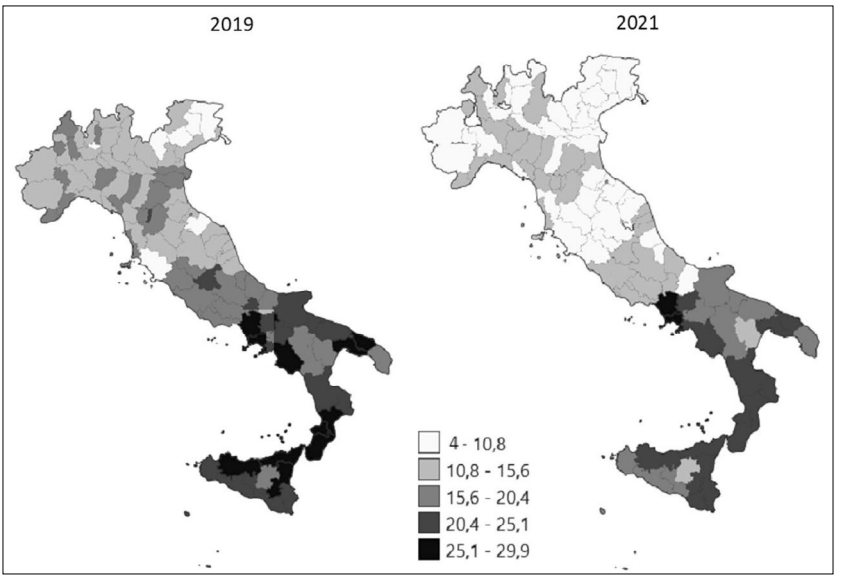


Fig. 4 – A geographic projection of the Out of Education (OUT-E) OUT-E in the pre and post covid time, school year 2018/2019 – and 2020/2021 by provinces, percentage value

Source: processing on MI data

3.3. OUT-E internal areas

An interesting result emerges from Figure 5 where OUT-e is calculated for Internal Areas, regions and capitals of regions. Dropping out of school as measured through our OUT-E indicator is a phenomenon that occurs mainly

in large urban cities (capitals of regions), rather than in less urbanized areas, as is clearly evident when comparing the "internal areas" of each region with the relative figure for the regional capitals. The OUT-E indicator records values that are far higher than those of the relative regional internal areas, with the exception of Lombardy.

This result opens up interesting insights into the ability of schools in inland areas to preserve their students from failure and the willingness of students to pursue their studies in a more determined way.

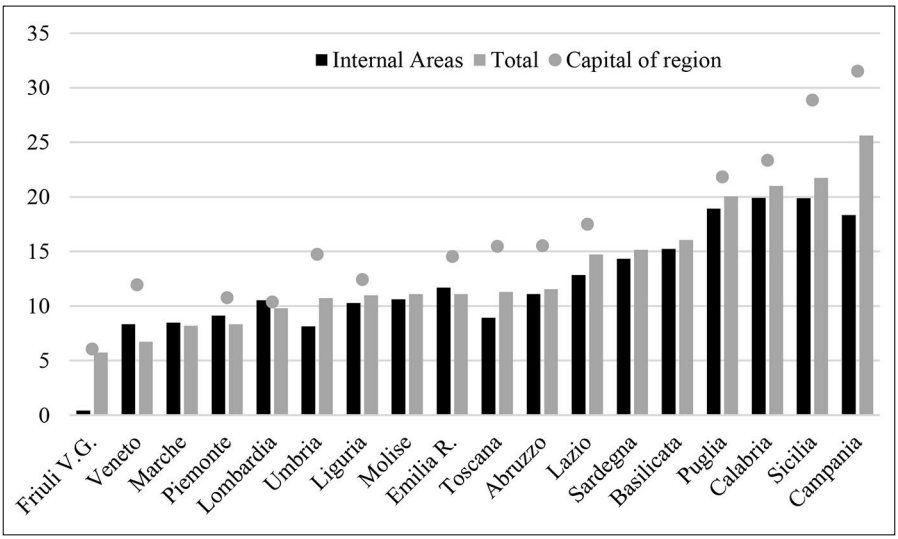


Fig. 5 – Out of Education (OUT-E) OUT-E by internal areas, total regions and capital of region, school year 2021/2022, percentage value

Source: processing on MI data

3.4. OUT-E and other sources of data

The possibility of comparing multiple data sources is a necessity and an opportunity to enrich and integrate information regarding education (Baldazzi, 2021). OUT-E results were compared with INVALSI tests with the objective to examine the implicit and explicit early school leaving rates and to understand the real capacity of a school to improve human capital and local capabilities.

