FOSTERING THE EVALUATION CAPACITY BUILDING FOR SCHOOL IMPROVEMENT

VIII 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 9788835191087

Copyright © 2026 by FrancoAngeli s.r.l., Milan, Italy. ISBN 9788835191087

FOSTERING THE EVALUATION CAPACITY BUILDING FOR SCHOOL IMPROVEMENT

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

edited by
Patrizia Falzetti

FrancoAngeli

 ISBN 9788835191087

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: 9788835181385

Isbn e-book Open Access: 9788835191087

Copyright © 2026 by FrancoAngeli s.r.l., Milan, Italy. ISBN 9788835191087

Index

 ISBN 9788835191087

Introduction

by Patrizia Falzetti

Recent changes to the school system have made assessment an increasingly important tool for reflection and improvement. Moving away from a purely measurement-based approach means recognising the value of assessment as a shared process that can inform educational and organisational decisions. With this in mind, we have collected contributions presented at the VIII Seminar, "INVALSI data: a tool for teaching and scientific research", held in Rome from 23 to 26 November 2023. This publication contributes to the ongoing dialogue surrounding the evolution of education systems, with a specific focus on evaluation, self-evaluation, and teaching innovation processes within Italian and European schools. The experiences and research presented all focus on the role of assessment and reflective practices in supporting the professional development of teachers and school leaders, improving teaching and learning processes, and strengthening the governance of educational institutions.

The first chapter opens the volume with a multiple case study conducted in Lombardy during the period of remote teaching that followed the Covid-19 pandemic. This qualitative and quantitative analysis relates the institutional context to teachers' beliefs and their declared practices regarding formative assessment, learning, and student engagement. The second chapter focuses on the integration of artificial intelligence in education, with a particular focus on the European Erasmus AI4T project. It considers the development of teachers' and school leaders' skills and awareness of AI, as well as the ethical and professional implications of its introduction into teaching practice. The theme of institutional self-evaluation is revisited in the third chapter, which analyses the results of the Valu.E for Schools research intervention, which aimed to strengthen teachers' and school leaders' evaluation skills. The fourth chapter shifts the focus to small schools, exploring their

specific contribution to the development of self-evaluation practices and cultures within comprehensive schools. Finally, the fifth chapter broadens the perspective to an international level by analysing the results of the IEA PIRLS 2021 survey with a specific focus on Italy. Together, these contributions provide a multi-level perspective on the ongoing changes in schools, combining empirical research, theoretical insights, and practical implications. The volume therefore aims to facilitate dialogue between research, practice and policy for those engaged in rethinking schools as organisations capable of learning, evaluating themselves and innovating in a responsible and sustainable manner.

1. Deepening teachers' beliefs, practices, and school contexts during Emergency Remote Teaching. The role of INVALSI data

by Marco Giganti, Emanuele Marcora

The contribution presents a multiple case study of three omnicomprehensive schools in Lombardy aimed at observing and describing their context, teachers' beliefs, and practice statements about formative assessment, learning, and student engagement in the Emergency Remote Teaching activated during the critical period of the Covid-19 pandemic, and the relationship between them and the context. Specifically, the qualitative-quantitative analyses carried out on the results emerging from the semi-structured interviews conducted with Headmasters, document analysis (the three-year plan of educational offerings, self-evaluation report, circulars, policy acts, etc.), the questionnaire submitted to teachers, and focus groups conducted with the HM, the internal evaluation team and selected teachers by school order will be presented. In the research in general and in the introduction to the last phase in particular, INVALSI data from the 2021 surveys were used in order to propose a comparison and initiate reflections on the performance of the schools. The case study is part of a larger research project of which hints will be provided.

Il contributo presenta un caso studio multiplo su tre Istituti omnicomprensivi in Lombardia finalizzato a osservare e descrivere il loro contesto, le credenze degli insegnanti e le dichiarazioni di pratiche sulla valutazione formativa, l'apprendimento e l'impegno degli studenti nell'Emergency Remote Teaching attivato durante il periodo critico della pandemia di Covid-19 e la relazione tra loro e il contesto. Nello specifico, saranno presentate le analisi quali-quantitative effettuate sui risultati emersi dalle interviste semi-strutturate condotte con i presidi, dalle analisi documentali (PTOF, RAV, circolari, atti di indirizzo ecc.), dal questionario somministrato agli insegnanti e dai focus group condotti con il dirigente, il NIV e due insegnanti selezionati per

ordine scolastico. Nella ricerca in generale e nell'introduzione all'ultima fase in particolare, sono stati utilizzati i dati INVALSI delle indagini del 2021 per proporre un confronto e avviare riflessioni sulle prestazioni delle scuole. Lo studio di caso fa parte di un progetto di ricerca più ampio di cui saranno forniti alcuni riferimenti.

1. Introduction and theoretical background

The centrality of assessment in the learning paradigm is a theme widely developed in the literature and finds particular expression in formative assessment (Scriven, 1967; Bloom, 1971; Black and William, 1998; Weeden, Winter and Broadfoot, 2009; Brown *et al.*, 1992).

The pedagogical potential and social effects of this kind of assessment have also long been considered by UNESCO and the Council of Europe (CoE), which, during the first phase of Covid-19 spread, saw fit to incentivize its use to support the engagement of students then their learning process (Engzell, Frey and Verhagen, 2020; UNESCO, 2020c, 2020d, 2020e; Hughes, 2020).

In spring 2020, national governments decided to close schools and abruptly switch to distance learning as measures to counteract infection and not disrupt educational activities. In most cases, school systems have adopted Distance Learning (DL) through e-learning or videoconferencing platforms (UNESCO, 2020a, 2020b). Some research sees formative assessment as a useful tool to foster student engagement and inclusion, especially in online education. A study by Z. Chen, J. Jiao, and K. Hu (2021), carried out in the pandemic period, finds that temporally and spatially separated students and teachers are subject to long-term criticality typical of distance education.

In this case, it is appropriate to take up the distinction postulated by Hodges *et al.* (2020) between online distance learning (DL) and emergency remote teaching (ERT). Unlike educational experiences designed to be online, ERT is a temporary and alternative mode of instructional delivery due to crisis circumstances. It involves fully distance teaching strategies; education, which would otherwise be delivered face-to-face or as blended or hybrid teaching, will revert to the traditional format when the emergency has receded. In this context, the main goal is not to recreate a robust educational ecosystem but to provide temporary access to education and learning supports, while also preventing possible risks of disengagement typical of emergencies.

Research shows that students feel isolated in online education (Hammond, 2009), resulting in high dropout rates (Hodges and Kim, 2010), high boredom rates, and poor performance (Fredricks, 2015). According to Chen,

Jiao, and Hu, the use of formative assessment in such contexts can be an effective solution. Implemented with forum discussion tasks, quizzes, and tests, it is effective in improving engagement and increasing learning. The studies mentioned explicitly refer to a mixed mode of teaching between online and presence, specifically concerning DL. According to the authors, there is limited evidence that this also works in the typical ERT mode.

Central to the management of such didactic has been the teaching profession, led by beliefs and transmuted into practice. The most problematic aspects emerging from empirical research (Richardson, 2002) concern the link between conceptions and effectiveness of teacher training (Balduzzi and Vannini, 2008). In this regard, there are studies inherent to teacher change (Floden, 2002; Richardson and Placier, 2002) and how beliefs and practices change: do the former follow the latter (and vice versa) or are they mutually interacting and synergistic (Goffman, 1973; Peterman, 1993)?

This contribution aims to present a multiple case study of three omnicomprehensive schools (primary to high school) in the Lombardy Region, to investigate the context in which teachers operated during COVID-19, their beliefs, and practice statements about formative assessment, learning, and engagement during ERT, what are the links between context and beliefs and practices. In the first stage, semi-structured interviews were conducted with principals, document analysis (the three-year plan of educational offerings, self-evaluation report, policy acts, circulars, etc.), and qualitative analysis was conducted on them. In the second stage, a questionnaire was administered to all teachers in the identified schools, aimed at investigating the constructs examined, and the data were analyzed quantitatively and qualitatively. In the third phase and from a constructivist perspective, a focus group was conducted with the school principal, the Internal Evaluation Core, and teachers selected by order of school to jointly interpret the data that emerged in the first two phases, and the data were analyzed qualitatively.

Analyses show that teachers favored formative assessment convinced that it could support student learning and engagement in a context considered emergent. Specifically, statistical analyses show that beliefs and practice statements about formative assessment are related to constructivist learning and student engagement; they are found to be consistent and applicable in the ERT context. It is interesting to note that the age and experience of the teachers are a conditioning factor as well as the order of teaching and the emergency. Also, the school influenced the formative assessment and learning beliefs: the smaller one, identifiably stronger with accompanying practices and teacher training was decisive on the agreement averages. The primary school also proved to be a favorable context for formative assessment (FA),

student engagement, and constructivist learning. Other hints of the analyses conducted will be given during the presentation of the contribution.

2. Research objectives and hypothesis

The main objectives are referred to three different areas: *theoretical-conceptual* to identify whether and what relationship exists between teachers' beliefs and practices on formative assessment in ERT; *methodological* to analyse and describe the school context, investigate teachers' beliefs and practices on formative assessment, student engagement and learning, and ERT; and *transformative* to engage school principals and teachers as active participants in the research and, at the end of the data analysis, design possible training paths in relation to the issues that emerged from the research. Three general questions were identified to guide the study: what is the school context in which teachers might have used formative assessment during ERT? (RQ1); What are teachers' beliefs and practice statements about formative assessment, engagement, and learning in the ERT context? Is there a relationship? What kind? (RQ2); how does the school context relate to teachers' beliefs and practice statements? (RQ3). After defining the problem and research questions, two general research hypotheses were formulated: there is a correlation between teachers' beliefs and practice statements about FA, learning, and student engagement in the ERT context (GHp1); school context is related to teachers' beliefs and practice statements (GHp2). The formulation of the following specific research hypotheses is intended to assume precision and concreteness with respect to the general ones: teachers' beliefs about learning are related to the application of FA in ERT context (SHp1); teachers' beliefs about student engagement are related to the application of FA in ERT context (SHp2).

3. Materials and methods

The multiple case study (fig. 1) took place in the context of the Lombardy Region (Italy) – specifically in the metropolitan city of Milan – and involved, through a non-probabilistic sampling of volunteers, three omnicomprehensive school institutes (from primary to high school, from first to thirteenth degree). The first is a state (girl's) boarding school in the city center of Milan and has 62 teachers, the second is a private Catholic inspired institute and is also located in the center and has 69 teachers, the third is in two cities in the province of Milan and has 147 teachers.

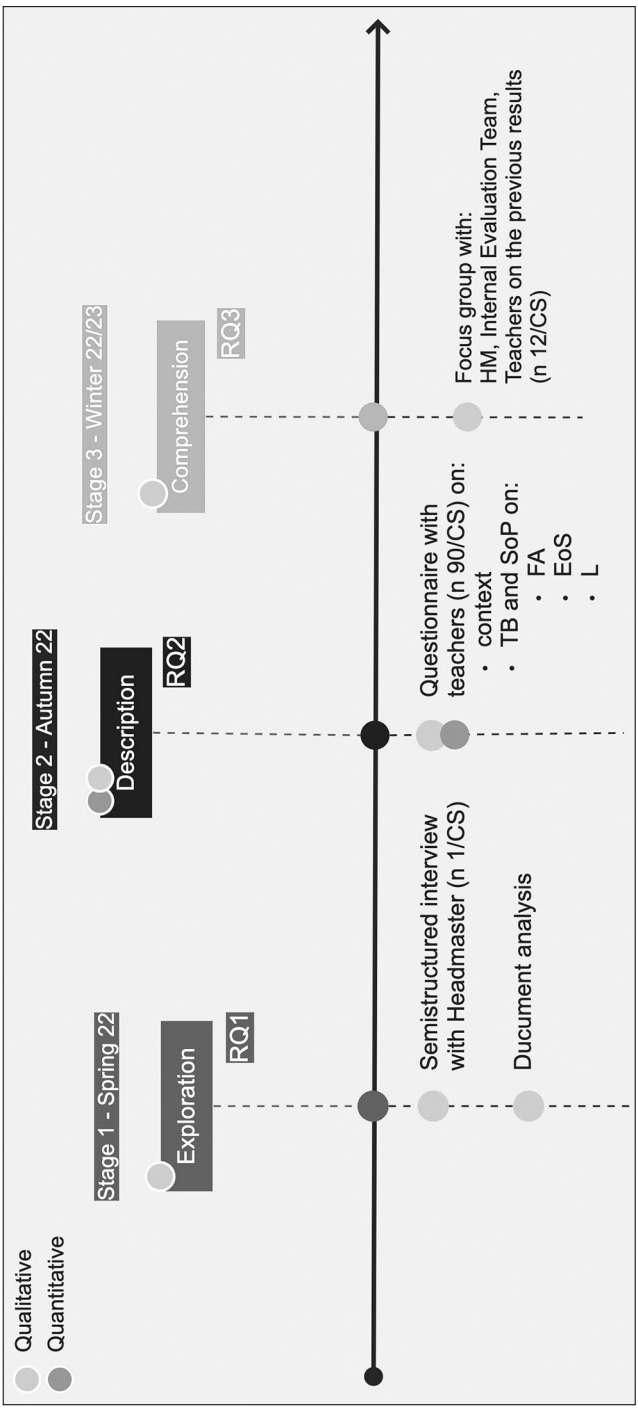


Fig. 1 – Research design

The study aims to investigate the context, beliefs, and statements of practice of teachers during two specific periods, both of which are part of the broader ERT period: remote learning in March-June 2020 (RL) and integrated digital education in September 2020-June 2021 (IDE).

Three semi-structured interviews were conducted on the Institutes' premises, one for each school principal and/or deputy, recorded and transcribed *verbatim*. The first of these, lasted one hour and forty minutes, was held at the School 3 (S3) HM office in May 2022 in the presence of the deputy at the same time; the second, lasted thirty-eight minutes, in the School 1 (S2) faculty room, in June 2022, with the presence of the deputy alone; the third was more articulated, took place in February 2023, and the first part of the interview was conducted with the rector's collaborator in his office in the presence of the principal of the junior high school (forty-four minutes), the second with the primary school principal in a classroom (thirty-five minutes), the third with the secondary school principal in his office (thirty-two minutes), the latter was also rector *pro tempore* at the time of ERT.

The focus group is aimed at sharing the analysis carried out by the researcher and based on this, at carrying out further investigations and reflections with the identified stakeholders. The analyses of the data collected are qualitative and are carried out using the MaxQDA software. The S1 focus group was 9 people, the S2 11, and the S3 7, and were conducted in spring 2023.

Transcribed interviews, focus groups, and documents were analyzed with the MaxQDA software using a code system (tab. 1) structured *ex ante* by reference to the variables identified for the questionnaire and ex-post because of the recurrence of certain themes important for the study of the context and the situations that occurred (Duverger, 1961; Rositi, 1971; Bruschi, 1999).

Tab. 1 – System of codes

System of codes		*Annotation*	*Frequency*
ENG		Beliefs and practices about engagement	0
	PRAXENG	Practices to support student engagement	19
	CONENG	Student engagement beliefs	8
APP		Learning beliefs and practices	4
	CONAPP	Beliefs about student learning	42
	INVALSI tests	Discussions about the results for INVALSI tests	15
	PRAXAPP	Practices to support student learning	11
	APP post-ERT	Learning after ERT	13
VAL		Beliefs and practices about assessment	1
	Training VAL	Assessment training	15
	CONVAL	Assessment beliefs	82
	PRAXVAL	Assessment practice statements	34
	Assessment in ERT	How to apply assessment during ERT	48
	VAL post-ERT	Post-ERT assessment modalities	8
Emotions		Pandemic emotions	13
	Emotions +	Emotions positives	16
	Emotions -	Emotions negatives	45
Context		Background information	83
	REL TER	School relations with the territory	20
	REL FAM	Relations between schools and families	44
	DS and collaborators' role	Management's role in the application of the FA	42
	Projects	Projects developed during Covid	9
ERT		Emergency Remote Teaching	6
	ERT training	ERT training	9
	DIGICOMP	Digital competences	14
	CONERT	Teachers' beliefs during the ERT	33
	INDICMIN	Relationship to ministerial or policy direction	10
	PRAX ERT	Teacher practice during the ERT	45
	STUD in ERT	Attitudes of students during the ERT	15

The questionnaire is subdivided into three blocks and aims to collect data on teachers' socio-professional status (16 questions) and, according to their point of view, the school context during the ERT period (9 questions). In addition, teachers' convictions and declarations of practice concerning the constructions and variables in question are recorded using eight scientifically validated and specially constructed Likert scales with four requests

for examples. A first version of the questionnaire was drafted, which was submitted to a tryout phase with teachers from different types of institutes and school orders who sent observations and suggestions. The questionnaire was administered online by sending a cover letter and using Qualtrics software, initially for two weeks, extended then for another; schools 2 and 3 were sent to winter 2022 and school 2 in spring 2023. The analysis of the data collected is quantitative, using SPSS-IBM and R software, and qualitative.

To further strengthen the methodological rigor of the study, it is important to clarify how the questionnaire data were treated.

Aware of the ordinal nature of Likert-type scales, we initially conducted descriptive analyses and parametric tests (ANOVA, etc.) as an exploratory step, in line with a strand of educational research that treats Likert data as approximately interval when sample sizes are sufficiently large (Norman, 2010; Carifio and Perla, 2008; de Winter and Dodou, 2010; Harpe, 2015). In a second step, in order to ensure robustness, we complemented these analyses with non-parametric tests (Kruskal-Wallis for multiple groups, Mann-Whitney for two-group comparisons) and ordinal logistic regression models. The results obtained through these procedures were largely consistent with the preliminary analyses, thus reinforcing the reliability of the findings. This multimodal approach allowed us to triangulate quantitative and qualitative evidence, thereby reducing the risk of distorted interpretations that might result from treating ordinal data as continuous.

4. Results[1]

4.1. Some findings from stage 1: INVALSI data 2021

During the document analysis process, it was crucial to access the INVALSI data of 2021 in order to observe the differences between schools. The results of grades 2 and 5 (average percentage score net of cheating), 8 and 13 (student outcomes at the same scale as the national report) for the subjects of Italian, Mathematics, English (listening and reading) were asked to the school principals (figures 2-15).

[1] In this paper not all results from the analyses (including interviews and focus groups) are considered. This section presents the main results of the statistical analysis of the questionnaire and INVALSI data. References can be found in the discussion section; for more information, please contact the author.

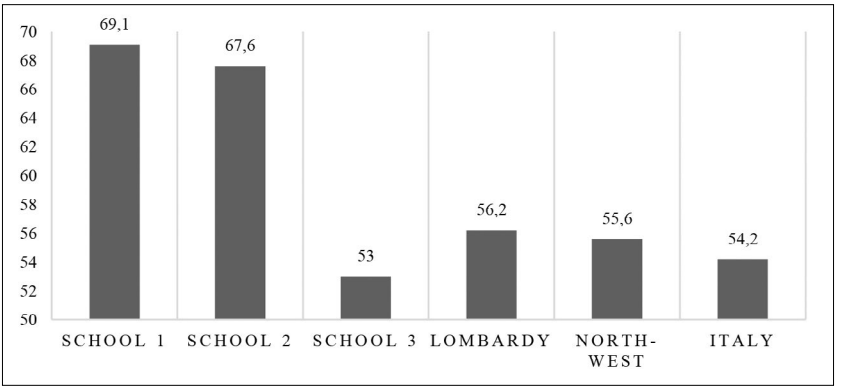


Fig. 2 – Italian (grade 2)

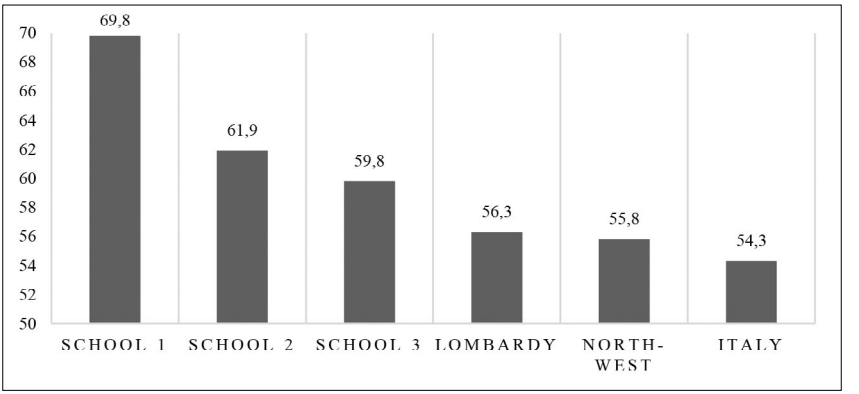


Fig. 3 – Italian (grade 5)

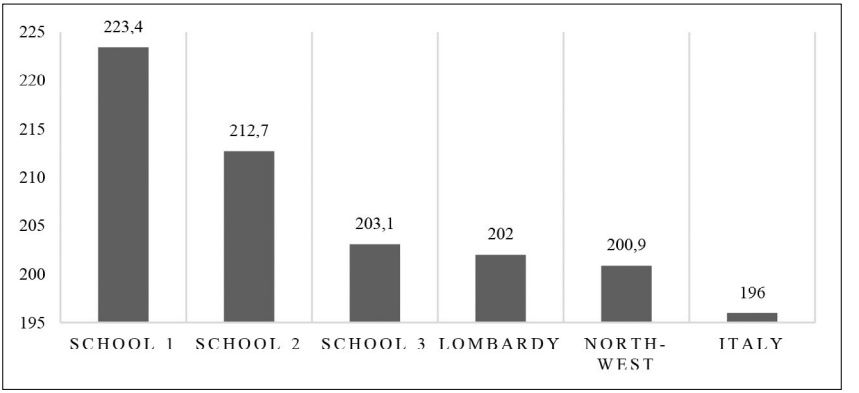


Fig. 4 – Italian (grade 8)

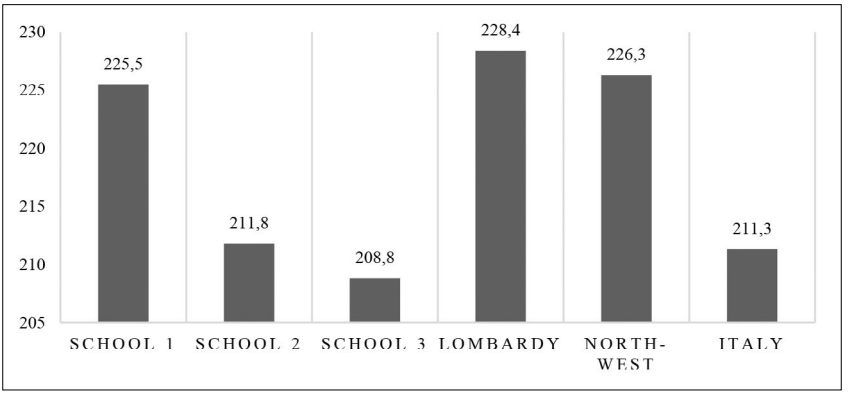


Fig. 5 – Italian (grade 13)

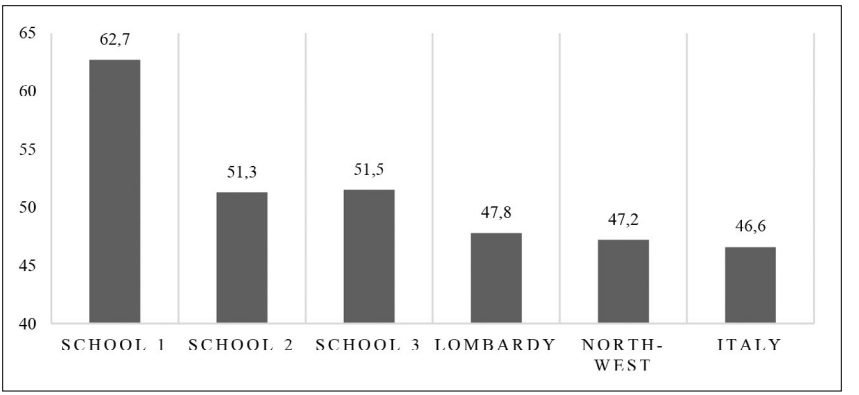


Fig. 6 – Mathematics (grade 2)

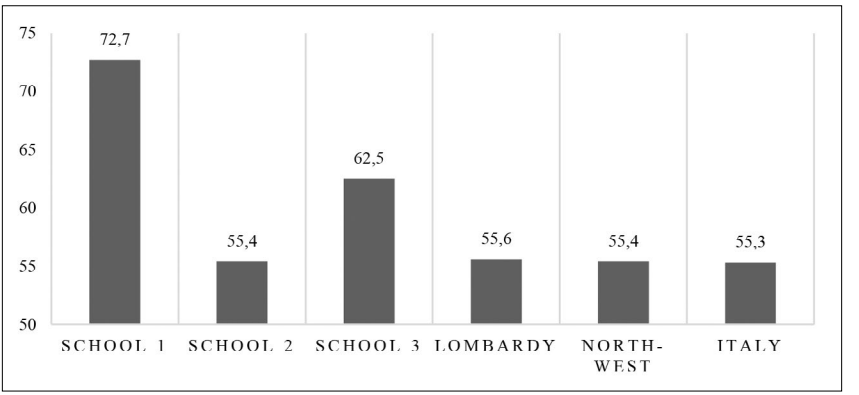


Fig. 7 – Mathematics (grade 5)

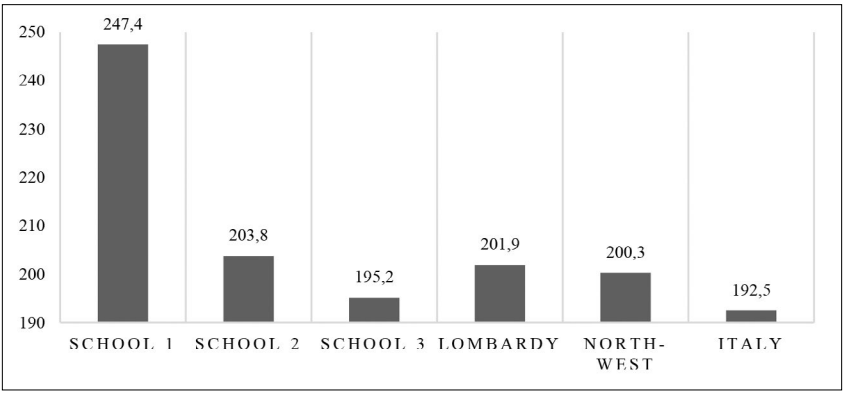


Fig. 8 – Mathematics (grade 8)

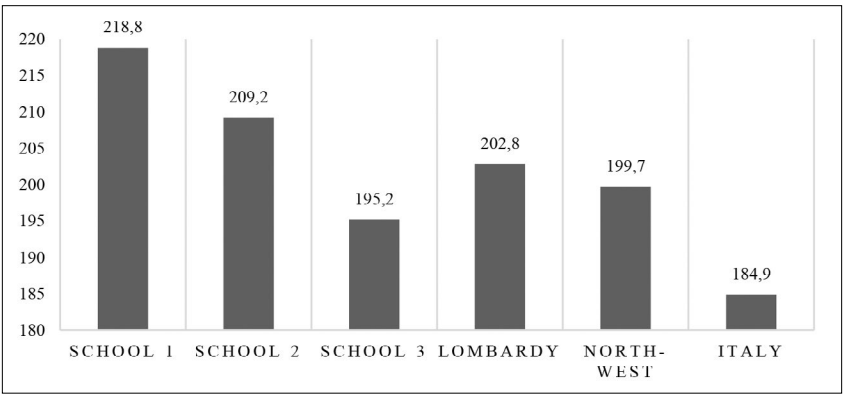


Fig. 9 – Mathematics (grade 13)

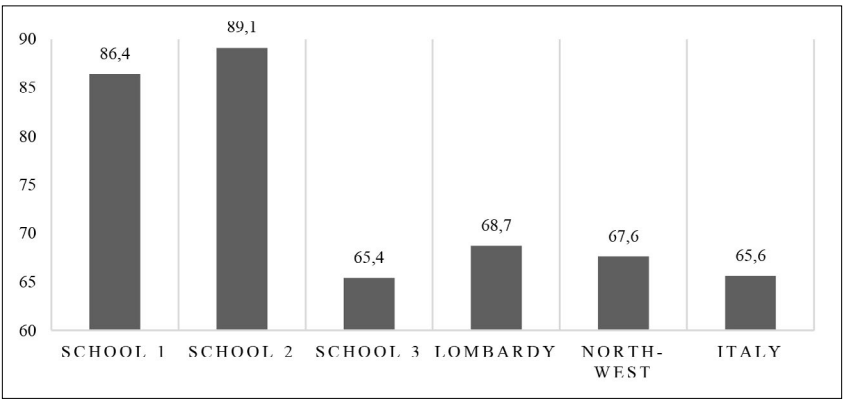


Fig. 10 – English (listening – grade 5)

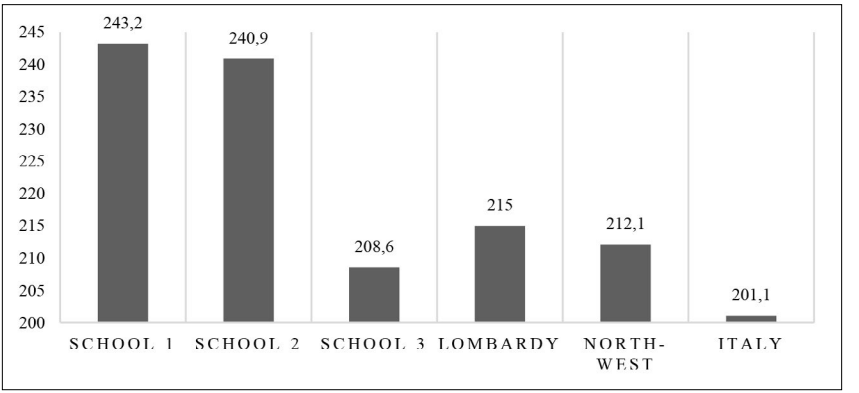


Fig. 11 – English (listening – grade 8)

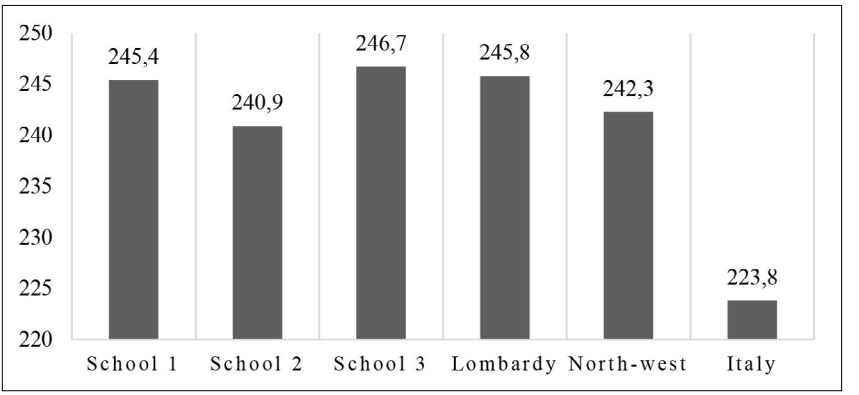


Fig. 12 – English (listening – grade 13)

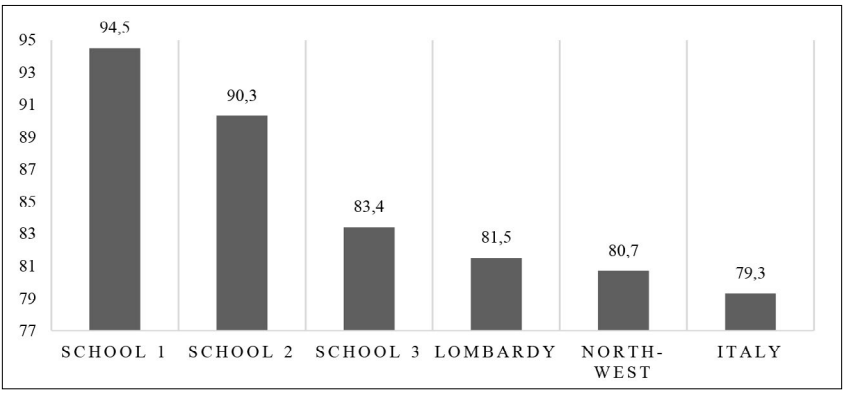


Fig. 13 – English (reading – grade 5)

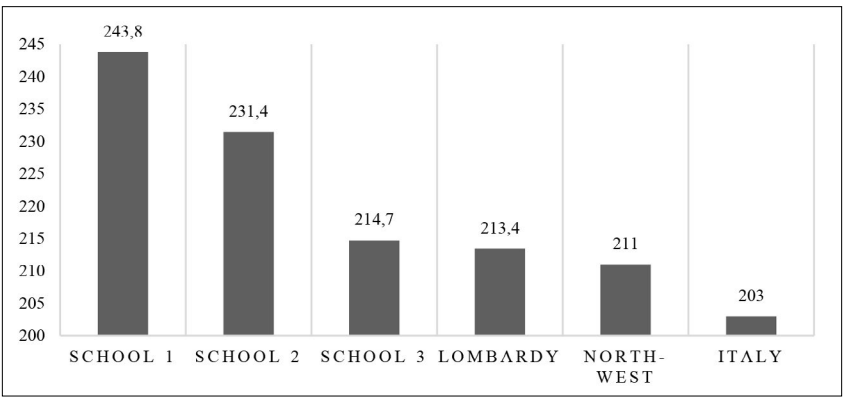


Fig. 14 – English (reading – grade 8)

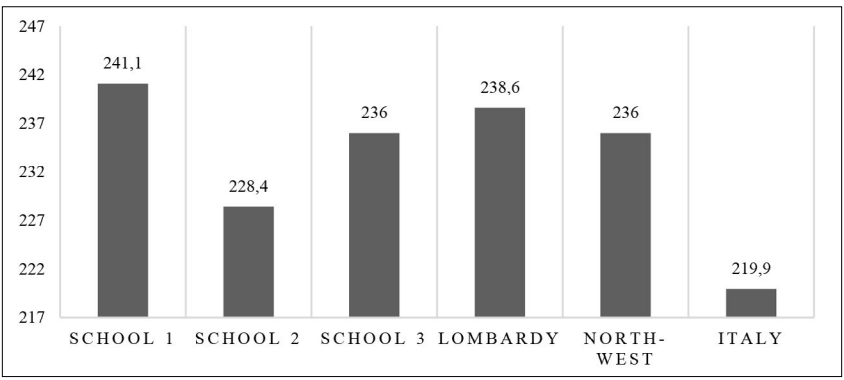


Fig. 15 – English (reading – grade 13)

In order to contextualize the interpretation of the graphs, it is necessary to briefly explain the rationale for using INVALSI data.

The choice of focusing on INVALSI 2021 data is justified by their temporal proximity to the most critical phase of the pandemic, making them particularly relevant to examine the potential effects of Emergency Remote Teaching on students' achievement. The selected domains (Italian, Mathematics, English) represent the core subjects assessed across all schools and are considered essential for monitoring basic competencies. These results were not treated as absolute measures of instructional effectiveness but rather as a reflective tool to be discussed with principals and teachers during the focus groups. Cross-school comparisons may be influenced by socio-economic and territorial variables. For this reason, INVALSI data were used within a triangulated design, integrated with qualitative findings, to avoid

simplistic or decontextualized conclusions. Moreover, considering the data shared by the principals, it was possible to carry out only descriptive analyses and not to conduct inferential analyses (e.g., regression models) to assess the impact of the school context on performance.

As can be seen in the graphs, there is almost a constant repetition of the placement of school results relative to each other. This suggests deepening the influence that context could have on outcomes. In particular, it must be verified that there is an effect of the school variable. This verification was carried out by analysis of the variance with the variables under investigation.

4.2. Structural description of the sample

A total of 193 teachers responded to the questionnaire. Of these, 41 taught in School 1 (S1), representing 21.2% of the sample and corresponding to a response rate of 66.1% (41 out of 62 teachers). Forty-seven teachers (24.4%) were from School 2 (S2), with a response rate of 68.0% (47 out of 69). The largest group, 105 teachers (54.4%), came from School 3 (S3), with a response rate of 71.4% (105 out of 147). The overall sample was predominantly female (77.7%). In addition, seven teachers (3.6%) preferred not to indicate their gender. In general, the response rate is higher than average (50%) with Computer Assisted Web Interviews (CAWI). Overall, the percentages remain unchanged for each school.

Most respondents are between the ages of 36 and 50 (39.4%), with the smallest being 25 years or younger. It should be noted that the bands identified are not homogeneous. Most teachers are in the 14-25 age group (28.5%). Except for those who have been teaching for about a year – and who would not be on duty during Covid anyway – the other bands have almost the same number of teachers. 69.9% of the sample is held, 22.3% were on an annual mission and 2.6% were on temporary duty during the Covid period. Teachers have been teaching for the longest time since 1980 (S1) and 8 since 2022; the oldest holder of an annual office since 2007 (S3) and 15 since 2022. Most teach in High School (40.4%), and minor in secondary school level I (24.4). It should also be noted that two teachers in secondary I also teach in primary and five in HS.

At the time of Covid, 22 worked in a different order (3 S1, 9 S2, 10 S3) and more specifically 7 in primary (4 S2, 3 S3), 6 sec. of 1 (1 S1, 2 S2, 3 S3) and 9 HS (S2). A little more than half (53.4%) have no missions other than teaching. The most common are class coordination, area or department coordination, instrumental function (pupils with disabilities and special ed-

ucational needs, programming, and evaluation, etc.), and project reference. 86.2% report having changed their teaching style during the ERT period.

It is interesting to note that the remaining 13.8% admit that they have not made any changes, essentially reproducing the ordinary modalities in an extraordinary context. More specifically, the highest percentage (17%) is reported by the S2 and the lowest by the S1 (7.5%), and the S3 (14.9%). Those who have varied have acted: on the learning environment that has necessarily been transferred to digital platforms, exploiting their potential such as, for example, the division into small groups to promote interaction; on the duration of the lessons, mainly by reducing their length following the ministerial guidelines; on programming, by limiting the content provided and by choosing those which are most inherent to the daily life of pupils; on evaluation, by giving priority to training; on didactic methods and tools, focusing on the classical room to give more time to interaction and some digital applications to support student engagement; on didactic materials, digitizing all necessary documents and giving preference to audio-video material already present on the web. 72.4% (N 134) of teachers were present in the same institute even during the Covid period (34 S1 – 72%; 69 S2 – 69.7%; 31 S3 – 79.5%); the remaining 27.6% were not asked the context description questions.

4.3. Background description

Relationships with families have been maintained through a variety of tools and modalities: most teachers say they prefer video conferencing platforms that allow eye contact (36.98%), followed by email (28.30%) and electronic communication (21%). Note that 22 teachers (7.07%) also used the private phone. Specifically, those in S1 did not use two tools (face-to-face interviews and the school phone); those in S2 did not identify other instruments and used the electronic register less (10%), and S3 tended to match the average of the total sample.

The electronic register was used primarily to record grades (31%) and, to a slightly lesser extent, for bureaucratic procedures (28%). About one quarter of teachers (23%) reported using it to communicate with families, a result that only partially aligns with the responses to the previous question (21.9%). By contrast, only a small proportion of teachers (15%) indicated that they used the register to provide feedback directly to students. This limited use is particularly relevant when examining formative assessment practices. Specifically, S1 teachers report targets that are consistent with the general sample;

S2 teachers, although they used it less than others, used it primarily to record scores (43%) and less to complete bureaucratic procedures (22%) and communicate with households (13%); S3 less to record votes (27%) and more to communicate with households (27%).

At the end of this section, teachers were asked to indicate some of the projects proposed by the school and considered significant for teaching management in the emergency period. We can distinguish some thematic cores: management of digital didactics and applications with support, STEAM and e-twinning methodology, review of assessment criteria, emotional and psychological support, book club production or architectural appetizers (aimed at student engagement), online dialogues with professionals from different parts of the world, lessons in common between professors from different disciplines.

4.4. Scales for detecting beliefs and practices

This part of the questionnaire is intended to collect teachers' statements on the beliefs and practices on survey themes. The scales[2] (tab. 2) were defined over four levels by reference to the theoretical constructions studied or to certain ministerial documents: convictions and declarations of practice on assessment, student engagement, learning, and ERT. In addition, exploratory factor analyses were conducted to identify latent factors. The hypotheses that guided this exploratory part of the research are oriented toward the analysis of the principal components and the verification of the validity of the scales for the next steps and analyses of the research. It is expected that the item batteries examined will be internally coherent so that they can be considered as valid scales concerning the theoretical background; they highlight some early correlations with other important research variables to support the assumptions made.

[2] The scales are shown in the English translation, the administration took place with the Italian one. In this paragraph are presented only the scales to detect the beliefs and practices of teachers about formative assessment and Emergency Remote Teaching.

Tab. 2 – Summary of variables and latent factors

Scale	*N. LF*	*Name of LF*	*p value*	*Sat.*	*Cronbach a*
VALFOR (Ciani and Vannini, 2017) *Teachers' beliefs about assessment*	2	Summative assessment (item 1-3-7-8): μ 1,74 Formative assessment (2-4-5-6): μ 3,25			
PRAXVAL *Teachers' practices about assessment*	2	Assessment actions of the teacher (item 1-7): μ 3,41 Peer/self-assessment (8-12): μ 2,26	< .001	53.39%	.842
CONERT *Teachers' beliefs about ERT*	2	Emergency of didactics (item 1-2): μ 3,24 Applicability of DDI (3-11): μ 2,20 → 8 e 10 (< 2)	< .001	58.55%	.791
PRAXERT *Teachers' practices about ERT*	2	Use of digital tools and materials (item 1-2): μ 3,32 Agreement with the teaching staff (3-5): μ 3,65	< .001	76.86%	.748

Tab. 3 – VALFOR scale

Variable code	*Variables related to the hypothesized theoretical construct (VALFOR battery)*
VALFOR 1	Sometimes it is necessary to attribute lower marks to urge the student to commit more
VALFOR 2	The main role of assessment during the quarter is to identify the learning difficulties of the student to help him overcome them
VALFOR 3	There is attention and interest in a class only if the student knows that it will be assessed on the concepts expressed by the teacher during the lesson
VALFOR 4	The assessments during the quarter are used to identify not only what the student has learned but also what the teacher must deepen
VALFOR 5	The assessment of students during the quarter allows the teacher to verify the validity of his work
VALFOR 6	Assessment during the quarter should help the student to better understand his learning process
VALFOR 7	The attribution of low marks (or judgments) during the quarter helps the teacher to be better respected
VALFOR 8	The teacher should not let the students know his assessment criteria

Fonte: Ciani and Vannini (2017)

The VALFOR battery items (tab. 3) come from a scale built and validated by A. Ciani and I. Vannini (2017) and are intended to measure formative and summative assessment beliefs. Descriptive battery statistics show medium-low agreement levels (not or not at all agree) with items 1, 3, 4, and 8, and medium-high (somewhat or very agree) with items 2, 5, 6, and 7. There is thus a polarization in two different typologies of assessment designs that the researchers defined by factor analysis that identified two latent factors: formative function (VALFOR 2, 4, 5, 6) and summative (VALFOR 1, 3, 7, 8) of the assessment. Descriptive statistics of the latent factors of the VALFOR battery indicate that this polarization is indeed related to the two different assessment designs during the ERT in favour of the formative function. Furthermore, it appears that there is no "very consensual" level for the summative one.

Tab. 4 – PRAXVAL scale

Variable code	*Variables related to the hypothesized theoretical construct (PRAXVAL battery)*
PRAXVAL 1	I shared with students learning goals before they started working
PRAXVAL 2	Before assigning a test, I clarified to the students what I would assess
PRAXVAL 3	In class, I used different methods of assessment (written, oral, graphic, practical, structured, and unstructured)
PRAXVAL 4	I made sure that the assignments could verify the progress of students' learning
PRAXVAL 5	In addition to the grade (rating), I gave feedback to students
PRAXVAL 6	If I gave feedback on it, I highlighted the strengths or weaknesses of the task
PRAXVAL 7	I provided tips to students to improve their learning
PRAXVAL 8	I asked the students to assess the work of the comrades
PRAXVAL 9	I asked students to provide feedback to fellow students to help them improve
PRAXVAL 10	I asked students to assess their work
PRAXVAL 11	I asked students to identify the strengths and/or weaknesses of their work
PRAXVAL 12	I asked students to propose strategies to improve their work

The PRAXVAL battery items (tab. 4) were defined based on research by D. Wiliam and M. Thompson (2007) and are intended to measure statements of the practical application of the five didactic strategies for the use of formative assessment. Descriptive battery statistics show medium-high agreement levels for almost all items except 8 and 9 that relate to peer review application practices. This could be due to the isolation of students and the didactic typology that limited peer-to-peer work in favour of predominant contact with the teacher. Practices aimed at promoting self-assessment have a medium level, unlike those where the teacher is a major player in which it is high.

Initially, a confirmatory factor analysis of the hypothesis made during the construction of the scale was carried out but did not have the expected results. Subsequently, an exploratory factorial analysis (Varimax method) was performed showing that the Kaiser-Meyer-Olkin sampling adequacy measure is good (.844), the Barlett sphericity test is significant (< .001) and indicates two latent factors (PRAXVAL 1 to 7 and 8 to 12) which saturate 53.39% of the total variance. The first covers the actions of exclusive competence of the teacher and combines the first three didactic strategies of Wiliam and Thompson, the second of the practices of the teacher that make the participating student active (peer and self-assessment) and combines the last two strategies. Finally, a reliability analysis of the scale was carried out which has a good Cronbach Alpha (.842), and it is not necessary to delete any element since the result would always be greater than .8. Descriptive statistics of latent factors indicate that teachers have more frequently carried out actions within their exclusive competence and less those involving student intervention.

Finally, the items of the CONERT and PRAXERT batteries are constructed on the studies of C. Hodges *et al.* (2020) and on the ministerial documents (Note 318 of 11.3.2020 and DM 26 June 2020, n. 39). They study teachers' beliefs and practices in emergency remote teaching.