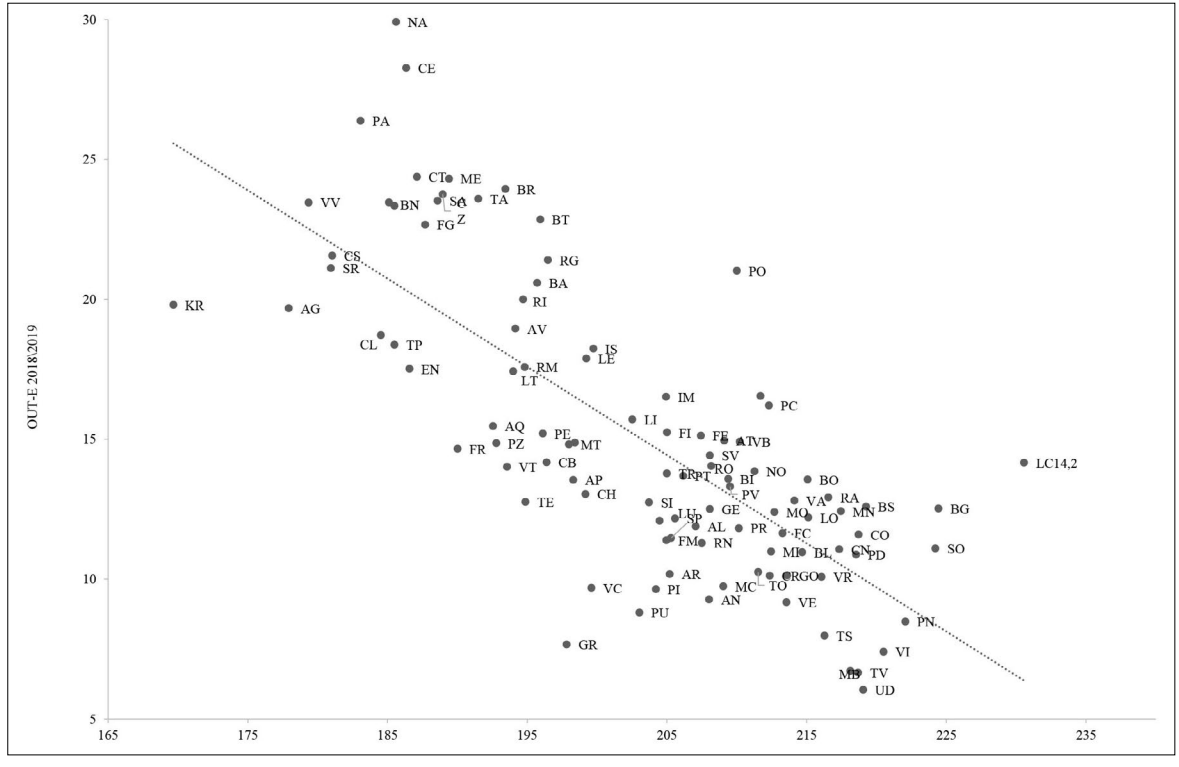


Fig. 6 – Out of Education (OUT-E), y axis, and Mathematics results in the INVALSI tests, x axis, by provinces, school year 2018/2019, percentage value and results

Source: processing on MI data and INVALSI data

For example, by comparing the results of the mathematics INVALSI tests with OUT-E (Figure 6), a correlation is found between the two indicators by province: the provinces where the results in the tests are better are also the provinces where the dropout rate is lower. At a wide enough territorial level (e.g. province or local labour system) this indicator prefigured as a measure of the school drop-out rate; the same indicator declined at a more punctual level of detail (specific school) can be indicative of the drop-out rate of a given school. In the latter case it shows a school's capacity/incapacity to "retain" its students (Save the children, 2019); this information, if linked with the INVALSI data, could provide an answer to the question on the effectiveness rate of a given school.