Tab. 5 – CONERT scale

Variable code	*Variables related to the hypothesized theoretical construct (CONERT battery)*
CONERT 1	RL/IDE or emergency remote teaching is a temporary transition to an alternative mode of providing education due to extraordinary circumstances
CONERT 2	RL involves the use of fully distance learning solutions and education will return to the traditional format once the crisis or emergency has subsided
CONERT 3	The IDE has effectively allowed the integration between teaching in presence and at distance
CONERT 4	The IDE has favored disciplinary and interdisciplinary insights
CONERT 5	The IDE has favored the customization of routes
CONERT 6	IDE has enabled the recovery from learning loss
CONERT 7	The IDE has encouraged the development of disciplinary skills
CONERT 8	The IDE has effectively fostered the development of students' soft skills (personal, social, citizenship, learning to learn, etc.)
CONERT 9	The IDE has effectively improved the correspondence between the teaching action of the teacher and the different learning styles of the students
CONERT 10	The IDE has indeed made it possible to meet the requirements of SEN, linguistic disadvantages, etc.
CONERT 11	IDE has favored the balanced combination of synchronous and asynchronous activity

The descriptive statistics of the CONERT battery (tab. 5) show medium-high levels of agreement (3) for items related to the urgency of the RL/IDE, the first where reference is made to the brevity and the extraordinary character of this didactic. The level for the remaining ten is medium-low (2), especially for the eighth concerning the development of transversal skills where there is no level 4, and the eleventh concerning the satisfaction of the needs of SEN students, etc. where the average is less than 2.

Exploratory factor analysis (by Varimax method) shows that the Kaiser-Meyer-Olkin sampling adequacy measure has a good value (.861), the Barlett sphericity test is significant (< .001), and extracted two components that could be defined as beliefs about the emergency of didactics (factor 1, CONERT 1 and 2) and the applicability of IDE as thought by the Department (factor 2, CONERT 3 to 11) that saturate 58.55% of the total variance. The reliability analysis of the scale shows an average Cronbach Alpha (.791); the correlation matrix between the elements has a negative correlation between the first two items and the other nine, plus Cronbach Alpha if removed at a good level (.838). For these reasons, the scale is divided into two distinct factors: CONERT and APPDDI. The descriptive statistics of the two scales identified confirm a medium-high level (3) of agreement with the emergency status of RL/IDE and a medium-low level on the effective applicability of IDE; the difference is about one point. This situation was also noted in the interviews and documents analyzed.

The descriptive statistics of the PRAXERT battery (tab. 6) show high agreement levels (4) for all items except the second where reference is made to the demand and monitoring of individual or grouped digital processing.

Tab. 6 – PRAXERT scale

Variable code	*Variables related to the hypothesized theoretical construct (PRAXERT battery)*
PRAXERT 1	I used e-learning and/or videoconferencing platforms
PRAXERT 2	I requested and monitored the processing of individual or group digital material
PRAXERT 3	Following the ministerial and college instructions I gave lessons of the appropriate duration (ca 45 min.), and I planned appropriate breaks
PRAXERT 4	Following what has been defined by the faculty, I have adopted the shared criteria for the design/implementation of RL/IDE
PRAXERT 5	Following what was defined by the teachers' college I adopted the organizational elements shared for the RL/IDE (duration of lessons, tools to use, etc.)

Exploratory factor analysis (using the Varimax method) shows that the Kaiser-Meyer-Olkin sampling adequacy measure has sufficient value (.707),

the Barlett sphericity test is significant (< .001) and extracted two components that could be defined as the use of tools and the production of digital material (factor 1, PRAXERT 1 and 2) and the agreement with the college of teachers (factor 2, CONERT 3 to 5) which saturate 76.86% of the total variance. The reliability analysis of the scale indicates an average value of Cronbach's Alpha (.748) and if items are deleted, the value will decrease anyway except for item 5 but the difference is .003. The descriptive statistics of latent factors confirm a medium-high level (3-4) of practices regarding the use of digital tools/materials and compliance with ministerial/college guidelines; more specifically, there is a higher average for seconds than for the first.

4.5. Study of correlations and effects of sociodemographic variables

As seen in the literature (Goffman, 1973; Peterman, 1993), in general terms, it can be said that in the three cases considered, there is a correlation between teachers' beliefs and statements of practice concerning assessment and the other constructions involved in the ERT and socio-demographic and contextual variables (tab. 7).

Tab. 7 – Correlations between LF and other variables

	Latent factors (ERT)	*Other variables*
Beliefs	IDE applicability	Formative assessment teacher actions (.190*, p .022) – Beliefs on constructivist learning (.204*, p .014)
Practices	Use of digital tools and materials	Beliefs on formative assessment (.184*, p .03) – Formative assessment teacher actions (.304**, p < .001) – beliefs on engagement (.279**, p < .001) – teacher practices for engagement (.307**, p < .001) – practices on constructivist learning (.316**, p < .001)
	Compliance with the ministerial indications	Beliefs on formative assessment (.180*, p .033) Formative assessment teacher actions (.291**, p < .001) – beliefs on engagement (.205*, p .014) – teacher practices for engagement (.272**, p < .001) – practices on constructivist learning (.258**, p. 002)

Summative assessment beliefs correlates with age (-.171*, p .027), seniority of service (-.184*, p.017), scholastic order (-.184*, p .017); formative assessment beliefs with summative assessment beliefs (-.199*, p .011), assessment actions of the teacher (.320**, p < .001); emergency of RL/IDE with the applicability of the IDE rules (-.204*, p .014); compliance with the ministerial indications with the use of digital tools (.330**, p < .001).

Tab. 8 – Analysis of effects

Outcome	*Grouping variable*	χ^2	*df*	*p*	*Significant pairwise (Dunn-Holm)*
CONVALF (Formative beliefs)	SCHOOL	8.23	2	.016	School 3 < School 1 (p = .014)
CONVALS (Summative beliefs)	ORDER	5.76	2	.050	High school > Primary (p = .050)
PRAXAPP (Learning practices)	ORDER	7.63	2	.022	Primary > High school (p = .025)
PRAXENG_TEACH (Engagement practices)	ORDER	7.64	2	.022	Primary > High school (p = .021)

Note. Kruskal-Wallis χ^2 values are reported with rounded test statistics. Pairwise comparisons are Dunn's tests with Holm correction. Only significant contrasts are shown.

Descriptive analyses indicated that teachers' beliefs and practices were generally stable across age and seniority groups, with median values consistently around 3 on the Likert scale and narrow interquartile ranges, suggesting homogeneous distributions. However, Kruskal-Wallis tests identified significant effects of school context and scholastic order for several constructs. Formative assessment beliefs (CONVALF) varied significantly across schools, $\chi^2(2) = 8.23$, $p = .016$, with Dunn *post-hoc* tests showing that teachers in School 3 reported significantly lower formative beliefs than those in School 1 ($p = .014$). Summative assessment beliefs (CONVALS) also differed by scholastic order, $\chi^2(2) = 5.76$, $p = .050$, with *post-hoc* analysis indicating stronger endorsement among high school teachers compared with primary school colleagues ($p = .050$). In addition, scholastic order significantly affected both constructivist learning practices (PRAXAPP), $\chi^2(2) = 7.63$, $p = .022$, and engagement practices (PRAXENG_TEACH), $\chi^2(2) = 7.64$, $p = .022$. Dunn's *post-hoc* tests revealed that primary teachers reported significantly higher levels of both learning practices ($p = .025$) and engagement practices ($p = .021$) compared with high school teachers. Altogether, these findings suggest that scholastic order – and, to a lesser extent, school context-represents the most relevant contextual factor in shaping teachers' beliefs and practices during Emergency Remote Teaching (tab. 8).

4.6. Some results from ordinal logistic regression

We estimated proportional-odds models for each construct with age (AGE), seniority of service (SENSERV), school (ID_SCUOLA), and scholastic order (ORDER) as predictors (tab. 9). For formative assessment beliefs (CONVALF) the model indicated meaningful between-school and scholastic-order effects: compared with the reference school, teachers in School 3 showed lower odds of endorsing higher categories on CONVALF ($\beta = -1.731$, OR = 0.18, $p < .001$), and teachers in upper-secondary (ORDER3) also had lower odds than the reference order ($\beta = -1.194$, OR = 0.30, $p = .0057$); AGE and SENSERV were not significant. The model's Nagelkerke $R^2 = 0.154$ suggests small-to-moderate explanatory power. For summative assessment beliefs (CONVALS), both School 3 ($\beta = 1.207$, OR = 3.34, $p = .0083$) and upper-secondary ($\beta = 1.315$, OR = 3.72, $p = .0013$) had higher odds of endorsing higher categories, with $R^2 = 0.200$. For engagement beliefs (CONENG), only seniority group 4 (14-25 years) was significant ($\beta = 2.192$, OR = 8.96, $p = .0259$), while other predictors were not; $R^2 = 0.135$. For constructivist learning beliefs (CONAPP), teachers in AGE3 (36-50 y.o.)

(β = -3.688, OR = 0.03, p = .0085) and AGE4 (51-60 y.o.) (β = -3.947, OR = 0.02, p = .0071) had lower odds than the reference age group; in addition, School 3 (β = -1.311, OR = 0.27, p = .0069) and upper-secondary (β = -1.350, OR = 0.26, p = .0015) were associated with lower odds; $R^2 = 0.207$. For engagement practices (PRAXENG_TEACH), School 3 (β = -0.911, OR = 0.40, p = .046) and upper-secondary (β = -1.260, OR = 0.28, p = .0021) showed lower odds than their reference categories; $R^2 = 0.117$. Finally, for learning practices (PRAXAPP), both lower-secondary (ORDER2) (β = -1.091, OR = 0.34, p = .043) and upper-secondary (ORDER3) (β = -1.364, OR = 0.26, p = .0045) were associated with lower odds relative to the reference order; $R^2 = 0.098$. Taken together, the results indicate that scholastic order (especially upper-secondary vs. the reference order) and, for some outcomes, school context are consistent correlates of teachers' beliefs and practices, whereas age and seniority exert limited or outcome-specific effects.

Tab. 9 – Significant predictors from ordinal logistic regressions

Outcome (DV)	*Predictor*	*β*	*OR*	*p*	*Nagelkerke R²*
CONVALF (Formative beliefs)	ID_SCUOLA3	-1.731	0.18	< .001	.154
	ORDER3	-1.194	0.30	.006	
CONVALS (Summative beliefs)	ID_SCUOLA3	1.207	3.34	.008	.200
	ORDER3	1.315	3.72	.001	
CONENG (Engagement beliefs)	SENSERV4	2.192	8.96	.026	.135
CONAPP (Learning beliefs)	AGE3	-3.688	0.03	.009	.207
	AGE4	-3.947	0.02	.007	
	ID_SCUOLA3	-1.311	0.27	.007	
	ORDER3	-1.350	0.26	.002	
PRAXENG_TEACH (Engagement practices)	ID_SCUOLA3	-0.911	0.40	.046	.117
	ORDER3	-1.260	0.28	.002	
PRAXAPP (Learning practices)	ORDER2	-1.091	0.34	.043	.098
	ORDER3	-1.364	0.26	.005	

Note. β = logit coefficient; OR = odds ratio (exp(β)); p = Wald test probability. Nagelkerke R^2 reported per model. Reference categories correspond to the baseline factor levels set in the dataset.

All these results seem to corroborate what was observed by reading the INVALSI 2021 data.

5. Discussion

On the basis of all the data collected, analysed and triangulated during the research (Giganti, 2025), several considerations can be made. The socio-economic level of the territory to which the schools belong, and the characteristics of the institute have been decisive, because where there are more human, instrumental, and economic resources, the smaller size, and stronger relationships have enabled more processes of change and innovation to be activated than other more fragmented, less cohesive, and resource-poor areas, despite the efforts of teachers and school leaders.

It is interesting to note that the age and experience of teachers are a conditioning factor as well as the order of teaching and the emergency. More experienced teachers do not need to use summative assessment to control the class and gain student respect. Similarly, primary school was the framework in which it was possible to offer formative assessment, a constructivist type of learning and to involve pupils more than in other orders, because of the different training of teachers, the reform of assessment introduced in a pandemic period and, as the data from the focus groups show, the tacit belief that older students did not need special support or attention. The school of belonging also influenced the beliefs of formative assessment and learning: the institutes numerically smaller, stronger identity with coaching and teacher training practices have been decisive on agreed averages. In all three schools, however, the management of integrated digital education has been complex but seen as an opportunity to innovate or improve certain teaching practices and to give new impetus to teaching professionalism.

From several points of view (open responses in the questionnaire, reflections emerging in the focus groups) it emerges that convictions and practices are difficult to change, especially if they are conditioned by the experience of the students and because of short and urgent delays. Training is seen as useful for this purpose but often too far removed from the daily practice of teachers, weak in its applicability in didactics, and ill-suited to profound transformation. In this regard, teachers report that they have not received adequate initial training on the subjects covered by the research and on assessment, apart from those in primary education with a university degree, and on digital education; continuing training is also judged in the same way as it was during the pandemic period.

INVALSI data has been fundamental to open questions about the school context and to start a comparison between institutes. At the same time, there were great difficulty for the researcher to access the data of the schools (difficulty for the schools to share, to find or to understand the needed data).

At this stage, it is also useful to comment on the statistical approach adopted.

In line with the INVALSI findings, the non-parametric analyses confirmed significant differences across schools and school levels, suggesting that contextual factors play a substantial role in shaping teachers' beliefs and practices regarding formative assessment. Nevertheless, the ordinal nature of the data requires interpretative caution: results should be understood as coherent tendencies rather than as precise measurements. The integration of quantitative and qualitative sources (interviews, focus groups, and document analysis) enhances the internal validity of the study and provides a more nuanced understanding of the dynamics at play, which numerical evidence alone cannot fully capture.

6. Researcher and Principal considerations

From the researcher point of view, it is possible to consider that an evolutionary process of assessment happened: from the confusion and replication of traditional methods teachers and schools have come to a functional approach. This had consequences on educational redesign in terms of major quality and minor quantity. Teaching practices developed according to the learning paradigm and students' engagement has been supported. The dissonances between implicit and explicit raise questions about the effectiveness of change models. Changes occasioned from contingent situations (such as pandemic) are not sufficient for lasting change. It is necessary to foster the development of paths consistent with the needs and beliefs of implicit/explicit training, at an appropriate time and within the communities of practice.

From the principal point of view. On 1st September 2021 the official report of the INVALSI test results (taken in spring 2021) was released. The accompanying letter from the Institute underlined: «Especially in this school year, the data is returned to the schools at the beginning of September to encourage a wider use of it in supporting the organization of the teaching activities in the school year».

There's no need to remember how hard the first two years into the Pandemic were for all the schools, but it's necessary to draw the attention on the different approaches to teaching and learning: in 2020, DAD (that is Emergency Remote Teaching) was the response to lockdown, in 2021, DDI (that is Integrated Digital Teaching) was the only way to face the continued use of quarantines. In the first case, the main obstacle was physical distancing, in

the second one the continuous alternation of setting between remote learning and presence at school of part of the students.

The INVALSI data arrived in a moment full of different emotions, but School staff experienced this moment with feelings of anxiety and fear, rather than confidence and hope.

General results in our school showed elements of fragility and uncertainty in different grades (for instance, in 2nd grade with regard to the Italian test 4 classes out of 6 were below the national average).

Frustration among the teachers, in particular among the most involved in the innovation process, was made evident by the words: «It was all in vain! […] Our resilience has been unproductive» and so on.

For sure, we focused all our efforts on reaching the aim of improving our students' skills and competencies. But, at this point, I can't fail to point out, here, that there's among teachers some difficulty in reading and interpreting the results: lack of skills in reading complex data make teachers continue to be limited to general results and it stresses a general tendency to be self-centered.

The lack of skills in reading more complex data makes the analysis of the results limited to the mere acknowledgement of the so-called overall scores, easily recognizable through the use of the colors of the traffic lights.

The attempt of going deeper into the reading of the analytical data often leads to finding justifications for individual cases known to the teacher.

Let me conclude with a modest proposal: in addition to training and professional development for teachers, could a task force of experts be created to support schools in a critical reading of their test results?

7. Conclusion

The deepening of constructs has allowed the integration of theoretical literature with empirical studies. In addition, the examination of beliefs and statements of practices provided knowledge to reflect on pre-service and in-service training practices and models.

There are three lines of reflection and development. Teachers need implicit/explicit based training and refresher courses, the feasibility of which must be considered on a large scale, but this is an effective direction that should be pursued. Researchers should deepen teachers' beliefs and practices to design effective training paths; furthermore, they should adopt a rigorous, conscious, and critical use of data also considering their social implications; greater interaction between public research institutions (e.g. INVALSI) and university research centers should be encouraged. Policymakers should val-

idate this type of educational research, help scientific research exit from the retreat on itself, aggregate economic, human, and material resources, stimulate work by geographical areas, and involve philanthropic organizations, enterprises, and the third sector.

In conclusion, the real way to prepare for the extraordinary is to substantiate the ordinary; well-trained and competent teachers are a real resource for the school to fulfil its social mandate.

References

Balduzzi L., Vannini I. (a cura di) (2008), *Nuovi insegnanti per una scuola nuova? Un'indagine tra i docenti formati alla Scuola di Specializzazione all'Insegnamento Secondario (SSIS) dell'Università di Bologna*, Clueb, Bologna.

Black P., William D. (1998), *Inside the Black Box: Raising Standards Through Classroom Assessment*, Phi Delta Kappan, Arlington.

Bloom B.S. (1971), *Handbook of formative and summative evaluation of student learning*, McGraw-Hill, New York.

Brown A.L., Campione J.C., Webb L.S., McGilly K. (1992), "A new look at assessment and instruction", in B.R. Gifford, M.C. O'Connor (eds.), *Changing Assessment Alternative views of aptitude, Achievement, and instruction*, Kluwer Academic, Boston.

Bruschi A. (1999), *Metodologia delle scienze sociali*, Bruno Mondadori, Milano.

Carifio J., Perla R. (2008), "Resolving the 50-year debate around using and misusing Likert scales", *Medical Education*, 42 (12), pp. 1150-1152.

Chen Z., Jiao J., Hu K. (2021), "Formative assessment as an online instruction intervention: Student engagement, outcomes, and perceptions", *International Journal of Distance Education Technologies*, 19 (1), pp. 1-16.

Ciani A., Vannini I. (2017), "Equità e didattica. Validazione di scale sulle convinzioni di insegnamento democratico", *Cadmo*, 2, pp. 5-32.

Dewey J. (1949), *Scuola e società*, La Nuova Italia, Firenze.

de Winter J.C.F., Dodou D. (2010), "Five-point Likert items: t test versus Mann–Whitney–Wilcoxon (addendum added October 2012)", *Practical Assessment, Research, and Evaluation*, 15 (1), 11.

Duverger M. (1961), *Méthodes de sciences sociales*, Presses Universitaires de France, Paris.

Engzell P., Frey A., Verhagen M.D (2020), *Learning Inequality During the Covid-19 Pandemic*, retrevied on December 5, 2024, from https://ideas.repec.org/p/osf/socarx/ve4z7.html.

Floden R.E. (2002), "Research on effects of teaching: a continuing model for research on teaching", in V. Richardson (ed.), *Handbook on research on teaching*, AERA, Washington DC, 4th ed., pp. 3-16.

Fredricks J.A. (2015), "Academic engagement", in J. Wright (ed.), *The International Encyclopaedia of Social and Behavioral Sciences*, Elsevier, Amsterdam, 2nd ed., vol. 2, pp. 31-36.
Giganti M. (2025), *La valutazione formative: risorsa per l'apprendimento. Cosa ci ha insegnato l'emergenza?*, FrancoAngeli, Milano.
Goffman E. (1963), *Behavior in public places: notes on the social organization of gatherings*, Free Press of Glencoe, New York.
Harpe S.E. (2015), "How to analyze Likert and other rating scale data", *Currents in Pharmacy Teaching and Learning*, 7 (6), pp. 836-850.
Hodges C., Moore S., Lockee B., Trust T., Bond M. (2020), *The Difference Between Emergency Remote Teaching and Online Learning*, retrieved on December 5, 2024, from https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning.
Hodges C.B., Barbour M.K. (2021), "Assessing learning during Emergency Remote Education", *Italian Journal of Educational Technology*, 29 (2), pp. 85-98.
Hodges C.B., Kim C. (2010), "Email, self-regulation, self-efficacy, and achievement in a college online mathematics course", *Journal of Educational Computing Research*, 43 (2), pp. 207-223.
Kuh G.D. (2009), "What Student Affairs Professionals Need to Know about Student Engagement", *Journal of College Student Development*, 50 (6), pp. 683-706.
Norman G. (2010), "Likert scales, levels of measurement and the 'laws' of statistics", *Advances in Health Sciences Education*, 15 (5), pp. 625-632.
Peterman F.P. (1993), "Staff development and the process of changing: A teacher's emerging constructivist beliefs about learning and teaching", in K. Tobin (ed.), *The practice of constructivism in science education*, Routledge, New York-London, pp. 227-245.
Piaget J. (1976), *Il linguaggio e il pensiero del fanciullo*, Giunti Barbera, Firenze.
Richardson V., Placier P. (2002), "Teacher Change", in V. Richardson (ed.), *Handbook of research on teaching*, AERA, Washington DC, 4th ed., pp. 905-947.
Rositi F. (1971), *Contraddizioni di cultur*a, Guaraldi, Rimini.
Scriven M. (1967), "The methodology of Evaluation", in R. Staker (ed.), *Perspectives of Curriculum Evaluation*, Rand McNally, Chicago.
UNESCO (2020a), *Covid-19 Education Response. Distance learning strategies in response to Covid-19 school closures*, Issue note, 2.1.
UNESCO (2020b), *Ensuring effective distance learning during Covid-19 disruption Guidance for teachers*, retrieved on December 5, 2024, from https://unesdoc.unesco.org/ark:/48223/pf00 00375116.
UNESCO (2020c), *Covid-19 Education Response Webinar. Ensuring quality of learning and well-being for young children in the context of the Covid-19*, retrieved on December 5, 2024, from https://unesdoc.unesco.org/ark:/48223/pf0000374005.
UNESCO (2020d), *Covid-19 Education Response Webinar. Managing high-stakes exams and assessments during the Covid-19 pandemic*, retrieved on December 5, 2024, from https://unesdoc.unesco.org/ark:/48223/pf0000373247.

UNESCO (2020e), *What have we learnt? Overview of findings from a survey of ministries of education on national responses to Covid-19*, retrieved on December 5, 2024, from https://unesdoc.unesco.org/ark:/48223/pf0000374702.
Vygotskij L.S. (1978), *Mind in Society: The development of higher psychological processes*, Harvard University Press, Cambridge.
Vygotskij L.S. (1987), *Studi sulla storia del comportamento*, Giunti, Firenze.
Vygotskij L.S. (2001), *Pensiero e linguaggio*, Laterza, Roma-Bari.
Weeden P., Winter J., Bradfoot P. (2002), *Valutazione per l'apprendimento nella scuola*, Erikson, Trento.
Wiliam D., Thompson M. (2007), "Integrating assessment with instruction: what will it take to make it work?", in C.A. Dwyer (ed.), *The future of assessment: shaping teaching and learning*, Lawrence Erlbaum Associates, Mahwah, pp. 58-82.

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

by Sara Mori, Jessica Niewint, Francesca Storai*

This study, conducted through the Erasmus AI4T project, explores the integration of artificial intelligence (AI) in education. The aim is to improve teachers' ability to use AI tools in the classroom and to self-assess their competences. The project promotes an approach that combines technology with practical, applicable teaching methodologies. The goal is to develop a useful framework for AI literacy among teachers and managers to improve awareness of these issues. Through semi-structured interviews with teachers and head teachers, the study explores the specific training needs of teachers and the potential of AI to enrich educational practices in the classroom and school. The results highlight teachers' positive attitude towards AI, emphasizing the importance of practical training materials and case studies. The study emphasizes the need for continuous professional development and the ethical use of AI in education, confirming the critical importance of digital and AI literacy. There is a clear need for comprehensive training modules and ethical guidelines to be developed in future to ensure the equitable and effective integration of AI in teaching and learning. A community of practice emerges as an ideal setting in which to develop expertise in this field, engage with colleagues, and enhance one's professional awareness.

Lo studio esplora l'integrazione dell'intelligenza artificiale (IA) in ambito educativo, in particolare attraverso il progetto Erasmus AI4T, con l'obiettivo di migliorare le capacità degli insegnanti di utilizzare gli strumenti di IA in classe e di autovalutare le proprie competenze. Il progetto promuove un approccio

* The paper was collaboratively designed by the authors. Specifically, Sara Mori wrote paragraphs 3, 4 and 5; Jessica Niewint wrote paragraphs 1 and 2, as well as the conclusions; and Francesca Storai wrote paragraphs 6 and 7, as well as the statistical analyses.

che coniughi la tecnologia con metodologie didattiche pratiche e applicabili in classe. L'obiettivo è stato quello di sviluppare un quadro di riferimento utile all'alfabetizzazione all'IA tra i docenti e i dirigenti, che migliorasse la consapevolezza rispetto a questi temi. Attraverso interviste semi-strutturate con insegnanti e dirigenti, lo studio approfondisce le esigenze di formazione specifiche e il potenziale dell'IA per arricchire le pratiche educative in classe e a scuola. I risultati evidenziano un atteggiamento positivo nei confronti dell'IA tra i docenti, sottolineando l'importanza di materiali formativi pratici e di casi di studio. Lo studio sottolinea la necessità di uno sviluppo professionale continuo e l'uso etico dell'IA in educazione, confermando la necessità critica di un'alfabetizzazione digitale e all'IA nel settore dell'educazione. Emerge la necessità per il futuro di sviluppare moduli di formazione completi e linee guida etiche per garantire un'integrazione equa ed efficace dell'IA nell'insegnamento e nell'apprendimento. La comunità di pratiche emerge come luogo privilegiato in cui sviluppare competenze in questo campo, confrontarsi con i colleghi e migliorare la consapevolezza rispetto alla propria professionalità.

1. Introduction

The integration of technology in education, particularly since the onset of the pandemic, is marking a significant shift in how teaching and learning processes are perceived and executed (Molinaar, 2021), making it even more necessary to ask how technology can improve teaching practice (Hattie, 2023, p. 408). Technologies like artificial intelligence (AI) are a powerful tool in this ongoing transformation (Selwyn, 2022) where areas like pedagogy, knowledge and societal needs converge (Pedro *et al.*, 2019).

This presents both, opportunities, and challenges for educational practices in applying the potential of emerging technologies to meet the needs of teachers and students. However, adapting systemically educational interventions to address these needs encounters often difficulties due to the lag in policy responses to technological progressions (Zawacki-Richter *et al.*, 2019; Korinek *et al.*, 2021).

The integration of AI into educational settings is seen as a promising step to revolutionize teaching methodologies, enhance learning experiences, and provide substantial support to educators and to transform pedagogical practices to the ethical considerations that go with its deployment (Miao *et al.*, 2021; Carvalho, 2022). As AI systems become more present in our daily lives, the necessity for a deeper understanding of these technologies and their implications becomes increasingly important (Fjelland, 2020).

The conversation around AI literacy must address the dual aspects of empowering educators and learners while ensuring ethical use, transparency, and accountability in AI applications (Miao *et al.*, 2021).

Frameworks as the TPACK (Mishra and Koehler, 2006) frequently has been adapted to the presence of AI technolgoy in educations (Mishra *et al.*, 2023). The framework describes three major teacher-knowledge areas: content (CK), pedagogy (PK), and technology (TK) with their intersections of pedagogical content knowledge (PCK), technological pedagogical knowledge (TPK), and technological content knowledge (TCK). Additional also the awareness of ethical (Celik, 2023) or societal influences (Crompton *et al.*, 2024) is required by teachers knowledge areas to successfully adapt AI in educational settings to have the «capacity to critically adapt and creativity explore innovative practices in the context of advancing capabilities of emerging AI iterations» (Miao and Mutlu, 2024, p. 23).

The impact of AI on pedagogical practices underscores the need for transversal human skills as creativity, complex problem-solving, and collaboration; which are essential for interacting with AI in various spheres of life (Carvalho, 2022).

In the vision in which AI technologies are seamlessly integrated into the educational context, it is imperative to address the challenges of accessibility and inclusivity. Ensuring that the use of AI benefits all students needs and to overcome the digital divide by investing in infrastructure and promoting digital literacy. Integrating AI tools in the educational practice that are accessible and responsive to the diverse needs of learners is critical in fostering an inclusive educational environment (EC, 2023). To do so, teachers need not only the necessary digital literacy but also how to combine this knowledge with the right pedagogical approach and content knowledge to integrate AI in classroom practices (Ning *et al.*, 2024).

2. AI for teachers

The use of AI in the classroom can be divided into three main areas: AI for, with and about AI (Niewint-Gori, 2023). Each of these domains demonstrates a different aspect of the integration of AI into educational settings and requires a variety of skills and competencies from teachers and learners (EC, 2023).

“Teaching for AI” focuses on building broader competencies in teachers and learners as a foundational aspect of incorporating AI into education. This approach promotes confident, critical and safe engagement with AI technol-

ogies, emphasising support for students' development and well-being, and ensuring inclusion, fairness and the protection of students' data and privacy (EC, 2022). The main goal is to introduce the definition of AI and its multiple applications across various fields, thereby fostering AI literacy and digital citizenship among students. This involves understanding not only the basic principles of machine learning, but also the ethical considerations and potential biases of AI (Holmes *et al.*, 2022).

The aim of teaching with AI is to enhance the teaching and learning process by emphasising the importance of teachers' knowledge of methodology, content, and technology, as well as their ability to utilise these aspects to improve teaching and learning processes (Chaipidech *et al.*, 2022; Ning *et al.*, 2024). This involves using AI for personalised learning, shifting the focus from a traditional one-size-fits-all approach to a more tailored educational experience (Calvani, 2020; Mori *et al.*, 2022). Digital educational tools empowered by AI promote self-regulated learning, helping students identify and adopt more effective learning strategies (Molenaar, 2021). Furthermore, these tools have the potential to align learning experiences with students' values and personal interests, thereby significantly improving engagement and outcomes (Walkington and Bernacki, 2019).

Integrating simulations and games makes complex systems tangible for students, fostering a deeper understanding of academic concepts (Asbell-Clarke, 2020).

Education about AI covers the technical foundations of this technology, including coding and machine learning. This subject area is crucial for equipping students with the skills they need to enter the labour market.

Teaching about AI focuses on the fundamentals and requires in-depth technical knowledge, including programming, mathematics and machine learning. This approach prepares students for the labour market and fosters a developer's perspective on AI (Holmes *et al.*, 2019; Markauskaite *et al.*, 2022). It is important to understand the fundamentals of AI, its ethical and social impact, and the practical application and development of AI technologies, including managing data cycles, the ethical challenges posed by AI and applying AI to solve real-world problems. Additionally, it involves understanding AI's broader societal implications and its applications beyond traditional computing fields, thereby fostering interdisciplinary knowledge and skills (Bellas *et al.*, 2022).

2.1. AI competences for teachers

Integrating AI into education requires educators to have a wide range of skills (see Table 1). A variety of teacher competencies and areas of knowledge are required for the responsible and effective use of AI in schools, and this list is growing. For example, UNESCO's competency frameworks for teachers (Miao and Mutlu, 2024) define what educators should understand about AI and what learners should be taught "about AI" across structured dimensions such as mindset, ethics, pedagogy and professional learning. The OECD's *Digital Education Outlook 2023* provides practical "guidelines and guardrails" for transparency, fairness, data protection and accountability, which practitioners can use when selecting and using systems (OECD, 2023). The question of what constitutes meaningful AI integration in classroom activities is also receiving increasing scientific attention. For instance, reviews report meaningful learning gains from AI-enabled feedback and adaptation. Automated writing feedback, for example, has a medium effect on writing performance ($g \approx 0.55$), while adaptive systems have a positive impact on cognitive outcomes, although design and contextual factors must be considered (Kim, 2025). Furthermore, ethics and safety are considered fundamental teacher responsibilities (EC, 2022), as well as the role of teachers in promoting student AI literacy. Collaboration across all stakeholder groups, such as educators, school leaders, developers and policymakers, is repeatedly recommended in order to align AI with pedagogy and equity goals. These AI-specific competencies map cleanly onto established digital competence baselines, such as DigCompEdu, including emerging AI supplements (US Department of Education).

At the same time, practice is clearly outpacing the scientific evidence base on generative AI in education, sending strong signals about gaps in teacher preparedness and ethics/safety concerns (Petrucco *et al.*, 2025). Classroom experiences now position GenAI as a dialogue partner and a tool to support self-regulated learning. However, it is important to keep the teacher involved to ensure human oversight (Tang *et al.*, 2024).

The focus of ethical awareness is on topics such as discrimination, privacy/autonomy and assessment integrity. It also documents misalignment between providers and deployers, emphasising the need for evaluation protocols, particularly in an educational context (Harvey *et al.*, 2025).

Tab. 1 – Skill set for AI in education

Skill	*Description*
Understanding of AI concepts	Gain a basic knowledge of AI principles and applications to understand how AI can enhance educational practices and outcomes The mindset and skills necessary to critically assess AI tools in terms of quality, effectiveness and ethical implications, ensuring that these tools align with educational goals
Critical evaluation of AI tools	The mindset and skills necessary to critically assess AI tools in terms of their quality, effectiveness and ethical implications, ensuring that they align with educational goals
Pedagogical integration skills	This includes the ability to incorporate AI technologies into pedagogy, such as designing AI-enhanced learning activities that promote engagement and personalisation, and using AI for formative assessment and feedback mechanisms
Ethical and safety considerations	The ability to address ethical concerns, including data privacy and potential bias, is essential. It is the responsibility of educators to ensure the ethical use of AI in the classroom
Facilitating AI literacy	Promote the development of AI literacy among students by fostering an understanding of how AI works, its various applications, and its impact on society
Continuous learning and professional development	Due to the rapidly evolving nature of AI, educators must be flexible and committed to ongoing professional development in this field
Collaboration and communication skills	The effective implementation of AI in education often requires a collaborative approach. Teachers need to be able to communicate with and work alongside AI experts, technologists and other stakeholders in order to align AI solutions with educational goals and student needs

2.2. Role of professional development

Professional development (PD) is most effective when it is job-embedded, iterative and directly connected to specific classroom tasks (Darling-Hammond *et al.*, 2017). This type of continuous professional development (CPD) is a cornerstone of teacher professionalism, influencing teaching practices in the classroom while enhancing teachers' self-efficacy and job satisfaction. In contemporary educational contexts, teachers have expressed an increased need for CPD in complex areas such as special educational needs, multicultural and multilingual classroom environments, and the integration of ICT skills for teaching (OECD, 2019).

Teachers are more likely to adopt and sustain innovative practices if professional development provides them with concrete, actionable resources

that align with existing curriculum and assessment frameworks. These resources are most effective when supported by iterative cycles of classroom implementation, targeted feedback and structured reflection (Kraft *et al.*, 2018). In the context of AI integration, ethics content has the greatest impact when it is structured as brief, repeatable classroom routines and implemented through school-wide initiatives. Current international guidance on generative AI and teacher AI competencies emphasises critical principles, including human oversight, transparency, data minimisation and age-appropriate implementation (Miao and Holmes, 2023).

Successful integration of AI into educational settings requires professional development that explicitly addresses the ethical dimensions of AI use. Teachers require comprehensive support, including access to curriculum resources and evidence-based teaching strategies. They also need to be able to share and critically reflect on their experiences of implementing AI in the classroom, and to do so within robust communities of practice (Hutchins *et al.*, 2025). This will ultimately promote effective classroom AI integration (Miao and Yao, 2021).

Teachers' confidence in using AI and digital tools is a critical predictor of successful classroom implementation. Teachers' confidence in using AI significantly influences their integration practices and intentions to continue using it in the classroom (Collie *et al.*, 2024; Kong *et al.*, 2024). This finding is consistent with broader research in the field of educational technology in K-12 settings, where technology-integration self-efficacy has been shown to predict actual technology adoption (Gómez Jr. *et al.*, 2021). Approaches to professional development that prioritise mastery experiences, such as classroom-ready exemplars, structured micro-pilots with coaching support and teacher-in-the-loop routines, correlate with substantial improvements in instructional practice. This provides a mechanism through which enhanced self-efficacy can be translated into sustained implementation (Kraft *et al.*, 2018).

Despite the existence of robust theoretical frameworks, there is limited empirical evidence on how educators navigate AI competency requirements in practice. This is particularly pertinent given the divergent needs of teachers seeking immediate classroom solutions and school leaders managing systemic change. The AI4T project addresses this gap by examining educators' experiences with AI in professional development (PD) and classroom activities.

3. AI4T Projects

The AI4T (Artificial Intelligence for and by Teachers) project, funded by the Erasmus+ programme for European policy experimentation in Education and Training led by high-level public authorities (project no. 626154-EPP-1-2020-2-FR-EPPKA3-PI-POLICY), aimed to raise teachers' awareness of the opportunities and implications of applying AI to teaching and learning processes.

This Erasmus+ Key Action 3 (KA3) research project resulted from collaboration at two levels: internationally, among a network of partners (including research institutes, universities, and European ministries) in France, Ireland, Italy, Luxembourg, and Slovenia, engaged in innovation and evaluation in education; and nationally, through a design group involving the Italian Ministry of Education and Merit (MIM), INDIRE (National Institute for Documentation, Innovation and Educational Research), and CNR-ITD (Institute for Educational Technology of the National Research Council of Italy) in Palermo. France Éducation international (formerly CIEP, Centre International d'Études Pédagogiques) requested collaboration from the research group in light of their previous experience in evaluating school improvement and innovation processes at organisational and didactic levels. The project's framework was based on earlier European studies examining the importance of AI-driven innovation in learning environments and its potential impact on teaching, learning, and education systems.

In the Italian context, CNR delivered the training for teachers, while INDIRE conducted the impact evaluation. The training was provided remotely via a Moodle platform and included a shared interactive book (translated into the respective languages) and video resources for teaching tools, as well as synchronous group discussions, across partner countries.

The AI4T project was structured as a mixed-methods evaluation study with experimental and control groups, considering two measurement points (pre and post) for both conditions. Questionnaires were administered to teachers at the beginning and end of the training and to principals and students only at the end.

The project involved 91 upper secondary schools (classes with pupils aged 15 to 17), which were selected by MIM across the country according to territorial distribution and school type criteria. A total of 104 school principals, 443 STEM and English teachers, and 1,554 students were also involved.

Implementation of the project took place between February 2021 and February 2024, comprising an initial pilot phase followed by a second, larger-scale administration phase with the selected sample.

4. Research aims and questions

This study provides an in-depth analysis of the broader data collection planned for the AI4T project. As previously mentioned, the AI4T project used a counterfactual design to evaluate the impact of the training programme on the use of AI in the classroom. This study aims to provide a qualitative analysis of interviews with teachers and head teachers from schools that underwent training (the experimental group). Exploring the opinions and attitudes of teachers and managers who participated in AI-specific training is believed to provide useful information for redesigning and developing future training courses on the subject. It will also be interesting to highlight any differences in the views of teachers and headteachers regarding training needs and the use of AI in education.

Particular attention has been devoted to analysing opinions on which aspects of the training were perceived as effective or, conversely, less useful by these key stakeholders. Teachers and school leaders are uniquely positioned to identify the programme's strengths and areas requiring revision, as their evaluations stem directly from classroom practice and organisational responsibilities (Damiano, 2013; Domenici, 2015). This type of reflective contribution is valuable because it moves beyond mere satisfaction assessment to identify conditions that enable or hinder the effective integration of AI into teaching (Calvani, 2011).

The hypothesis is that teachers and head teachers who have participated in a training course will be able to provide meaningful insights on a new and increasingly necessary topic for schools. These insights are important for understanding not only the perceived usefulness of the training, but also which components support professional growth and which are less relevant or applicable (Lucisano and Salerni, 2002).

The study addresses the following questions:

– Q1: Do the insights gained from the experiences of teachers and DSs through participation in the AI4T project enable the identification of specific elements for artificial intelligence training activities and tools?
– Q2: Could participation in the AI4T project provide suggestions for strengthening one's skills in this area?

Being able to answer these questions will be useful in terms of both furthering the impact assessment and informing further redesigns, especially in the Italian context.

5. Methodology

5.1. Participants

At the conclusion of the training course, one school principal and three teachers participated in the study. All participants were female. The principal was employed in a school located in Northern Italy, while the teachers worked in three different schools situated respectively in the North, Center, and South of the country. Two of the teachers specialized in mathematics and one in English language instruction. The mean age of the participants was 47.3 years (SD = 3.7).

These participants represented a subset of the experimental group. They were selected on the basis of both geographical distribution and subject specialization, with the aim of ensuring the inclusion of perspectives from different categories relevant to the counterfactual evaluation study (i.e., schools across Italy and teachers of scientific subjects or English language). Although the participants cannot be defined as representative of the entire population, they were among those who attended almost the entire course and participated in the planned activities. For this reason, they were also deemed suitable for participation in the interviews.

5.2. Tools

A semi-structured interview was used for the final evaluation of the course and the tools used for the evaluation, due to its usefulness in the field of evaluation and in the educational and social spheres (Ciucci, 2012).

Interviews allow researchers to grasp the complexity of individual experiences and perceptions. Unlike surveys or questionnaires, which may limit responses to predefined options, interviews can reveal nuances and details that may not have been anticipated by the researcher (Bryman, 2012). The flexible nature of interviews also allows for the exploration of unexpected topics or ideas that arise during the conversation, adapting questions based on the interviewee's responses. This adaptability can lead to a richer and more comprehensive understanding of the topic (Kvale, 1996). The depth of information gathered through interviews can be significantly greater than that typically obtained through other qualitative methods. The detailed descriptions and stories collected can provide a rich set of data for analysis, allowing for an in-depth examination of the phenomenon under study (Corrao, 2005; Della Porta, 2014).

The interview with teachers included the following sections:

- specific training requirements for AI;
- the role of training in self-assessment of individual skills;
- perceptions of AI application in the classroom and potential areas for growth;
- key perceived ethical risks;
- suggestions for implementing AI in the classroom.

The interview with the Principal also explored the systemic requirements for implementing AI in the classroom and the motivation perceived by teachers to engage with what they learned during the training course.

The tools used for the impact assessment of the AI4T project, which, as mentioned, used a counterfactual research design to explore teachers' professional learning experiences (Guskey, 2000), incorporate the constructs explored in the Technology Acceptance Model (Davis *et al.*, 1989; Kemp *et al.*, 2019).

The constructs discussed in the interviews also used this framework as a backdrop to explore both the elements that were useful to teachers and those that were critical and could be implemented in new ways.

5.3. Procedures and data analysis

The interviews were conducted at the end of the training course (in September- November 2023), within the broader context of impact assessment. Teachers were interviewed online, with the camera on, using Webex software. An interviewer and an observer were present during the interview. This allowed us to collect textual data, but also to note down some reflections that may have enriched the information.

The qualitative data in this study were examined through a primary content analysis approach (Downe-Wamboldt, 1992; Britt-Marie Lindgren *et al.*, 2020), supported using Voyant Tools, a web-based environment for reading and analysing digital texts. In this context, Voyant was used to conduct a systematic exploration of the corpus, focusing on word frequencies and co-occurrences. These functions facilitated the identification of thematic patterns and the organization of textual data in a transparent and replicable manner. While not intended to replace in-depth interpretative analysis, the software provided a complementary layer of evidence, helping to ground qualitative insights in recurring linguistic features.

6. Results

From the corpus of the interviews with the teachers, the most frequently occurring words (occurrences) were analysed, they were: *training* (24) training provided by the project; *videos* (20) training in the MOOC; *guys* (18) word used by teachers indifferently by students; *intelligence* (18) referring to artificial intelligence; *tools* (16) provided by the project.

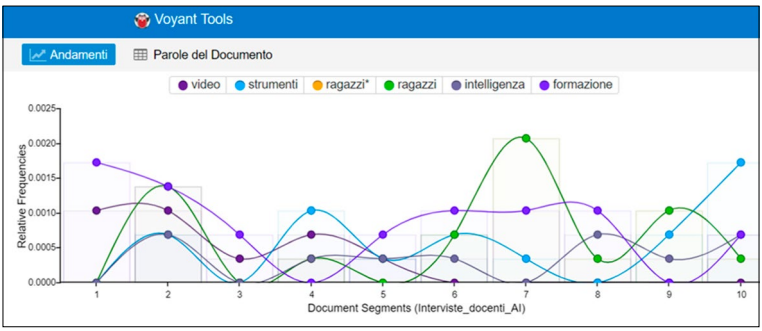


Fig. 1 – Trends and occurrences: teachers

The analysis provided an initial graphical representation of the trend of these words during the interviews time. In the first part of the interview, the one concerning the *AI4T professional learning experience*: the teachers' attention focused on the words training and video. More emphasis was placed on the students, the interview then went on to the second topic related to the *Impact of the professional learning experience on AI knowledge and perception*, where the word training was always highly mentioned. In the last theme *Experience of using AI*, the word most frequently used is *tools*.

6.1. Correlations and significance of the terms

The term student* appears to have a high correlation with the terms *changed* and a strong significance with the term aware (Tab. 2). Extracts from the interviews read: "I have done research and bought books but the students' attitudes towards their use have changed. The students are well aware that they themselves have an e-mail which determines responsibility'. The term student* despite having a high correlation with *teaching resources*, but is not significant (Tab. 3)

Tab. 2 – Correlations and significance terms student and ragazz**

Term 1		*Term 2 Correlation (r)*	*Significance (p)*
Changed	Student*	0.921132400	0.00015376818
Aware	Student*	0.921132400	0.00015376818
Privacy	Student*	0.921132400	0.00015376818
Teaching resources	Student*	0.075377840	0.83604467000
Artificial	Student*	0.072595400	0.84203184000
Coaching	Student*	0.065795160	0.85669450000
Training	Ragazz*	0.063741670	0.86113024000
Content	Ragazz*	0.057486450	0.87466310000
Subject	Ragazz*	0.052999900	0.88438785000
Important	Ragazz*	0.052999900	0.88438785000
Knowledge	Ragazz*	0.035333257	0.92280490000

*Tab. 3 – Correlations and significance terms teacher**

Term 1		*Term 2 Correlation (r)*	*Significance (p)*
Attitude	Teacher*	0.80403024	0.005055611
Deepen	Teacher*	0.67363310	0.032717230
English	Teacher*	0.67363310	0.032717230
Meetings	Teacher*	0.62882817	0.051469225
Change	Teacher*	0.62882817	0.051469225
Share	Teacher*	0.62882817	0.051469225
Example	Teacher*	0.49123332	0.149345760
Create	Teacher*	0.45226702	0.189403670
Teaching practise	Teacher*	0.45226702	0.189403670
Difficulties	Teacher*	0.45226702	0.189403670
Different	Teacher*	0.45226702	0.189403670
Motivation	Teacher*	-0.14997116	0.679217600

«We received an indication from the school managers about the project, our expectations were that there would be an impact on the students in terms of teaching resources of the applications. In my opinion, the accompaniment should have been a real experimentation with a laboratory».