4. Next step and conclusions

The availability of fine-grained statistics on OUT-E, combined with INVALSI data on student assessments and implicit "dispersion", could well serve the education system to signal areas and even schools in need of intervention.

The preliminary analysis carried out thus makes it clear how the current scenario is prodromal to generating a new further widening in income disparities and to follow a non-development path in the near future, above all in those territories in which there will be a contemporary presence of a demographic contraction and a low instruction rate of the population. But it also makes it clear how important it is to act quickly and in a targeted way to reduce these disparities and how important it is to have an indicator with a smaller territorial detail (province, internal area, school).

References

Armenise M., Baldazzi B. (2022), "Abbandono scolastico e territorio: una misura più dettagliata utile per le politiche", *Menabò*, 183, retrieved on April 7, 2023, from: https://eticaeconomia.it/abbandono-scolastico-e-territorio-una-misura-piu-dettagliata-utile-per-le-politiche/.

Baldazzi B. (2021), "The INVALSI data for the 2030 Agenda", in P. Falzetti (a cura di), *INVALSI data to investigate the Characteristics of Students, Schools and Society – IV Seminar "INVALSI data: a tool for teaching and scientific research"*, FrancoAngeli, Milano, pp. 110-123.

Baldazzi B., Cascioli R. (2019), *The Early School Leavers as Indicator of Educational Poverty*, presented at "Educational Poverty: Definitions, measures and Mapping, Lunch-to-lunch Workshop", Pisa, November 4-5.

Bugamelli M., Lotti F. (2018), "Productivity growth in Italy; a tale of a slow-motion change", *Questioni di economia e finanza*, 422, Banca d'Italia.
INVALSI (2022), *Rapporto INVALSI 2022*, retrieved on April 7, 2023 from: https://invalsi-areaprove.cineca.it/docs/2022/Rilevazioni_Nazionali/Rapporto/Rapporto_Prove_INVALSI_2022.pdf.
ISTAT (2021), *Ciclo di audizioni sul tema della dispersione scolastica*, retrieved on April 7, 2023 from: https://www.istat.it/it/archivio/259206.
ISTAT (2022a), *Previsioni della popolazione residente e delle famiglie – Base 1/1/2021, Statistiche Report*, retrieved on April 7, 2023 from: https://www.ISTAT.it/it/files//2022/09/REPORT-PREVISIONI-DEMOGRAFICHE-2021.pdf.
ISTAT (2022b), *Rapporto SDGs 2022. Informazioni statistiche per l'Agenda 2030 in Italia*, retrieved on April 7, 2023 from: https://www.istat.it/it/archivio/275718.
Save the children (2019), *Il tempo dei bambini. Atlante dell'infanzia a rischio*, retrieved on April 7, 2023, from: https://s3.savethechildren.it/public/files/uploads/pubblicazioni/x-atlante-dellinfanzia-rischio-il-tempo-dei-bambini_2.pdf.

 ISBN 9788835179498

The authors

Massimo Armenise is a researcher at the Italian National Institute of Statistics (ISTAT), focusing on the evaluation of public policies and the development of territorial indicators related to business, research, and innovation. He holds a degree with honors from the University of Bari, a Master's in International Economics from the University of Rome Tor Vergata, and a PhD in Theories and Methods of Individual and Collective Decision-Making from the University of Bari.

Barbara Baldazzi is a researcher at the Italian National Institute of Statistics (ISTAT) since 1997, focusing on the Sustainable Development Goals and the Equitable and Sustainable Well-being. She was 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.

Claudia Busetti has been working at ISTAT since 2010, mainly dealt in data processing and statistical analysis. She is currently working in the sector of the Education and Training statistics and collaborating in the design and development of the new Statistical Register of Education and Training. Previously she also worked in the ISTAT quality' sector.