The term teacher* has a high correlation and significance with *attitude*, (Tab. 3). «I informed myself and bought books and my attitude towards using it changed. After the training, I tried to supplement the information I had received a bit and contacted some start-up founders in the area to learn more about Big Data, it all came about in the wake of the course». The correlation with the term *knowledge* was negative and not significant: «At the level of

content exposure, knowledge was done well, but it did not change the way I perceived AI».

*Tab. 4 – Correlations and significance terms tool**

Term 1	*Term 2*	*Correlation (r)*	*Significance (p)*
Training	Tool*	0.83802360	0.0024647545
App	Tool*	0.83581483	0.0025946465
Classroom	Tool*	0.75827795	0.0110250700
Aeas	Tool*	0.74182534	0.0140392175
Applications	Tool*	0.74182534	0.0140392175
Subjects	Tool*	0.74182534	0.0140392175
Colleagues	Tool*	0.71567810	0.0199443800
Can	Tool*	0.68727815	0.0280945800
Intelligence	Tool*	0.50404865	0.1374072100
Feedback	Tool*	0.44899955	0.1930212400

The term tool* appears to have a high correlation with the terms: training, app and classroom, with which it is significant (Tab. 4). «I suggest that the training course should be a little more practical in nature by creating a kit, a manual for the various areas with tools that are dropped into everyday classroom teaching». Tool* has a low correlation with the term feedback with which it is not significant.

6.2. School leaders' views on AI

Similarly to what was done for interviews with teachers for the analysis, we took into account the words that were most frequently mentioned in the corpus of interviews with school heads (occurrences): *schools* (15); *guys* (12) used indifferently with the word students; *important* (10); *teachers* (10); *exemple* (9) as good practices.

The graphic representation (Fig. 2) shows the trend of these words during the interview time. In the first part of the interview, the part concerning *Support for teacher training*, the focus of the school leaders was on the word's *schools* and *important*. The words can be contextualised in the following statements. «It is important to move on to the practical part. By participating in the project, I would like to bring added value to the school and the participating teachers». The second part of the interview was about *Knowledge of AI*, the words *important* and *example* were the ones most used: «As a school we have worked a lot with the Ministry managing the digital school award. I

try to communicate what we do through the website and GSuite so that even teachers who do not participate in these activities are aware of what others are doing and that it serves as an example and good practice». In the last topic discussed with the two school leaders was *AI integration in schools*, the focus is on the students: «It is important that the children understand the importance of these AI tools, but also that they never lose touch with everyday reality, that they don't just experience it as a playful dimension. It is important to include AI projects in the curriculum to involve children, to make them understand what the use and purpose of this tool can be».

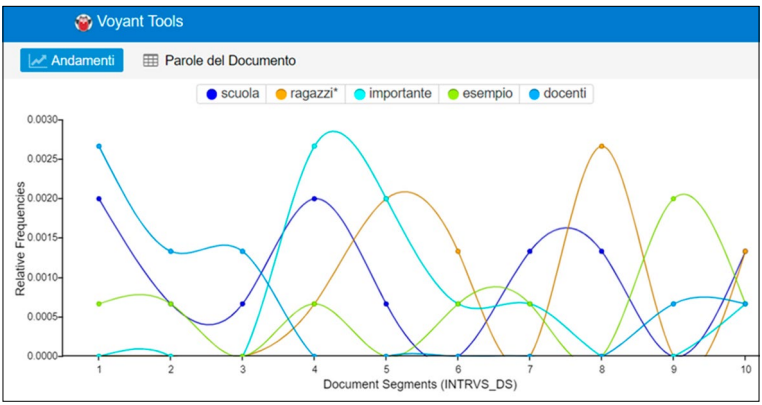


Fig. 2 – Trends and occurrences: school leaders

6.3. The contests: a storytelling

The analyses of words *important* and *example* had received 10 and 9 occurrences respectively. Based on the words on the left and right, a narrative was constructed within the interview's contexts.

As far as the term important is concerned, in the context of the use of artificial intelligence in schools, the focus of the school leaders was on the playful use by students that can be taken as a motivational drive to be transferred into teaching. Another theme concerns reflection on the ethical level, again referring to students, for a conscious use.

The term example* refers mainly to how to include practical examples within curricula to be more interesting, but also to teacher training.

*Tab. 5 – Contexts left/right: important**

Left	*Word*	*Right*
For the student as well is	Important*	understand that what you do online
it is linked to ethical choices, it is	Important*	and could result from this
in the use of data, big data is	Important*	as the playful motivational part
motivational play that I believe is	Important*	and I believe that a school
and whatnot, I think it is	Important*	because the boys on
to artificial intelligences, however, is	Important*	for boys in development
of all these activities	Important*	that young people understand these tools but also that they never lose touch with everyday reality
of a playstation so it is	Important*	to give clearer indications
things. In my opinion it is very	Important*	like curiosity as well. The school
that artificial intelligence is	Important*	and that therefore in some

*Tab. 6 – Contexts left/right: example**

Left	*Word*	*Right*
Interested in future teachers training for	example*	the one on metaverse. The school
are interested groups, for	example*	something has been done on the
and ethical fallout. Also for	example*	on the use of
as in the medical field for	example*	but one should try to include it
all their companions. for	example*	in this sense was
of the situation. This year for	example*	I began by saying that we must
history curricula because for	example*	at most you get to the
that should be dealt with, for	example*	as a student I didn't even
within the school to for	example*	in civic education, is a

6.4. Comparison interview teachers and school leader

The last analysis involved comparing the two interviews, that of the teachers and that of the students.

As Fig. 3 shows, the attention of the school leaders focused more on the words *school* and *students* and less on the words *training* and *tools*. Conversely, the teachers focused their attention on *tools* and *training* and mentioned the words *school* and *students* less. Since the project is based on teacher training, the school leaders' reflective gaze focused mainly on the school as a system dimension and on the students as the final recipients of all

teaching and educational actions, in this case based on the topic of artificial intelligence. The teachers, on the other hand, focused on topics that concern them closely to teaching and their professional development.

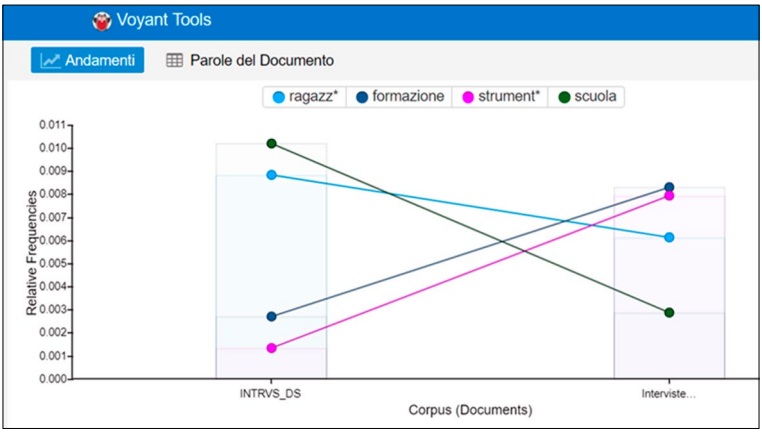


Fig. 3 – Comparision interview: teachers and school leaders

7. Discussion

We will attempt to answer the main question that initiated our investigation below: (Q1) Can the insights gained from the experiences of teachers and DSs through participation in the AI4T project be used to identify specific elements for artificial intelligence training activities and tools?

Having a favourable attitude and a willingness to change is certainly the first step in tackling such an innovative topic as artificial intelligence (Koehler *et al.*, 2014; Taopan *et al.*, 2020). The teachers interviewed seem in favour of this new approach, but they are most concerned about the educational use of AI tools. Although on a very small scale, the survey showed how useful training can be in understanding the opportunities and challenges of AI pedagogy (Liu *et al.*, 2021; Yang, 2022; Çelik, 2023). A technological understanding also enables more conscious addressing of teaching challenges and facilitates understanding of the potential of AI, even if this does not seem sufficient for working with students in the classroom. The teachers interviewed believe that it is important to have material and case examples available for teaching, stating: «It is important to move on to the practical part».

In this sense, one of the main findings is the discrepancy between theoretical and practical training in the classroom. Teachers emphasised that under-

standing "what AI is" is less useful without clear educational applications. This confirms the importance of designing training modules that integrate case studies, practical sessions, and collaborative experimentation so that teachers can acquire and test knowledge immediately in their teaching contexts. Furthermore, the correlations that emerged from the interviews highlight the importance of concepts related to attitudes and awareness (e.g. the relationship between students and awareness). This suggests that training should focus not only on tools, but also on how AI changes the relationships between students, between teachers and students, and the responsibilities of students, who must be aware of, and participate in, the use of AI tools in their learning process.

Our second question was specifically about the knowledge acquired on the AI4T training course and how it could benefit teachers: could the knowledge gained from participating in the AI4T project be used to suggest ways to strengthen one's skills in this area?

To better understand the usefulness of the project, we analysed interviews with teachers, who are direct stakeholders in teaching practice, and with school leaders, who are involved in system-wide technological change that may also affect schools in the future. As in Q1, the teachers' opinions concerned the possibility of bringing concrete examples into the classroom: «By participating in training, I would like to be able to add value to the classroom». Give practical examples of what artificial intelligence does. School leaders focused on how the curriculum could be improved using AI tools, as well as the ethical implications (Gocen and Aydemir, 2020; Shapiro and Stefkovich, 2016). «The topic of ethics is an interesting issue, especially for children, but it needs to be made explicit because they do not grasp it. Students often experience artificial intelligence as a parallel imaginary world».

Teachers focused on their own professional development and AI tools (Celik *et al.*, 2022), as well as how these can be used in the classroom.

This comparison reveals an integrated and complementary vision: teachers require operational tools and methodologies, while school leaders emphasise the importance of systemic change, from integrating AI tools into curricula to encouraging ethical reflection. Therefore, professional development should be designed at two interconnected levels: (1) the micro level, which focuses on classroom teaching practices and provides ready-to-use tools and applications; and (2) the macro level, which focuses on developing digital skills for teachers and students and provides guidelines for the ethical use of data and privacy protection (UNESCO, 2023). Furthermore, the strong correlation between teachers' attitudes suggests that promoting reflective practices and communities of practice could be crucial for long-term skill

development. From this perspective, AI-related professional growth cannot be limited to individual technical skills; it must be embedded in a broader pedagogical culture in which teachers, students, and leaders collaborate to develop meaningful, ethical, and sustainable applications of AI in education.

8. Conclusion and future developments

Within the context of the Erasmus AI4T project, the study aimed to explore tools and constructs that enhance educators' abilities to self-assess their proficiency in using digital tools and artificial intelligence (AI) in the classroom. The study emphasised the increasing importance of digital and AI literacy in education, considering the dual aspects of empowering educators and promoting the ethical use of AI. It was confirmed that improving the use of AI in the classroom requires an understanding of its pedagogical contributions (Celik, 2023). The research findings revealed a positive attitude and willingness among teachers to embrace AI in education. The study emphasised the importance of integrating technical, pedagogical and subject-specific knowledge to encourage the use of AI (Koehler *et al.*, 2014). Creating a community of practice through training sessions enabled teachers to identify good practices for using AI in the classroom (Ligorio, 2022). This approach addresses the need to bridge the gap between current educational outcomes and the evolving demands of the global labour market by promoting 21st-century skills among students and fostering greater alignment between their existing use of AI in formal, informal, and non-formal contexts (Carvalho *et al.*, 2022).

Looking ahead, the study proposes several avenues for further research and development. Building on the demand for practical, hands-on examples of AI in teaching, future initiatives could involve developing comprehensive training modules incorporating case studies, simulation exercises, and live projects.

These modules would help educators understand and apply the theoretical aspects of AI in real-world classroom settings. However, as AI becomes more integrated into education, ethical considerations must be at the forefront. Future research should therefore explore frameworks and guidelines for the ethical use of AI in education, with a focus on issues such as privacy, bias and equity. This includes developing inclusive and accessible AI tools for all students to ensure fairness in AI-assisted educational outcomes. About self-assessment of professional competence, the development of AI-based assessment and evaluation tools could provide educators with insights into

their own performance and their students' learning progress. Future developments should focus on creating reliable and valid AI tools that can assess a wide range of skills and competencies, including academic knowledge and soft skills.

References

Asbell-Clarke J. (2020), "The Importance of Teacher Bridging in Game-Based Learning Classrooms", *Global Perspectives on Gameful and Playful Teaching and Learning*, IGI Global Scientific Publishing, Palmdale.

Bellas F., Guerreiro-Santalla S., Naya M. (2022), "AI Curriculum for European High Schools: An Embedded Intelligence Approach", *Int. J. Artif. Intell.*, 33, pp. 399-426.

Bryman A. (2016), *Social research methods*, University Press, Oxford.

Lindgren B.M., Lundman B., Graneheim U. H. (2020), "Abstraction and interpretation during the qualitative content analysis process", *International Journal of Nursing Studies*, 108, retrieved on December 5, 2024, from https://www.sciencedirect.com/science/article/pii/S0020748920301164.

Calvani A. (2011), Principi dell'istruzione e strategie per insegnare, Carocci, Roma.

Carvalho L., Martinez-Maldonado R., Tsai Y.S, Markauskaite L., De Laat M. (2022), "How can we design for learning in an AI world?", *Computers and Education: Artificial Intelligence*, 3, retrieved on December 5, 2024, from https://www.sciencedirect.com/science/article/pii/S2666920X2200008X.

Calvani C. (2020), Tecnologie per l'inclusione, Carocci, Roma.

Chaipidech P., Srisawasdi N., Kajornmanee T., Chapah K. (2022), "A personalized learning system-supported professional training model for teachers' TPACK development", *Computers and Education: Artificial Intelligence*, 3, retrieved on December 5, 2024, from https://www.sciencedirect.com/science/article/pii/S2666920X22000194.

Celik I., Dindar M., Muukkonen H., Järvelä S. (2022), "The promises and challenges of artificial intelligence for teachers: A systematic review of research", *TechTrends*, 66 (4), pp. 616-630.

Celik I. (2023), "Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education", *Computers in Human Behavior*, 138, retrieved on December 5, 2024, from https://www.sciencedirect.com/science/article/pii/S0747563222002886.

Ciucci F. (2012), *L'intervista nella valutazione e nella ricerca sociale. Parole di chi non ha voce. Parole di chi non ha voce*, FrancoAngeli, Milano.

Collie R.J., Martin A. J., Gašević D. (2024), "Teachers' generative AI self-efficacy, valuing, and integration at work: Examining job resources and demands", *Computers and Education: Artificial Intelligence*, 7 (3), 100333.

Corrao S. (2005), "L'intervista nella ricerca sociale", *Quaderni di sociologia*, 38, pp. 147-171.
Council of Europe (2022), *Artificial Intelligence and Education – A Critical View Through the Lens of Human Rights, Democracy, and the Rule of Law*, retrieved on December 5, 2024, from https://book.coe.int/en/education-policy/11334-pdf-artificial-intelligence-and-education-a-critical-view-through-the-lens-of-human-rights-democracy-and-the-rule-of-law.html.
Crompton H., Burke D., Nickel C., Chigona A. (2024), "The SETI Framework and Technology Integration in the Digital Age", *Asian Journal of Distance Education*, 19 (1), retrieved on December 5, 2024, from https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/771.
Damiano E. (2013), *La mediazione didattica. Per una teoria dell'insegnamento*, FrancoAngeli, Milano.
Davis F.D. (1989), "Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology", *MIS Quarterly*, 13 (3), pp. 319-339.
Davis F.D., Bagozzi R.P., Warshaw P.R. (1989), "User acceptance of computer technology: a comparison of two theoretical models", *Management Science*, 35, pp. 982-1003.
Darling-Hammond L., Hyler M.E., Gardner M. (2017), *Effective Teacher Professional Development (Research Brief)*, Learning Policy Institute, Palo Alto.
Della Porta D. (2014), *L'intervista qualitativa*, Laterza, Roma-Bari.
Domenici G. (2015), *Manuale della valutazione scolastica*, Armando, Roma.
Downe-Wamboldt B. (1992), "Content analysis: Method, applications, and issues", *Health Care for Women International*, 13 (3), pp. 313-321.
European Commission, Directorate-General for Education, Youth, Sport and Culture (2022), *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*, Publications Office of the European Union, https://data.europa.eu/doi/10.2766/153756.
European Commission, European Education and Culture Executive Agency (2023), *AI report: by the European Digital Education Hub's Squad on artificial intelligence in education*, Publications Office of the European Union, retrieved on December 5, 2024, from https://data.europa.eu/doi/10.2797/828281.
Fitria T.N. (2021), "Artificial intelligence (AI) in education: Using AI tools for teaching and learning process", *Prosiding Seminar Nasional & Call for Paper STIE AAS*, 4, 1, pp. 134-147.
Fjelland R. (2020), "Why general artificial intelligence will not be realized", *Humanit. Soc. Sci. Commun.*, 7 (10).
Gocen A., Aydemir F. (2020), "Artificial intelligence in education and schools", *Research on Education and Media*, 12 (1), pp. 13-21.
Gomez F.C., Jr Trespalacios J., Hsu Y.C., Yang D. (2022), "Exploring Teachers' Technology Integration Self-Efficacy through the 2017 ISTE Standards", *TechTrends: for leaders in education & training*, 66 (2), pp. 159-171.
Guskey T.R. (2000), *Evaluating Professional Development*, Corwin Press, Thousand Oaks.

Harvey E., Koenecke A., Kizilcec R.F. (2025), "'Don't forget the teachers': Towards an educator-centered understanding of harms from large language models in education", *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*, Association for Computing Machinery, 1064, pp. 1-19.

Hattie J. (2023), *Visible Learning: The Sequel: A Synthesis of Over 2,100 Meta-Analyses Relating to Achievement*, Taylor & Francis, New York.

Holmes W., Bialik M., Fadel C. (2019), *Artificial intelligence in education: Promises and implications for teaching and learning*, Center for Curriculum Redesign, retrieved on December 5, 2024, from https://curriculumredesign.org/wp-content/uploads/AIED-Book-Excerpt-CCR.pdf.

Holmes W., Tuomi I. (2022), "State of the art and practice in AI in education", *European Journal of Education*, 57, retrieved on December 5, 2024, from https://onlinelibrary.wiley.com/doi/10.1111/ejed.12533.

Hutchins N.M., Zhang S., Barrett J. R., Israel M. (2025), "Empowering Educators in AI: Insights from Co-Designing an AI Microcredential with and for K-12 Educators", *Proceedings of the AAAI Conference on Artificial Intelligence*, 39 (28), pp. 29137-29144.

Kemp A., Palmer E., Strelan P. (2019), "A taxonomy of factors affecting attitudes towards educational technologies for use with technology acceptance models", *British Journal of Educational Technology*, 50 (5), pp. 2394-2413.

Kim J. (2025), "Perceptions and preparedness of K-12 educators in adopting generative AI", *Research in Learning Technology*, 33, retrieved on December 5, 2024, from https://journal.alt.ac.uk/index.php/rlt/article/view/3448.

Kong S. C., Yang Y., Hou C. (2024), "Examining teachers' behavioural intention of using generative artificial intelligence tools for teaching and learning based on the extended technology acceptance model", *Computers and Education: Artificial Intelligence, Advance online publication*, 7, retrieved on December 5, 2024, from https://www.sciencedirect.com/science/article/pii/S2666920X24001310.

Kvale S. (1994), *Interviews: an introduction to qualitive research interviewing*, Sage, Thousand Oaks.

Koehler M.J., Mishra P., Kereluik K., Shin T.S., Graham C.R. (2014), "The Technological Pedagogical Content Knowledge Framework", in J. Spector, M. Merrill, J. Elen, M. Bishop (eds.), *Handbook of Research on Educational Communications and Technology*, Springer, New York.

Korinek A., Schindler M., Stiglitz J. (2021), "Technological Progress, Artificial Intelligence, and Inclusive Growth", *IMF Working Papers*, 166.

Ligorio M.B. (2022), "Intelligenza artificiale e apprendimento", *Sistemi intelligenti*, 34 (1), pp. 21-26.

Liua Y., Salehb S., Huang J. (2021), "Artificial intelligence in promoting teaching and learning transformation in schools", *Artificial Intelligence*, 15 (3), pp. 892-902.

Lucisano P., Salerni A. (2002), *Metodologia della ricerca in educazione e formazione*, Carocci, Roma.

Markauskaite L., Marrone R., Poquet, O., Knight S., Martinez-Maldonado R., Howard S., Tondeur J., De Laat M., Buckingham Shum S., Gašević D., Siemens G.

(2022), "Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? Computers and Education", *Artificial Intelligence*, 3, retrieved on December 5, 2024, from https://www.sciencedirect.com/science/article/pii/S2666920X2200011X.

Miao F., Holmes W., Huang R., Zhang H. (2021), *AI and education: Guidance for policy-makers*, UNESCO, Paris, retrieved on December 5, 2024, from https://unesdoc.unesco.org/ark:/48223/pf0000 376709.

Miao Y., Yao Y. (2021), *Professional Development of College Teachers in the Era of Artificial Intelligence: Role Rebuilding and Development Path*, Springer, Cham.

Miao F., Mutlu C. (2024), *AI competency framework for teacher*s, UNESCO, Paris.

Miao F., Holmes W. (2023), *Guidance for generative AI in education and research*, UNESCO, Paris, retrieved on December 5, 2024, from https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research.

Molenaar I. (2021), "Personalisation of learning: Towards hybrid human-AI learning technologies", *OECD Digital Education Outlook 2021: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots*, OECD Publishing, Paris.

Mori S., Rosa A., Niewint J. (2022), "Personalisation of teaching and use of digital technologies between face-to-face and distance education", *Proceedings Apprendere con le tecnologie tra presenza e distanza*, Convegno SIREM 2022.

Niewint-Gori J. (2023), *A snapshot of the evolving landscape of artificial intelligence in education*, retrieved on December 5, 2024, from https://www.ital-ia2023.it/submission/92/paper.

Ning Y., Zhang C., Xu B., Zhou Y., Wijaya T.T. (2024), "Teachers' AI-TPACK: Exploring the Relationship between Knowledge Elements", *Sustainability*, 16 (3), 978.

Owston R. (2007), "Contextual factors that sustain innovative pedagogical practice using technology: an international study", *J. Educ Change*, 8, pp. 61-77, retrieved on December 5, 2024, from https://doi.org/10.1007/s10833-006-9006-6.

OECD (2023), *OECD Digital Education Outlook 2023: Towards an Effective Digital Education Ecosystem*, OECD Publishing, Paris.

OECD (2019), *How teachers update their teaching skills to cope with the rapidly changing world (Teaching in Focus No. 28)*, OECD Publishing, Paris, retrieved on December 5, 2024, from https://www.oecd.org/en/publications/how-teachers-update-their-teaching-skills-to-cope-with-the-rapidly-changing-world_282b9edd-en.html.

Pedró F., Subosa M., Rivas A., Valverde P. (2019), *Artificial Intelligence in Education: Challenges and opportunities for sustainable development*, UNESCO, Paris, retrieved on December 5, 2024, from https://unesdoc.unesco.org/ark:/48223/pf0000366994.

Petrucco C., Favino F., Conte A. (2025), "Teachers' perceptions on the introduction of Generative AI in schools: A mixed-method study on the opinions of 1,223 teachers in the Veneto Region, Italy", *Education Sciences & Society – Open Access*, 15 (2).

Shemshack A., Spector J.M. (2020), "A systematic literature review of personalized learning terms", *Smart Learn. Environ.*, 7, 33.

Selwyn N. (2022), "The future of AI and education: Some cautionary notes", *European Journal of Education*, 57, pp. 620-631, retrieved on December 5, 2024, from https://doi.org/10.1111/ejed.12532.
Shapiro J.P., Stefkovich J.A. (2016), *Ethical leadership and decision making in education: Applying theoretical perspectives to complex dilemmas*, Routledge, London.
Stahl B.C., Wright D. (2018), "Ethics and privacy in AI and Big Data: implementing responsible research and innovation", *IEEE Secur. Priv.*, 16 (3), pp. 26-33.
Tang K.S., Cooper G., Rappa N., Cooper M., Sims C., Nonis K. (2024), "A dialogic approach to transform teaching, learning & assessment with generative AI in secondary education: a proof of concept", *Pedagogies: An International Journal*, 19 (3), pp. 493-503.
Taopan L.L., Drajati N.A., Sumardi S. (2020), "TPACK Framework: Challenges and Opportunities in EFL classrooms", *Research and Innovation in Language Learning*, 3 (1), pp. 1-22.
Walkington C., Bernacki M. (2019), "Personalizing Algebra to Students' Individual Interests in an Intelligent Tutoring System: Moderators of Impact", *International Journal of Artificial Intelligence in Education*, 29 (1), pp. 58-88, retrieved on December 5, 2024, from https://doi.org/10.1007/s40593-018-0168-1.
Yang W. (2022), "Artificial Intelligence education for young children: Why, what, and how in curriculum design and implementation", *Computers and Education: Artificial Intelligence*, 3, 100061.
Zawacki-Richter O., Marín V.I., Bond M., Gouverneur F. (2019), "Systematic review of research on artificial intelligence applications in higher education – where are the educators?", *International Journal of Educational Technology in Higher Education*, 16, 39.

3. Exploring school personnel perspectives on a development program for school self-evaluation

by Giuseppe C. Pillera, Letizia Giampietro, Donatella Poliandri*

Italian schools have conducted a self-evaluation process since 2014, mandated by DPR 80/2013, which requires careful consideration of staff evaluative skills and training strategies. Key aspects for fostering an evaluation culture include supporting processes, networking with stakeholders, and promoting collaborative spaces. Valu.E for Schools (VfS) aimed to strengthen self-evaluation skills of school teachers (STs) and school principals (SPs) through training courses delivered between 2020 and 2021. Thirteen online focus groups were conducted, transcribed, and subjected to textual analysis to compare the results of SPs and STs. The study explores the alignment of training with schools' needs, collaboration activities, and dissemination of outcomes. Data from 37 SPs and 101 STs were analyzed, revealing differences in thematic articulation and level of detail. SPs presented more extensive self-evaluation work and were more aware of emerging training needs, while STs discussed training course details more extensively. This study highlights SPs' adherence to training objectives and their role in preparing evaluation activities for staff. Targeted courses can strengthen SPs' competences for school-wide self-evaluation and support STs' evaluation and data reading skills, enhancing evaluation and governance processes within schools.

Le scuole italiane conducono un processo di autovalutazione dal 2014, come previsto dal DPR 80/2013, che richiede un'attenta considerazione delle competenze valutative dello staff e delle strategie formative. Gli aspetti chiave per favorire una cultura della valutazione includono il supporto ai processi, il networking con gli stakeholder e la promozione di spazi col-

* The chapter is the result of joint work. Specifically, Giuseppe C. Pillera is responsible for par. 4 and 5, Letizia Giampietro is responsible for par. 1, 2 and 3, Donatella Poliandri is responsible for par. 6 and 7.

laborativi. La ricerca intervento Valu.E for Schools (VfS) mirava a potenziare le competenze di autovalutazione dei docenti e dei dirigenti scolastici attraverso corsi di formazione svolti tra il 2020 e il 2021. Sono stati condotti tredici focus group online, trascritti e sottoposti ad analisi testuale per confrontare i risultati di docenti e dirigenti scolastici. Questo studio esplora l'allineamento della formazione con le esigenze delle scuole, le attività di collaborazione e la diffusione dei risultati. I dati relativi a 37 dirigenti scolastici e 101 docenti sono stati analizzati, rivelando differenze nell'articolazione tematica e nel livello di dettaglio. I dirigenti sono propensi a fare un lavoro di autovalutazione più esteso e mostrano una maggiore consapevolezza delle nuove esigenze formative, mentre i docenti hanno discusso più approfonditamente i dettagli dei corsi di formazione. Questo studio evidenzia l'aderenza dei dirigenti scolastici agli obiettivi formativi e il loro ruolo nella preparazione delle attività di valutazione per il personale. Corsi mirati possono rafforzare le competenze dei dirigenti scolastici per l'autovalutazione d'istituto e sostenere le competenze di valutazione e di lettura dei dati dei docenti, migliorando i processi di valutazione e di governance all'interno delle scuole.

1. Introduction

Since 2014, Italian schools have conducted a process of self-evaluation, as required by DPR 180/2013. The introduction of a semi-structured self-evaluation report required careful consideration of data teams' evaluative skills and the training strategies to support evaluation capacity building (Poliandri, Freddano and Molinari, 2019).

According to the OECD (2013), key aspects for fostering a culture of evaluation in schools include supporting evaluative processes, networking with stakeholders in the local community and promoting spaces – including online ones – for discussion, exploration, and collaboration within and among schools.

In this scenario, the *Valu.E for Schools* (VfS) project[1] aimed to promote the capacity building of schools and to strengthen school teachers' (STs) and school principals' (SPs) professional skills in self-evaluation, through three professional development paths on evaluation involving 42 schools and 440 participants, including principals and members of data teams.

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

In this contribution, we examine the perspectives of both school principals and teachers who have participated in the training course, focusing on the medium-term progress of the project. This exploration encompasses several key topics, including the alignment of the provided training with schools' expectations and needs, the promotion of collaboration and peer learning activities, and the dissemination of learning outcomes within the school community. This comparative analysis has the potential to significantly enhance the development of tailored professional development pathways that effectively address the diverse needs and roles within the educational context.

2. Sustaining evaluation capacity building in schools: some notes

The Presidential Decree 80/2013, the Regulation on the National Evaluation System (NES), implements European directives sent to the Italian government in November 2011 for the construction of a National School Evaluation System. The regulation articulates a recursive process of school evaluation that starts with self-evaluation and proceeds with an improvement phase, providing for an external evaluation phase for a given sample of schools.

The Presidential Decree 80/2013 views the self-evaluation phase as the starting point for continuous improvement action within an essentially formative development approach. Some studies report that school improvement, as a result of evaluation, seems to occur through indirect development processes, where important elements include stakeholder involvement and teacher collaboration (Gustafsson *et al.*, 2015); a decentralized perspective in supporting school self-evaluation (Ehren *et al.*, 2017); the promotion of interventions to strengthen the evaluative culture in schools (Schildkamp *et al.*, 2016); the building of evaluative skills of school data teams (Hubers and Poortman, 2018); the creation of spaces for in-depth study, including online, in a horizontal mode among peer schools and the allocation of more human and financial resources (OECD, 2013); the role of school networks in supporting self-evaluation processes as a strategic opportunity for mediation between the application of standards and the needs and practices of schools (Diaz *et al.*, 2014; Azorin and Muijs, 2017; Fiore and Torelli, 2019).

In NES, the aims are both improvement and accountability, in a horizontal perspective towards the local context and families and vertically towards the central administration. Therefore, since the administration can determine

intervention measures based on the results of the self-evaluation process, the quality of the self-evaluation processes conducted by schools becomes crucial (Giampietro *et al.*, 2016). The implementation of NES therefore led to greater attention to data literacy of school personnel, to the ability of school data teams to master the different phases of self-evaluation, and to assume an evaluative perspective of school processes aimed at improvement (Muzzioli *et al.*, 2016; Robasto, 2016), as also reported by other international research about the capacity of data teams to carry out effective and valid self-evaluation (Blok *et al.*, 2018).

The literature also points out that the development of evaluative capacity can be fostered by adopting a perspective that is sensitive to situated contexts, considering certain elements such as school culture, teacher collaboration, leadership, communities of practice for sharing and building knowledge, and attitudes toward evaluation (Schildkamp, 2019). As a result of the *VALES*[2] and *Evaluation and Improvement*[3] experimental projects, the need has emerged for professional development paths that are targeted and as close as possible to the specific needs of similar schools (Fortini *et al.*, 2016) and that involve the participation of all components of the school, both principals and teaching and administrative staff (Poliandri *et al.*, 2019). In fact, as Tagle and Casavola (2004) argue, the construction of evaluative capacity building requires broader reflection on the role and the different real capabilities of the administration as a whole: there is no single way of building capabilities; rather, there are different ways of combining common elements depending on the opportunities that each system and/or experience faces.

3. Sustaining evaluation capacity building in schools, a proposal: value for schools projects

The overall purpose of VfS is to test and evaluate the effectiveness of different training models in supporting schools' evaluation processes. The training and support actions are aimed at improving the capacity of school personnel in self-evaluation (Giampietro and Romiti, 2019).

[2] *VALES – Valutazione e Sviluppo della Scuola*: https://www.invalsi.it/invalsi/ri/vales/.
[3] *Valutazione e Miglioramento*: https://www.invalsi.it/invalsi/ri/audit/.

The objectives are multiple:

– support the skills of school leaders so that they can provide effective feedback, assist, advise, and evaluate staff activities, and lead the whole-school evaluation process;
– strengthen the evaluation and planning skills of teachers to improve their ability to read the data offered by the system with possible improved effects on the design and monitoring of improvement plans;
– give insights into the training and support models for schools that can potentially be implemented in the future in a broader scope;
– offer information on the provision of informal and formal networks to support self-evaluation and design for improvement.

VfS included a public call to select three entities (universities, associations, public and private training institutions, etc.) to be entrusted with the design and implementation of training and support activities for schools to improve the staff's ability to intervene in self-evaluation. The selected subjects – one for each macro-region of the country (North, Central, and South of Italy) – carried out specialized training and support activities, through the provision of advanced technological solutions, at a sample of schools identified by INVALSI[4], during the emergency period determined by the Covid-19 pandemic, from spring 2020 to December 2021.

In designing the support action, efforts were made to adapt the interventions to the sampled schools by providing activities that were accessible online, spatially, and contextually proximate, and tailored to the school's needs (Vanhoof, 2013).

Consistent with the literature on the professional development of school personnel, training models that involve sharing and building knowledge in accordance with the models of professional communities that learn collaboratively (Wenger, 1998) were selected; training methodologies had to support critical-reflective aspects (Schön, 2006; Mezirow, 1997) and contextualize evaluation processes within the school community (Brown and Poortman, 2018).

The methodologies for training activities, referred to as Lines of Action Training (LAF) (Gomez Paloma *et al.*, 2020), included peer learning between schools and situated training. The first line of training action, peer

[4] The sample of 45 schools (primary, lower-secondary, and comprehensive institutes, including all three types) – 15 for each macro-region (North, Central, South) – was selected through a two-stage stratified sampling: provinces were considered as the first stage units (a random sampling selected one province for every intervention region: Lombardia, Piemonte, Emilia-Romagna, Lazio, Toscana, Umbria, Puglia, Campania, Sardegna); schools were considered as the second stage units (selection of five schools in each province through random sampling).

learning between schools from a professional learning perspective (Topping, 2007), provided joint interventions for paired schools, enhancement of co-operative learning, and peer evaluation. The second line of action situated learning (Rossi, 2011; Rivoltella, 2013), involved enactive teaching interventions, focusing on the social and contextual conditions of knowledge and skill-building processes.

The training courses were designed using a mixed methodology, with in-person and distance (blended) modules, utilizing the MOODLE platform (Modular Object-Oriented Dynamic Learning Environment) for distance activities. The delivery of training activities on the platforms prepared by the partners and video conferencing tools allowed the project to be launched and subsequently developed even during the resurgence phase of the pandemic caused by the Covid-19 virus.

Schools participating in VfS were identified through a two-stage stratified sampling procedure; 15 schools were sampled in each macro area of reference of the call (North: Lombardia, Emilia Romagna, Piemonte; Central: Lazio, Toscana, Umbria; South: Puglia, Campania, Sardegna), for a total of 45. The total participants in the support and training activities were 42 school principals and about 400 teachers.

4. Method and data

As part of the project's evaluation design, 13 Online Focus Groups (OFG) were conducted (Poliandri *et al.*, 2023) with a convenience sample of project participants – school principals and teachers – to debate together the strengths and limitations of the training course and the medium-term progress of the project (Poliandri *et al.*, 2022). Four online focus groups were carried out with SPs: one for each macro-region plus an additional catch-up one. Nine focus groups were carried out with STs: one for each region involved. The overall sample of focus group participants consists of 37 SPs and 101 STs (on average, 9.3 and 11.2 participants per focus group, respectively). The spatial-temporal context of OFGs is the pandemic one, during spring 2021.

In this contribution, we compare the main results of the computer-aided textual analysis conducted on the focus groups with SPs and with STs, exploring redundancies and peculiarities of viewpoints emerging from similarities and differences in the articulation of the discussion topics:

– alignment of the provided training with schools' expectations and needs;
– promoted collaboration and peer learning activities;
– dissemination of the learning outcomes within school.

The recordings of the 13 OFGs were automatically transcribed by the online software Cabolo[5], then the transcriptions were manually reviewed by the VfS project research team (Poliandri *et al.*, 2023). The corpus contains 979 interventions (excluding those relating to the moderator and those in the final farewell phase of each OFG): 364 by SPs (37.2%) and 615 by STs (62.8%).

A computer-aided textual analysis was collaboratively conducted by four researchers, using a coding methodology (Adu, 2019; Pagani, 2020) inspired by an abductive approach (Fereday and Muir-Cochrane, 2006). This means that we combined the scheme deriving from the focus groups' stimulus questions with the evidence emerging from the direct exploration of the empirical base. In other words, we used:

– a top-down approach, to define general categories as top-level topics, deriving them from the four focus group stimuli, which results in four categories (Needs and expectations, Training course, Collaboration and discussion, Outcomes and impacts);
– a bottom-up approach, to specify several codes for each category, as topics at a deeper level of detail, basing them on the evidence emerging from direct exploration of the empirical base.

We have coded each intervention using one or more codes, thus obtaining 1,259 encoded fragments[6]. The main results of the coding process, including synthetic definitions of the codes, are reported in Table 1. Then, we examined the corpus by means of text data mining techniques[7].

[5] Cabolo: https://www.cabolo.com.

[6] We frequently utilized different codes within a single speaker's intervention because it often encompassed various specific topics. While the topics were often distinct, allowing for sequential coding, there were instances where it was necessary to overlap two codes partially or fully. This was done to accommodate the complexity, different nuances, and potential interpretations of the same intervention fragment.

[7] The coding and data analysis processes were carried out by VfS research team using the text analysis software QDA-Miner: https://provalisresearch.com/products/qualitative-data-analysis-software.

Tab. 1 – Categories and codes: synthetic definitions, code frequency (number of fragments encoded with a given code) and coded cases (number of cases containing a given code)

Category	*Code*	*Definition*	*Code frequency*		*Coded cases*	
			SPs	*STs*	*SPs*	*STs*
Needs and Expectations (N&E)	Self-presentation	Speaker's self-presentation, information about participating group and descriptions of the school context.	66	104	57	90
	Expectations	Explicit expectations about various aspects of training, including statements as to whether these are met	54	73	47	57
	Dynamics and quality of participation	Elements that qualify attendance (way of perceiving oneself in the training, interest, motivation, engagement, constancy, etc.), as well as information on changes in the group during the training (turnover, contractions/expansions, change of SP, etc.)	29	88	27	73
Training course (TC)	Aims	Comments on communication/explanation, clarity/clarification, understanding and sharing or not of the aims and objectives of the training	15	28	14	25
	Contents and materials	Descriptions, comments, and judgements relating to the training topics, learning materials and teaching aids	25	88	23	80
	Teaching methods	Reflections and perceptions with respect to the methodological-didactical choices and their implementation, including reflections on the "distance" modes	9	88	8	80
	Timing and calendar	Descriptions, comments, and judgements on the scheduling of meetings and other planned activities, as well as on the duration of the webinars	5	30	5	28
	Judgements about the course	General and generic evaluations of the training course (used when there are no further elements to support the judgement made and when the elements are various and accumulated in the same unit of analysis)	17	28	16	26

Tab. 1 – Categories and codes: synthetic definitions, code frequency (number of fragments encoded with a given code) and coded cases (number of cases containing a given code) (to be continued)

Category	*Code*	*Definition*	*Code frequency*		*Coded cases*	
			SPs	*STs*	*SPs*	*STs*
Collaboration and Discussion (C&D)	Intra-group C&D	Considerations regarding moments, opportunities and ways of exchange, dialogue, debate, and shared reflection between participants from the same school and between participants and their own DS, including limitations, critical issues, problems and conflicts	19	67	19	61
	C&D among schools	Considerations regarding moments, opportunities and ways of exchange, dialogue, debate and shared reflection between participants from different schools and between participants from one school and DSs from another, including limitations, critical issues, problems and conflicts	27	60	24	54
	C&D with course teachers and tutors	Considerations regarding moments, opportunities and ways of exchange, dialogue, debate and shared reflection between participants and trainers or tutors, including limitations, critical issues, problems, and conflicts, as well as comments on activities and interventions of guidance, mediation, facilitation, clarification and provision of feedback by trainers or tutors	5	30	5	29
	Discussion during INVALSI OFG	Considerations regarding the ongoing discussion within the INVALSI OFG, both with the conductor and among the participants	3	7	3	6

Tab. 1 – Categories and codes: synthetic definitions, code frequency (number of fragments encoded with a given code) and coded cases (number of cases containing a given code) (to be continued)

Category	*Code*	*Definition*	*Code frequency*		*Coded cases*	
			SPs	*STs*	*SPs*	*STs*
Outcomes and impacts (O&I)	First general outcomes	General and generic considerations concerning the first results, effects, or individual/group impacts (used also when the elements are various and accumulated in the same unit of analysis)	12	14	12	12
	Awareness of new learning needs	Information and considerations regarding the emergence of new training or research needs (at individual, group or institute level) relating to the topics covered by the training course	24	9	20	9
	Implications for data team's work and evaluation skills	Information and considerations relating to the first individual or group results related to the training course and specifically inherent to the work of the data team and to the self-assessment and improvement competences	32	62	31	55
	Impacts on classroom work	Information and considerations related to awareness/awareness, results, impacts, developments (missing, desired, planned or achieved) related to the training course and concerning daily educational-didactic work (planning, collaboration, classroom management, teaching methods, students' assessment, etc.)	0	10		10
	Shared outcomes and impacts on the school	Information and considerations concerning missing, desired, planned or achieved impacts: sharing, at school level, of information, materials, tools, results, reflections and future perspectives related to the training course	40	91	36	88
		Tot.	382	877	347	783

In our analysis, each of the 979 interventions represents a case, to which several background variables – descriptive of the sample – are associated, first and foremost, the professional role. The encoded fragments, both in terms of general categories and specific codes, have been counted for each of the two professional categories and analysed through column percentages, to directly compare the relative weight of the various topics (codes) into the SPs' and STs' focus groups[8]. Moreover, the coded cases have been compared between the two professional categories row by row, enabling us to conduct a chi-square test for each code, to evaluate the association between the role and the cases containing that code[9].

The statistical procedures for the chi-square test were carried out by verifying the basic assumptions (independence of observations and adequacy of expected cell counts) and should therefore be considered valid. Nevertheless, they are interpreted as exploratory, mainly aimed at guiding the targeted selection and re-reading of textual material, fostering critical reflection, and supporting the choice of illustrative excerpts[10].

Statistical significance was set at $p < 0.05$. Results with p-values between 0.05 and 0.1 were not considered significant and are reported only as non-conclusive trends.

5. Results

The comparison of SPs and STs sub-corpora at a general level (i.e., the top-level categories) through the column percentage analysis (Fig. 1) reveals the first main differences regarding the thematic articulation of the discussion, especially concerning the categories "Needs and expectations" and "Training course", to which SPs respectively allocate more and less attention (+8.8% and -11.3% of encoded fragments). There are also notable differenc-

[8] In the analysis of the column percentages, the codes for each group are scaled to 100, allowing for immediate comparison of their distributions.

[9] The row-by-row analysis (applied to top-level categories and then to each code, also in column analysis) corresponds to a 2x2 contingency table, with professional roles on one side and the presence or absence of a code on the other. This design satisfies the independence requirement of the chi-square test. In the tables of paragraph 5 we report row percentages only, since the difference between expected and observed values is already captured by the chi-square statistic.

[10] Although several code-level differences between the two professional roles reached statistical significance in the chi-square test, the strength of the associations, measured by Cramér's V, was generally small, indicating effects that are statistically reliable but limited in magnitude (at most 0.1827).

es observed in “Outcomes and impacts” (+7.1%) and “Collaboration and discussion” (-4.6%).

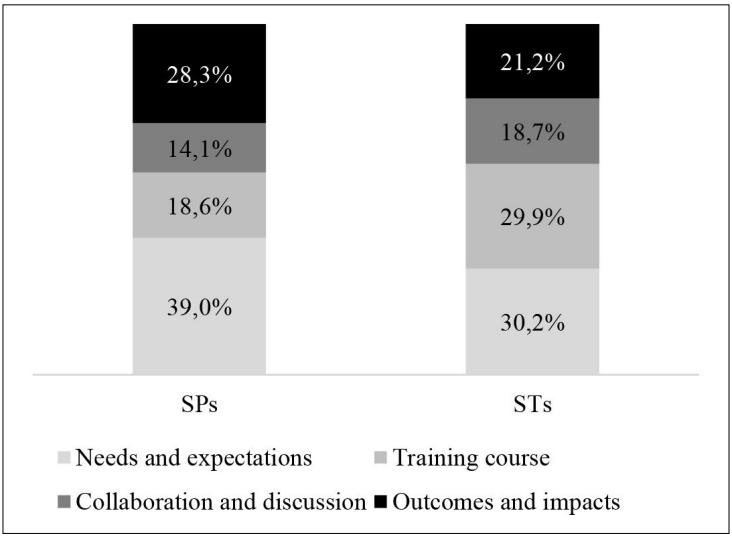


Fig. 1 – Top-level categories: frequency column percentages comparative analysis

However, the chi-square test reveals significant differences between SPs and STs only for the categories “Training course” and “Collaboration and discussion” (Tab. 2), both of which, as we have seen, are treated to a greater extent by STs.