Luciano Canova, PhD Economics. Lecturer of Behavioural Economics at Master MEDEA (Scuola Enrico Mattei – Università di Pavia). Professional trainer for financial literacy labs and workshops, he works with Feduf and Museo del Risparmio.

Roberta Cipriano is a PhD student in the "Life Course Research" national PhD programme at the University of Florence and the University of

Siena. Her research focuses on educational inequalities from a life course perspective, with particular attention to gender dynamics and the intersection of various life domains, including family structures, social contexts, and the influence of gender norms on family and career aspirations.

Eleonora Faggiano is associate professor in Mathematics Education at the University of Bari Aldo Moro. She is involved in mathematics teacher pre-service and in-sevice professional development and her research interests mainly concern the integration of digital resources in meaningful teaching activities and the learning of mathematics in STEAM contexts.

Giulio Ghellini is Full Professor under retirement in Social Statistics at the Department of Economics and Statistics, University of Siena. He has directed several projects for Eurostat related to survey design and data quality of EU Statistics on Income and Living Conditions and EU Social Survey. His current research interests are focused on Educational System Evaluation, Youth Transition to work, and Educational Mobility.

Gabriele Lombardi is a Research Fellow in Social Statistics at the University of Florence. His research interests include internal migration of Italian university students, university-to-work transition, adolescent gambling addiction, and the propensity to adopt pro-sustainable behaviors among adolescents.

Silvia Montecolle, researcher at ISTAT and PhD in Applied Research in the Social Sciences. She has worked mainly on sample surveys on households and individuals. In the past, she has done research on issues related to quality of life and work-family reconciliation; she is currently working on the labour market and households.

Sara Mori, PhD in Evaluation of Educational Systems and Processes, psychologist-psychotherapist. She is a researcher at INDIRE: she carries out research activities in the field of the development of students' transversal competences and the promotion of well-being at school.

Jessica Niewint is a Technologist at INDIRE and the Head of the Research Department on innovative methods for laboratory education and new technologies and principal investigator in international and national research projects based on STEM education and didactics supported by Artificial Intelligence tools.

Giovanna Paladino, PhD, is Head of the Chairman's Staff at Intesa Sanpaolo and founder and director of the Museum of Saving. Formerly at the IMF and LUISS University, she studied at HEI (CH) and Brown(US). Her research spans international economics, finance, and financial education, with publications in top journals including Energy Economics, JKE and REHO.

Alessia Rosa is an INDIRE Researcher and Degree Programme Direc-tor in psychology course and Techniques of Human Resources at the IUL Telematic University. She holds an honors degree in Educational Sciences and a PhD in Media Education from the University of Turin. Her research interests mainly concern technological applications in teaching and neuro-science. She works on experimenting with diversified teaching proposals in the preschool.

Marilena Salsano, primary school teacher for about twenty years, currently headteacher of a primary and secondary institute. She is now attending a PhD course in Public Sector Economics at DISES-UNISA.

Camilla Spagnolo is a researcher of Mathematics Education at University of Ferrara. She studies perceived difficulty in mathematical tasks, adopting both qualitative and quantitative approaches. She investigates the interplay between emotions and perceived difficulty across all school levels, exploring how cognitive and affective factors shape students' problem-solving experiences and how teachers' perceptions align with students' perspectives.

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.

 ISBN 9788835179498

Over time, the seminars on INVALSI data have increasingly highlighted the importance of using the data published by the Institute, also in combination with data from other bodies and institutions, which in various ways provide important elements to shed light on the Italian school system. Indeed, the possibility of comparing different data sources is both a necessity and an opportunity to enrich and integrate information on education.
Evidence of this can be found in the contributions collected in this volume. They come from the seventh edition of the event "INVALSI data: a tool for teaching and scientific research" (Rome, 27-30 October 2022), in which – and not only – the data provided by INVALSI are tools to study the field of education from different perspectives.
As a statistical service, we hope that this volume will not only help to increase knowledge about the potential of the data, but also stimulate curiosity and new research questions.

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

FrancoAngeli
La passione per le conoscenze