Tab. 2 – Top-level categories: cases row percentages comparative analysis

Category	*SPs*	*STs*	*Chi-square*	*p value*
Needs and expectations	36.5%	63.5%	0.066	0.798
Training course	23.3%	76.7%	21.949	0.000
Collaboration and discussion	25.2%	74.8%	12.112	0.001
Outcomes and impacts	33.9%	66.1%	1.315	0.251

By delving deeper into the analysis of the column percentages (Fig. 2) and row percentages (in the tables that follow) at the more granular level of the codes, it is possible, respectively, to specify the differences observed above (at the category level) in terms of articulation of discussion (counting codes frequencies), and to identify additional differences in terms of association between professional role and a given code (counting coded cases), many of which are statistically significant.

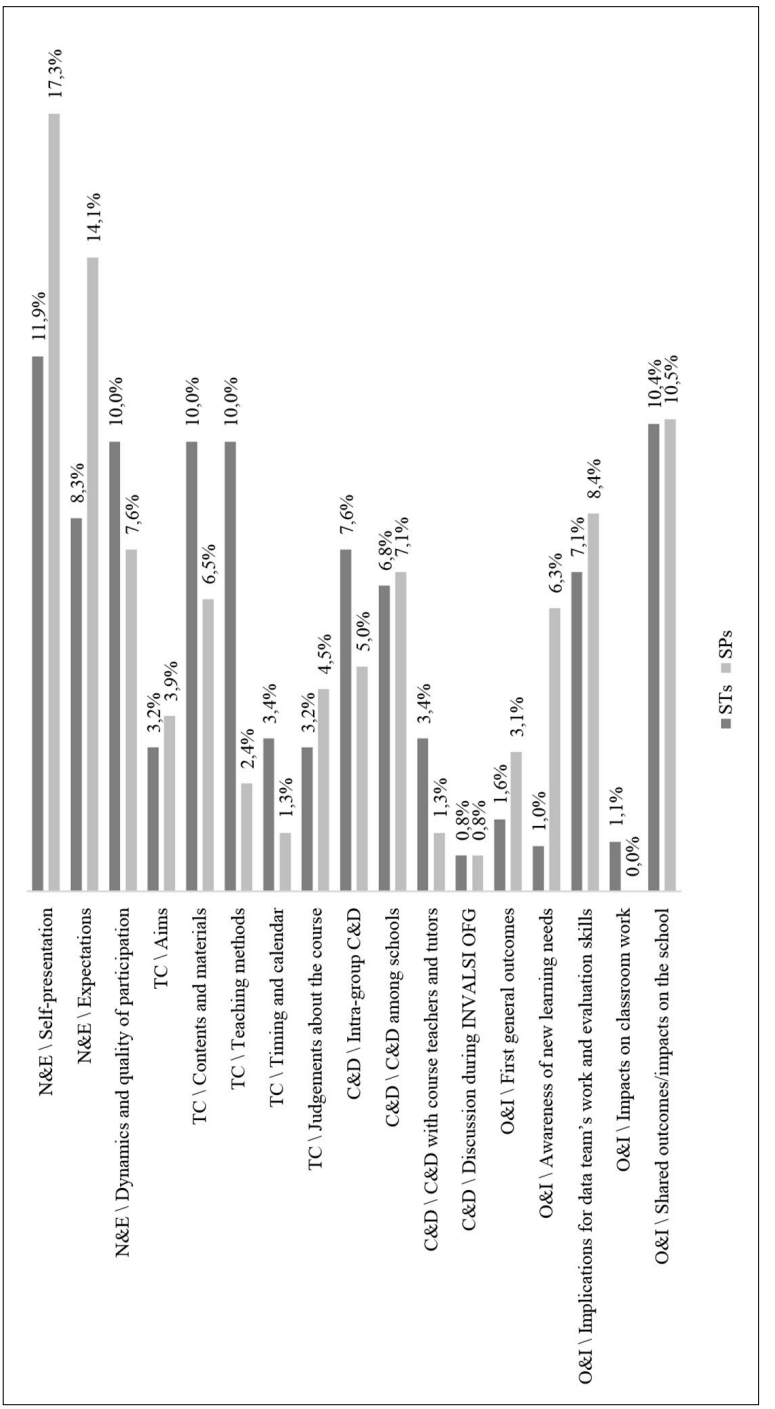


Fig. 2 – Code-level frequency column percentage comparative analysis

5.1. Expectations and needs

Observing the codes in the category "Needs and expectations" (Tab. 3), in comparison to the STs, SPs offer more extensive presentations of their own school context and self-evaluation work being carried out[11], put in relation to the more in-depth examination of expectations on the training course.

Tab. 3 – Codes in the category "Needs and expectations": row percentages comparison and chi-square test results

Code	*SPs*	*STs*	*Chi-square*	*p value*
Self-presentation	38.8%	61.2%	0.188	0.664
Expectations	45.2%	54.8%	3.198	0.074
Dynamics and quality of participation	27.0%	73.0%	4.942	0.026

«Bearing in mind that this training clearly comes […] in different contexts in terms of history, experiences already acquired and practised» (#625[12]), SPs' expectations are linked to different contextual needs, as exemplified by the following selected fragments: sharing and consolidating self-evaluation reflections and practices or creating and/or strengthening *young* data-team.

I must say, we were deeply invested in this project because it was crucial for us to refine the multitude of reflections and meta-reflections that had been initiated within the school [SP #337].

Our aim is to establish a group that not only fosters but also embodies the culture of self-assessment. Unfortunately, in my school, I've observed that it often remains merely a formal obligation [SP #66].

Moreover, the needs to develop skills aimed at identifying priorities and designing improvement («a strategic planning of the school, which is fundamental and often instead becomes a formal act or not well understood by the teaching board»: #23) are highlighted in several cases.

[11] We observed that this phenomenon is significantly more pronounced among SPs with less seniority, as noted across three experience bands: 2-3, 5-10, and 14-25 years (Spearman's rho = -0.119, p-value = 0.043).

[12] All quotations from OFG in this paper refer to the SPs or to STs. The hash character precedes the sequential number assigned to the intervention from which the quotation is taken, identifying the case within our text database. All the long quotations are numbered in order of appearance in the text and reported in the appendix in the original language.

Maybe we could have benefited from such a crucial and detailed intervention as far back as 2013-2014, when we first initiated the RAV, the Triennial Educational Offer Plan, and the improvement plan. This area remains the most formidable to tackle. Despite our painstaking efforts to identify priorities to address, we find it extremely challenging to enact actions geared towards improvement [SP #29].

In this sense, a SP (#176) makes it clear precisely how his expectation «to understand how self-evaluation should be a process that fits into a much broader discourse within the school […]. Self-evaluation as a moment, as a building block of this […] cyclicality of strategic planning».

STs' expectations, differently, are often linked to more practical and technical evaluation needs, as the following fragments testify.

The theme of self-assessment is both complex and technical. While it's true that there are ministry guidelines, tackling a RAV – what a RAV is, constructing it, or being able to apply it – is quite complex. As a member of the data team, I hoped that the course would not only be an opportunity for reflection but also for learning the technical aspect [ST #2065].

Perhaps from this project, we would have expected something more technical, meaning that they would have given us maybe […] some sort of format, something to reason, work, and think on. Instead, it seems like we're being asked to slowly build it ourselves [ST #1493].

However, both STs and SPs stress the productive disorientation due to disorientation from expectations of a more technically oriented course, as evidenced by the following fragments.

For the sake of paradox, I'd say the course didn't meet my expectations. As a newly appointed manager, I was hoping for assistance in understanding how to carry out the RAV. Thankfully, that wasn't the case; it turned out to be invaluable instead [SP #172].

At first, I genuinely believed that the course would be filled with grids, tables, all neatly structured. Indeed, I anticipated something rigid. However, to my surprise, despite all the valid criticisms my colleagues expressed – because deep down, I too acknowledged the confusion it caused – I enjoyed the course from the beginning. It met my expectations because it was something unstructured [ST #1670].

5.2. Training course

SPs comment less on the issues related to the category "Training course" (Tab. 4), especially on learning content and materials, teaching methods, and timing, while the STs have widely debated these topics, even with critical positions about questions such as timing and calendar (Poliandri *et al.*, 2022).

Tab. 4 – Codes in the category "Training course": row percentages comparison and chi-square test results

Code	*SPs*	*STs*	*Chi-square*	*p value*
Aims	35.9%	64.1%	0.029	0.866
Contents and materials	22.3%	77.7%	10.869	0.001
Teaching methods	9.1%	90.9%	32.665	0.000
Timing and calendar	15.2%	84.8%	7.096	0.008
Judgements about the course	38.1%	61.9%	0.016	0.900

Teachers not only provide the most widespread but also the most precise and detailed opinions on the "Training course", particularly regarding the contents and materials, teaching methods, and the timing and calendar (the second, third, and fourth codes in Tab. 4), as emerges from the following highly illustrative and comprehensive fragment.

> The structure is also nice, in the sense that some delve into the lesson, even if clearly at a distance, but still they hear people and so a series of information falls into place in the mind, then it is elaborated within the group and then there is a synthesis that however is a synthesis made up of multiple voices coming from multiple schools and this is a bit of a novelty because the RAV has always been instead perhaps something internal to the school to be managed by the instrumental functions that deal with it. And I also liked the platform a lot, in the sense that there is very clear material to view, but there is also the way, the time, and the request to put in information that not only forces you to listen to the videos or read, but to contribute your own and to read what others write, and I find this interesting because over time one rereads, revisits, discovers, and so it is a journey that one also takes alone, but experiences at home [ST #1067].

Several STs' interventions focus not only on the completeness, richness, usefulness, and usability of the materials offered by trainers but also on the fact that "the circularity and sharing of materials that are already in use in the various institutes is a great asset" [ST #1428].

Moreover, several STs – in line with what is reported in par. 4.1 regarding a pleasant failure to meet expectations – underline the dialogic and reflective style on which the proposed teaching was based.

So, I believe that the strength of this training lies in the mode of delivery, namely it's definitely a course that has focused a lot on self-reflection and becoming aware of self-assessment practices, putting us operators in a critical way, urging us to review our... and it has allowed us to review our practices [ST #1362].

Are the exercises helpful? Absolutely yes, they are, so to speak, the pedagogical artifact that allows you to self-analyse and understand within your institution, how things are going, where you are at. You stop and reflect [ST #1416].

However, this does not mean that the SPs did not pay attention to these issues, as emerges from the following fragments, where they appreciate some learning contents and materials of the course.

The other, let's say, formative polarity, which the project has certainly intercepted at the moment, but on which we still have to work a lot, is precisely this need to find effective ways to communicate self-assessment, but with a language that is accessible, obviously also to users, a sort of simple self-assessment, I don't know how to say, because in fact the complexity that characterizes even technical language and that is often difficult to make understandable even to teachers, translates into an equal complexity in communicating those that are the internal processes [SP #144].

And I also want to emphasize that what has been very important, at least for us as a group, is the significant work that has been done on the entire aspect of the work environment, the importance of relationships. In our group, there has been a lot of productive work on this [SP #315].

I believe that all the material represents a goldmine, obviously, so of excellent quality [SP #588].

Regarding the training methods, a large part of SPs' reflections as well as several of those of STs were devoted to the unexpected remodelling of the course in a completely distant mode. This situation was experienced with weariness and a sense of having missed out on something, especially in large and geographically articulated institutes, where the restrictions due to the pandemic further limited the already complicated opportunities to get together.

So, subjectively, the pandemic has significantly changed the atmosphere in schools. The project was initially set up and has been reconsidered, I would say, admirably, leveraging all possible tools, I believe even with immense organizational effort, but it has been rearranged in the best possible way. The issue is that the context in which it is implemented is different from the context in which it was conceived [SP #379].

However, some SPs, on the contrary, point out the positive aspects of the online mode in which the training took place:

> The opportunity to take this course remotely has enabled me to be more engaged and involved. I can't imagine what would have happened if we had to do everything in person; perhaps my participation wouldn't have been as meaningful [SP #66].

Finally, still in relation to the training methods, we want to highlight this extract from an intervention by a teacher, which anticipates the issues addressed in the following paragraph:

> However, the necessity of practically doing, experimenting in the field, seeking alternatives, and exploring strategies that extend beyond our familiar situations has been incredibly stimulating. It immerses you in a perspective of responsibility and different complexity. This aspect, I must say, is what I've enjoyed the most about this course [ST #1085].

5.3. Collaboration and discussion

In the category "Collaboration and discussion" (Tab. 5), some topics were significantly less addressed by the SPs, specifically collaboration and discussion within the participating school group and collaboration and discussion with course teachers and tutors (the first and third codes in Tab. 5).

Tab. 5 – Codes in the category "Collaboration and discussion": row percentages comparison and chi-square test results

Code	*SPs*	*STs*	*Chi-square*	*p value*
Intra-group collaboration and discussion	23.8%	76.3%	6.728	0.009
Collaboration and discussion among schools	30.8%	69.2%	1.492	0.222
Collaboration and discussion with course teachers and tutors	14.7%	85.3%	7.618	0.006
Discussion during the INVALSI OFG	33.3%	66.7%	0.058	0.810

This does not imply that the SPs did not, at least, follow and support the teamwork of the participating teachers from their own school[13], as the following fragments by SPs show.

So, what happened, especially in my school, is that there were moments that were not perceived as training because they were more about guiding and assisting groups, which on the other hand is part of those peer-to-peer techniques that we have all studied [SP #4].

As a group, therefore, we met, even more than once; there was a confrontation as we progressed along the path, which also allowed us to discuss, analyse, and even grasp those initial perplexities, understand why they were there, and then proceed and reflect, in short, on what the ongoing process was and enter, this allowed us to penetrate it well [SP #602].

The active participation of SPs was recognized as a key-point by several teachers, with functions of facilitating reflection among the group of participants and sharing it with larger groups, supporting communication among school participants, monitoring and interfacing with trainers, as exemplified as follows:

Our school principal, in addition to fairly regularly attending the meetings, also created a chat for all the participants of the course, where we can confront each other, and continuously monitors the situations that arise, those that might be perhaps points of reflection or even issues, challenges that we highlight, and she still has a preferential channel with the trainers, so if we have brought forward requests, she has still taken steps to ensure that these requests are conveyed to the trainers. So, there is continuous monitoring [ST #2152].

However, the STs provide us with some more precise accounts of the working dynamics within the school group, underlining the constant exchange of points of view. The complementarity of different previous experiences or the different levels of evaluation skills itself spontaneously enabled peer learning methods among the teachers involved, particularly well exemplified by the following fragment.

We are a large group, so we had the opportunity to have active exchanges during the meetings, to discuss, also... I mean, to have some doubts about the documents, because obviously those who are integral part of the data team (NIV) and have parti-

[13] We observed that SPs' interventions relating to the intra-group confrontation are somewhat directly proportional to their age, traced into three bands: 44-50, 51-56, 57-67 years (Spearman's rho = 0.117, p value = 0.045).

cipated in the drafting of various documents may have a different perspective, unlike other colleagues who have chosen, for example, to join the course to learn and have operational tools to then immerse themselves for the first time in this field [ST #2086].

The sharing, peer reviewing, and discussion of the group tasks assigned by the trainers also allowed:

A reflection on the different interpretations that each institution can give regarding the output to be produced [...], the working method has allowed within the data team to reflect on the different ways in which different institutions, different groups interpret and approach a request, precisely due to a matter of – let's say – willingness to delve into a certain topic; in the sense that the request was very clear, but for us it meant going in a certain direction because we wanted to use and implement something that serves us at this moment [ST #1410].

Both the STs and the SPs really appreciate the opportunity provided by different activities – depending on the specific macro-area training pathway (group works, visiting, comparison of the respective self-evaluation report, etc.) – to contact, discuss, collaborate with participants from other schools, as the following fragments testify.

I also really like the idea of checking, of seeing what other institutions have done to make comparisons and to understand effectively what our strengths are and what instead are the areas where we are a little more lacking [ST #1366].

On the other hand, I would say that the aspect of exchange with other school realities is also fundamental. We have participated as observers in a visiting. Soon, we will be the hosting school for a visiting. This is also an opportunity for the school to reflect on how it works, what objectives it is pursuing, on which underlying principles it operates, and therefore it is a sort of constant self-assessment. We have indeed framed our visiting precisely by considering the innovative experience, so to speak, of the kindergarten, precisely to enhance this participation because the biggest problem, previously mentioned as motivation, is to stimulate motivation on the part of teachers [SP #82].

In my opinion, one thing that [the course] has responded to very well in terms of our expectations, the project has been to allow us to move away from self-referentiality, so, we don't just read and sing the RAV among ourselves, but we also share it with other schools; so, it truly becomes a reason for exchange and comparison, to see what others do, how others do it, and so this is certainly the most positive aspect that we have highlighted of this project [SP #168].

Several STs also define the group work as the privileged aspect of the training course:

In group work, we begin with a shared content, but then various sensitivities from the experiences and characteristics of our schools, as well as our personal backgrounds, intersect, intertwine, and converge. Being part of a group of adults starting from a common ground entails... putting oneself at stake in a different manner. Thus, there's some struggle, some reorientation, some recalling of difficulties, and some searching for perhaps small pieces that may have been overlooked [ST #1121].

In this respect, the SPs specifically have exploited the few opportunities for specific exchange and discussion between the participating SPs (#363 asked a «dedicated communication channel for SPs within this project»), to take advantage of the experience of senior SPs or those working in different contexts and to plan future developments in terms of networks, as shown in the following fragments.

Certainly, it's an important moment of comparison, and also hearing from principals who may have many years of experience in school management and find themselves in very different contexts [SP #226].

I have already scheduled a time after this discussion with other principals to work specifically on implementing some of these learnings, these pathways, to our concrete situation, thus applying the model [...] to the particular situation [SP #80].

Finally, although it is a minority topic and mainly addressed by teachers, there was appreciation for the collaboration with the trainers' staff, also because they brought an external perspective to the school world. In particular, the effectiveness, consistency, and punctuality of the support provided by the tutors and coaches involved in the training courses were praised.

5.4. Outcomes and impacts

Apart from the result on the outcomes of the course related to the classroom work, obviously highlighted exclusively by the teachers, SPs and STs interventions appear to be in proportional balance for the most important codes of the category "Outcomes and impacts" (Tab. 6): implications for data team's work and evaluation skills; sharing of outcomes and impacts on the school.

Tab. 6 – Codes in the category "Outcomes and impacts": row percentages comparison and chi-square test results

Code	*SPs*	*STs*	*Chi-square*	*p value*
First general outcomes*	50.0%	50.0%	1.731	0.188
Awareness of new learning needs	69.0%	31.0%	12.927	0.000
Implications for data team's work and evaluation skills	36.0%	64.0%	0.052	0.820
Impacts on classroom work	0.0%	100.0%	5.98	0.014
Shared outcomes/impacts on the school	29.0%	71.0%	4.036	0.045

This code, even if 50% attributed to each of the two professional categories, therefore in a manner not proportional to the number of participants and interventions in the corpus, it is not significant in the chi-square test due to the relatively low occurrence of the code itself.

About the implications for data team's work and evaluation skills, it should certainly be emphasized that STs and especially SPs appear – at different levels, also depending on the starting points of each school – to be satisfied. The SPs, of course, reflect more at an organizational level, stating that the course is giving: «confirmation that some practices that were already being used are going in the right direction» (SP #234); help, especially for newly appointed school leaders, «to team up and get to know [...] the evaluation culture of the institute» (SP #168); guidance and support on prioritization and improvement planning, as seen in the fragments below.

The added value from what I have seen, and on which they are working step by step […], but according to us with significant developments in the teachers, is that of a conscious definition of priority levels and what is the priority, because in the end it is what interests us, to understand well what the priorities of the school are [SP #23].

I mean that we went from a cold document like the RAV, which took four, five days to be completed, to a process that should be part of the school institute and that accompanies us throughout the year and helps us collect as much data as possible [SP #154].

Another result noted by the SPs is that the training course provided an opportunity for awareness acquisition as well as an opportunity to acquire a habitus of reflection, as demonstrated, respectively, in the following two fragments.

We are also realizing the necessity of acquiring certain tools and the importance of occasionally pausing to reflect on the direction our school is taking, to contem-

plate, to take a snapshot of the school, to ascertain where we need to go; and we are experimenting with some reflective tools on certain practices that were already in place [SP #399].

During the visiting [...] I revisited the entire history of the school since we conducted the first RAV, and it was, for me – speaking personally, but then also confirmed by the teachers – an excellent opportunity to reflect on the school's journey [SP #456].

The STs, for their part, also when reasoning about the results, highlight more technical and punctual aspects, as exemplified below:

Reflecting and maturing new ways of analysing data [ST #933].

A specific analysis grid allowed us [...] to change even a little the way we think, to review the data, to review our forms […] of collection and monitoring [ST #1513].

And a world opened to me because we realized [...] how useful this tool can be if analysed and observed as we were taught [ST #917].

Ultimately in each course we analysed it in a slightly more precise manner and therefore this allowed us to read it and perhaps even try to implement it in such a way as to overcome our critical points [ST #2166].

Acquiring new knowledge, skills, especially regarding the formulation of new indicators [ST #823].

Moreover, the STs reflect on deeper and metacognitive aspects, as in the following excerpts.

However, in the end, no tool can be truly effective if we are not able to read them, not only with scientific skills, but with the right attitude. [...] that is, before going to monitor you need to understand what and how to monitor it [ST #737].

I realize that [...] I am aiming more at the analytical discourse, at the analytical question, because I have acquired this mentality [ST#1875].

Through this course, I also came to the awareness that starting from the same school reality, therefore from the same RAV, we arrive at two different interpretations, two photographs of the same school, because they were taken by different teachers [ST #1987].

These STs' interventions, in short, illustrate a process of capacity building, or maybe of reorganization and systematization of knowledge and skills («skills have not increased, perhaps we are simply reorganizing them, we are

recreating a map of our skills, which may perhaps be useful in the coming years», ST #2209), that starts from the processing of information through operations that the acute intervention of a teacher defines as comparison, historicization, testing in one's most typical professional context, the classroom, and that finally reach the group, the place where evaluative skills, such as relational skills of shared reflection, are really formed[14].

Regarding the sharing of outcomes and impacts on the school, we refer to the most significant point of view of the SPs, among whom most of the interventions emphasized how the project was able first of all to «set in motion experiences and practices among participants» (SP #208) but, «despite illustration in the Teachers' Board or Department meetings, it is far from involving all staff» (SP #113): «the process of – shall we say – internal spillover [...] certainly needs a much longer time frame» (SP #144).

6. Discussion

The findings of this study highlight significant differences in the perceptions and experiences of school principals (SPs) and school teachers (STs) participating in a training course focused on self-assessment practices within Italian schools. Through a comparative analysis of their contributions, several key themes emerged, highlighting disparities in thematic articulation, engagement with course content, collaboration dynamics, and perceptions of outcomes and impacts.

The observations made by SPs regarding their adherence to the training's objectives and methods shed light on their commitment to self-evaluation processes within their schools. Despite participating less intensively than teachers, SPs expressed alignment with the course's goals, emphasizing the importance of preparing training activities on evaluation for all staff. This suggests a recognition among SPs of the value of self-assessment practic-

[14] «Knowledge means putting together a lot of information [...] and forces us to a vision that is clearer, no, even more interconnected [...]. When instead I systematize information in a comparative way, that is, I re-think and then position it historically, according to the learning environment, and then reconstruct it, and then create pathways, and then arrive at an answer, then I create that change that leads to competence. This happens with the training process in our interaction with the group, but it also happens in the classroom, obviously. We would be schizophrenic if it weren't so, no. That is, and so here is the beauty of having a knowledge pathway given by trainers, who are very competent, but then a construction of skills that we carry out in the group and that makes us then more aware – capable we won't know, future generations will... – but certainly more capable of creating competence» (ST #1140).

es and their willingness to incorporate them into their school's governance processes. Moreover, their interpretation of the training course as consistent with their training needs underscores the relevance of professional development initiatives tailored to the specific roles and responsibilities of school leaders. These differences are consistent with the frameworks outlined in the introduction, where evaluation capacity building is described as a multidimensional process requiring both technical skills (teachers' focus on content and methods) and leadership competences (principals' attention to organisation and governance), in line with OECD (2013) and subsequent literature on data teams and collaborative evaluation (Hubers and Poortman, 2018; Schildkamp, 2019).

At a general level, SPs tended to allocate more attention to categories such as "Needs and expectations", emphasizing contextual needs and the importance of self-evaluation processes within a broader strategic planning framework. In contrast, as emerged in other studies (Poliandri *et al.*, 2022), STs exhibited a greater focus on technical aspects of the training course, such as learning content, teaching methods, timing, and calendar. This result is probably attributable to the fact that the SPs attended training courses in different and less intense ways than teachers, as suggested also by some results in the "Collaboration and discussion" category (Tab. 5 in par. 4.3) and by the Learning analytics in the Moodle platforms, on which the three courses were delivered (Pillera, Giampietro and Poliandri, 2023). However, both groups acknowledged a positive crisis-effect resulting from the course's departure from their initial expectations.

Disparities in collaboration dynamics between SPs and STs were evident.

While SPs may have contributed less extensively to discussions on collaboration dynamics and the quality of participation, their acknowledgment of the importance of intra-group collaboration within their schools reflects an understanding of the collaborative nature of effective school governance. Furthermore, the expressed desire for a dedicated communication channel for SPs within the project highlights their recognition of the value of peer exchange and collaboration among school leaders. This underscores the importance of fostering networks and communities of practice among educational leaders, which, as Brown and Poortman (2018) note, are crucial for promoting knowledge sharing and supporting professional growth. STs provided more precise accounts of working dynamics, emphasizing the exchange of viewpoints and peer learning methods. Both groups recognized the importance of collaboration among schools, albeit with varying levels of engagement. STs, in particular, highlighted the benefits of sharing experiences and engaging in reflective dialogue with peers and instructors. This confirms what has been highlight-

ed in previous studies on the role of collaboration and networks in fostering evaluative capacity (Ehren *et al.*, 2017; Azorín and Muijs, 2017), supporting the view that peer interaction and distributed forms of leadership are central to sustaining evaluation capacity in schools. Our findings are also consistent with Topping's (2007) evidence that structured peer learning among professionals in similar roles fosters reciprocal knowledge exchange and produces concrete improvements in professional practice.

Regarding outcomes and impacts, SPs and STs exhibited proportional representation regarding implications for data team's work and evaluation skills, as well as the sharing of outcomes and impacts on the school. SPs emphasized the course's contribution to confirming existing practices, fostering teamwork, and guiding improvement planning processes within their schools. They also appreciated the opportunity for self-reflection and the acquisition of awareness and reflective habits. Conversely, STs focused on technical aspects and the acquisition of new knowledge and skills, highlighting a process of capacity building and reorganization of skills within their professional context.

In addition to their focus on technical aspects, STs reflected on deeper metacognitive aspects of their learning experiences, such as changes in attitude, the subjectivity of evaluation, and the establishment of positive communication climates within data teams. These reflections indicate a profound engagement with the course content and a commitment to continuous professional growth. Moreover, their emphasis on the development of specific evaluation vocabulary and terminology highlights the importance of building a common language and understanding among educators to enhance collaboration and effectiveness in evaluation processes. This multiplicity of learning dimensions – from technical skills to reflective attitudes and shared evaluative language – reflects the idea that capacity building does not follow a single pathway but rather emerges through different combinations shaped by specific contexts and opportunities (Tagle and Casavola, 2004).

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

Grasping these differences based on the role makes it possible to set up targeted courses according to SPs' and STs' specific responsibilities within the evaluation and governance processes of the school: on the one hand, it is a question of strengthening SPs' competences, so that they can conduct

self-evaluation processes at the level of the entire school, and on the other hand, of supporting the teachers' evaluation and data reading for the improvement, as theorized in the preparation of VfS.

7. Conclusions

This study provides valuable insights into the perceptions and experiences of SPs and STs participating in a training course on self-assessment practices within Italian schools. By highlighting the nuanced differences in thematic articulation, collaboration dynamics, and perceptions of outcomes and impacts, the findings contribute to our understanding of effective professional development initiatives in education. Ultimately, fostering collaboration and addressing the diverse needs and expectations of stakeholders are essential for promoting continuous improvement in educational practice and enhancing student outcomes.

This discussion serves to contextualize the findings within the broader literature on educational leadership, professional development, and school improvement. It underscores the importance of understanding the unique perspectives of different stakeholders in educational settings and the need for targeted interventions to support their professional growth and development.

The findings of this study underscore the need for targeted professional development initiatives that recognize and address the diverse needs and perspectives of school leaders and teachers. By understanding the specific roles and responsibilities of SPs and STs within the evaluation and governance processes of schools, educational stakeholders can design training programs that effectively support their professional growth and enhance their capacity to drive meaningful change within their schools. Moreover, the insights gained from this study can inform the development of future professional development initiatives, ensuring their relevance and effectiveness in promoting a culture of continuous improvement within Italian schools.

Overall, these findings reinforce the literature reviewed in Sections 2 and 3, which stresses that evaluation capacity building in schools depends on combining context-sensitive professional development, leadership for self-evaluation, and opportunities for collaboration and networking (OECD, 2013; Fortini *et al.*, 2016; Poliandri *et al.*, 2019). The VfS project provides empirical evidence of how these elements can be operationalised in practice, differentiating the contributions of principals and teachers within the same evaluative framework.

The results also confirm the relevance of adopting training methodologies inspired by communities of practice (Wenger, 1998) and by reflective

learning approaches (Schön, 2006; Mezirow, 1997), which contextualize evaluation as a collective process within the school community (Brown and Poortman, 2018).

While this study provides valuable insights into the perceptions and experiences of SPs and STs participating in a training course on self-evaluation practices, it is not without limitations. The sample size and scope of the research may limit the generalizability of the findings. Additionally, the study relied on self-reported data, which may be subject to biases and inaccuracies. Future research could address these limitations by employing larger and more diverse samples and incorporating multiple data collection methods to enhance the validity and reliability of the findings.

In conclusion, the findings of this study highlight the diverse perspectives and experiences of SPs and STs participating in a training course focused on self-evaluation practices within Italian schools. Despite differences in thematic articulation, engagement with course content, and collaboration dynamics, both groups demonstrate a commitment to professional growth and the enhancement of evaluation practices within their schools. By recognizing and addressing these differences, educational stakeholders can design targeted professional development initiatives that effectively support the needs of school leaders and teachers, thereby promoting a culture of continuous improvement within Italian schools.

This study emphasizes the importance of recognizing and addressing the diverse needs and perspectives of school leaders and teachers in professional development initiatives. By understanding these differences, educational stakeholders can tailor training programs to better meet the various needs of participants, thereby enhancing the effectiveness and relevance of professional learning opportunities within schools.

References

Adu P. (2019), *A Step-by-Step Guide to Qualitative Data Coding*, Routledge, Oxford.

Azorín C., Muijs D. (2017), "Networks and collaboration in Spanish education policy", *Educational Research*, 59 (3), pp. 273-296.

Brown C., Poortman C.L. (2018), *Networks for Learning: Effective Collaboration for Teacher, School and System Improvement*, Routledge, Oxford.

Casavola P., Tagle L. (2004), "Costruzione di capacità di valutazione: riflessioni sulla recente esperienza nella valutazione dei programmi di sviluppo", *Rassegna italiana di valutazione*, 29, pp. 93-114.

Cunningham B. (1976), "Action research toward a procedural model", *Human Relations*, 29 (3), pp. 215-238.

Díaz-Gibson J., Civís-Zaragoza M., Guàrdia-Olmos J. (2014), "Strengthening education through collaborative networks: leading the cultural change", *School Leadership & Management*, 34 (2), pp. 179-200.
Ehren M.C.M., Janssen F.J.G., Brown M., McNamara G., O'Hara J., Shevlin P. (2017), "Evaluation and decentralised governance: Examples of inspections in polycentric education systems", *Journal of Educational Change*, 18, pp. 365-383.
Fereday J., Muir-Cochrane E. (2006), "Demonstrating Rigor Using Thematic Analysis: Hybrid Approach of Inductive and Deductive Coding and Theme Development", *International Journal of Qualitative Methods*, 5 (1), pp. 80-92.
Fiore B., Torelli C. (2019), *Azioni a supporto delle iniziative di valutazione e miglioramento delle scuole nei sistemi decentrati*, Working Paper INVALSI, n. 36/2019, retrieved on December 5, 2024, from https://www.invalsi.it/download2/wp/wp36_Fiore_Torelli.pdf.
Fortini F., Muzzioli P., Poliandri D., Vinci E. (2016), "Il Sistema Nazionale di Valutazione: conoscere le scuole per supportarle", *Scuola Democratica*, 2, pp. 451-466.
Gomez Paloma F., Poliandri D., Giampietro L. (2020), "Il Progetto Value for Schools. Ricerca pedagogica e learning analytics per l'autovalutazione delle scuole, Formazione & Insegnamento", *Rivista internazionale di Scienze dell'educazione e della formazione*, 18 (1), pp. 294-307.
Hubers M.D., Poortman C.L. (2018), "Establishing Sustainable School Improvement through Professional Learning Networks", in C. Brown, C.L. Poortman (eds.), *Networks for Learning: Effective Collaboration for Teacher, School and System Improvement*, Routledge Taylor and Francis Group, London, pp. 194-204.
Muzzioli P., Perazzolo M., Poliandri D., Quadrelli I. (2016), "La qualità del percorso di autovalutazione", *Scuola Democratica*, 2, pp. 421-438.
Organisation for Economic Co-operation and Development (2013), Synergies for better learning: an international perspective on evaluation and assessment, OECD Publishing, Paris.
Mezirow J. (1997), "Transformative learning: Theory to practice", *New Directions for Adult & Continuing Education*, 74, pp. 5-12.
Pagani V. (2020), *Dare voce ai dati: l'analisi dei dati testuali nella ricerca educativa*, Edizioni Junior, Parma.
Pillera G.C., Giampietro L., Poliandri D. (2022), *Esiti e impatti Vfs rilevati nei focus group docenti*, Convegno nazionale AIV, 21-23 settembre, Pescara.
Poliandri D., Giampietro L. (a cura di) (2023a) (in press), *Sostenere la cultura valutativa delle scuole*, FrancoAngeli, Milano.
Poliandri D., Giampietro L. (a cura di) (2023b) (in press), *La formazione a supporto dei processi di autovalutazione delle scuole: cosa funziona e perché*, FrancoAngeli, Milano.
Poliandri D., Perazzolo M., Pillera G.C., Giampietro L. (2023), "Dematerialized participation challenges: Methods and practices for online focus groups", *Frontiers in Sociology*, 8, 1145264.
Rivoltella P.C. (2013), *Fare didattica con gli EAS. Episodi di Apprendimento Situato*, La Scuola, Brescia.

Robasto D. (2016), "Una nuova expertise del corpo docente a supporto dell'autovalutazione scolastica", *Form@re – Open Journal per la formazione in rete*, 2 (16), pp. 121-136.
Rossi P.G. (2011), *Didattica enattiva. Complessità, teorie dell'azione, professionalità docente*, FrancoAngeli, Milano.
Schildkamp K. (2019), "Data-based decision-making for school improvement: Research insights and gaps", *Educational Research*, 61 (3), pp. 257-273.
Schildkamp K., Poortman C.L., Handelzalts A. (2016), *Data teams for school improvement, School Effectiveness and School Improvement*, 27 (2), pp. 228-254.
Schön D. (2006), *Formare il professionista riflessivo. Per una nuova prospettiva della formazione e dell'apprendimento nelle professioni*, FrancoAngeli, Milano.
Vanhoof J. ,Verhaeghe G., Petegem P.V., Valcke M. (2013), "Improving data literacy in schools: Lessons from the school feedback project", in K. Schildkamp, M.K. Lai, L. Earl (eds.), *Data-Based Decision Making in Education: Challenges and Opportunities*, Springer, Dordrecht, pp. 113-134.
Wenger E. (1998), *Communities of Practice: learning, meaning and identity*, Cambridge University Press, Cambridge.

Appendix. Fragments of speech from online focus groups of teachers and school principals in the original language

Devo dire che abbiamo fortemente voluto questo progetto proprio perché ci serviva per mettere a punto tutta quella serie di riflessioni, meta-riflessioni avviate dentro la scuola [SP #337].

[…] è quella di creare un gruppo che semini, ecco, che sia un gruppo portatore di questa cultura dell'autovalutazione perché, ripeto, nel mio istituto purtroppo io ho dovuto constatare che è solo un adempimento formale [SP #66].

Forse noi ne avremmo avuto bisogno, di questo intervento così importante e anche dettagliato nel suo percorso, già dal 2013-2014, quando noi abbiamo cominciato con il RAV, con il Piano Triennale dell'Offerta Formativa, con il piano di miglioramento, che rimane ancora il settore più ostico da affrontare. Perché è lì che, nonostante andiamo faticosamente ad individuare quali sono, come dire, le priorità di cui tener conto, poi facciamo una grande fatica a mettere in atto quelle azioni orientate verso quel miglioramento [SP #29].

Il tema dell'autovalutazione è un tema complesso e anche tecnico. È vero che ci sono le guide del ministero, però, insomma, affrontare un RAV, cos'è un RAV, costruire il RAV o poterlo applicare è qualcosa di abbastanza complesso. Speravo come componente del nucleo di autovalutazione che il corso fosse anche una un'occasione non solo di riflessione, ma anche per apprendere l'aspetto tecnico [ST #2065].

Forse noi da questo progetto ci saremmo aspettato un qualche cosa più di tecnico, cioè che ci avessero dato magari […] un qualche format, un qualche cosa su cui poter ragionare, lavorare, pensare e invece ci sembra di capire che ci viene chiesto lentamente di costruirlo [ST #1493].

Per amor di paradosso, dico che allora il corso non ha risposto alle mie aspettative, perché io, da dirigente neonominato, mi aspettavo un aiuto acché mi spiegasse come si fa il RAV. Ecco, grazie al cielo, non è stato questo, appunto è stato prezioso [SP #172].

Io sinceramente all'inizio pensavo sicuramente che sarebbe stato un corso pieno di griglie, di tabelle, tutto strutturato. Infatti, io sinceramente mi aspettavo qualcosa di piatto e invece sinceramente a me, con tutte le criticità che comunque le mie colleghe hanno espresso benissimo, perché in fondo in fondo anch'io mi rendevo conto del disorientamento che ci portava, però a me il corso è piaciuto dall'inizio, perché era quello che mi aspettavo, qualcosa di destrutturato [ST #1670].

È bella anche la struttura, nel senso che c'è chi approfondisce la lezione, anche se chiaramente a distanza, ma comunque si ascoltano delle persone e quindi nella testa si rimettono a posto una serie di informazioni, poi si rielabora nel gruppo e poi c'è una sintesi che però è una sintesi fatta da più voci che vengono da più scuole e questa è un po' una novità perché il RAV è sempre stato invece forse una cosa interna alla scuola da gestire con le funzioni strumentali che se ne occupano. E a me è piaciuto anche molto la piattaforma, nel senso che c'è il materiale da visionare molto chiaro, ma c'è anche il modo, il tempo e la richiesta di mettere dentro delle informazioni che costringono non solo ad ascoltare i video oppure a leggere, ma a mettere del proprio e a leggere quello che scrivono gli altri e questa la trovo una cosa interessante perché col tempo uno rilegge, riguarda, scopre, e quindi è un percorso che uno fa anche da solo, ma vive a casa [ST #1067].

Allora, io ritengo che il punto di forza di questo percorso di formazione stia nella modalità di conduzione, ovvero sicuramente è un corso che ha puntato molto sull'autoriflessione e sulla presa di coscienza delle pratiche di autovalutazione, ponendo noi operatori in modo critico, ponendoci un po' in modo critico e andando a rivedere le nostre… e ci ha permesso di andare a rivedere le nostre pratiche [ST #1362].

Le esercitazioni sono di aiuto? Assolutamente sì, sono, come dire, l'artefatto pedagogico che ti permette di fare autoanalisi e che ti permette di capire dentro al tuo istituto, come sta andando, a che punto sei. Ti fermi e rifletti [ST #1416].

L'altra, diciamo, polarità formativa, che comunque, al momento il progetto sicuramente ha intercettato, ma sulla quale poi dobbiamo lavorare ancora molto è proprio questa necessità di trovare delle modalità efficaci per comunicare l'autovalutazione, ma con un linguaggio che sia accessibile, ovviamente anche all'utenza, una sorta di autovalutazione semplessa, non so come dire, perché di fatto la complessità

che caratterizza comunque il linguaggio anche tecnico e che spesso è difficile da far capire anche ai docenti, si traduce in un'altrettanta complessità nel comunicare poi appunto quelli che sono i processi interni [SP #144].

E poi ci tengo a dire che molto importante, almeno per noi, come gruppo, è stato anche il lavoro importante che è stato fatto su tutto il discorso del clima di lavoro, dell'importanza delle relazioni, nel nostro gruppo si è lavorato molto bene su questo [SP #315].

Ritengo che tutto il materiale rappresenti una miniera ovviamente, quindi di ottimo livello [SP #588].

Allora, la pandemia soggettivamente ha cambiato molto il clima delle scuole. Il progetto era stato impostato ed è stato ripensato direi in modo mirabile, valorizzando tutti gli strumenti possibili, credo anche con una fatica organizzativa immane, ma è stato riorganizzato nel modo migliore possibile. Il problema è che il contesto nel quale si cala è diverso dal contesto nel quale è stato pensato [SP #379].

Il fatto di poter usufruire di questo corso anche a distanza mi ha consentito di essere più presente, di partecipare al corso. Non so cosa sarebbe successo nell'eventualità che avessimo dovuto fare tutto in presenza, la mia partecipazione, forse non sarebbe stata così importante [SP #66].

Invece, il fatto di dover fare, concretamente, provare nel campo, cercare alternative, cercare strategie che non riguardano la propria situazione alla quale ormai siamo un pochino assuefatti, ma situazioni differenti, devo dire l'ho trovato molto molto stimolante perché riesce a farti entrare, in un'ottica anche di responsabilità, di complessità differente, ecco quindi questo mi è… è la parte che più mi è piaciuta di questo corso, ecco [ST #1085].

Quindi la cosa che si è creata, soprattutto nella mia scuola è stata che ci sono stati momenti che non venivano vissuti come formazione perché erano più momenti in cui si guidava e si aiutavano i gruppi, che d'altra parte fa parte poi di quelle tecniche di peer-to-peer, che noi tutti abbiamo studiato [SP #4].

Come gruppo, quindi, ci siamo incontrati, anche più di una volta, c'è stato un confronto man mano che procedevamo lungo il percorso, che ci ha permesso pure di discutere, di analizzare, anche di cogliere, appunto, quelle perplessità iniziali, capire quindi perché c'erano e poi procedere e riflettere, insomma, su quello che era il processo in corso ed entrare, questo ci ha permesso proprio di penetrarlo bene [SP #602].

La nostra dirigente, oltre ad aver seguito abbastanza frequentemente gli incontri, ha creato anche proprio una chat di tutti i partecipanti del corso, dove noi ci possiamo confrontare, e comunque monitora continuamente quelle che sono, diciamo, le

situazioni che si vengono a creare, quelli che possono essere appunto magari o spunti di riflessione o magari anche delle cose, delle criticità che noi poniamo in evidenza e lei comunque ha un canale preferenziale con i formatori, per cui magari se abbiamo portato avanti delle richieste, lei si è comunque attivata perché queste richieste venissero portate ai formatori. Per cui c'è un continuo monitoraggio [ST #2152].

Siamo un gruppo folto, numeroso, quindi abbiamo avuto l'occasione di avere, durante gli incontri, degli scambi attivi in cui confrontarci, anche per… cioè per avere alcuni dubbi sui documenti, perché ovviamente chi è parte integrante del NIV e ha partecipato alla stesura dei vari documenti può avere anche un occhio diverso, al contrario di altri colleghi che hanno scelto, per esempio, di entrare nel corso per conoscere e avere strumenti operativi per poi calarsi per la prima volta in questo ambito [ST #2086].

Una riflessione rispetto all'interpretazione diversa che ogni istituto può dare in merito all'output da produrre [...]., la modalità di lavoro ha permesso all'interno del NIV di riflettere sulle modalità diverse con le quali i diversi istituti, diversi gruppi interpretano e si pongono di fronte ad una richiesta, proprio per una questione anche di – come dire – volontà di approfondire un determinato tema; nel senso che la richiesta era molto chiara, però per noi ha significato andare in una certa direzione perché volevamo utilizzare e mettere in campo qualcosa che serve a noi in questo momento [ST #1410].

E mi piace molto anche l'idea di andare a verificare, a vedere quello che hanno fatto gli altri istituti per operare dei confronti e per capire effettivamente quelli che sono i nostri punti di forza e quelli che sono invece le cose in cui siamo un pochino più carenti [ST #1366].

Dall'altro lato, direi che anche l'aspetto dello scambio con altre realtà scolastiche è fondamentale. Noi abbiamo partecipato come osservatori a un visiting. Prossimamente, siamo noi scuola ospitante per un visiting. Anche questo è un'opportunità per la scuola di riflettere su come lavora, quali obiettivi va perseguendo, su quali e linee di fondo si muove e quindi è una sorta di autovalutazione costante questa. Noi abbiamo improntato appunto il nostro visiting proprio prendendo in considerazione l'esperienza innovativa, per così dire, della scuola dell'infanzia, proprio per valorizzare questa partecipazione, perché il problema più grosso, prima si è parlato di motivazione, è stimolare la motivazione da parte degli insegnanti [SP #82].

Quindi, secondo me, la cosa su cui ha risposto molto rispetto alle nostre aspettative, il progetto è stato quello di permetterci di uscire dall'autoreferenzialità, quindi, il RAV non ce lo leggiamo e ce lo cantiamo solo tra di noi, ma lo condividiamo anche con altre scuole; quindi, diventa davvero un motivo di scambio e di confronto, di vedere cosa fanno altri, come fanno altri, e quindi questo senz'altro è l'aspetto più positivo che noi abbiamo rimarcato di questo progetto [SP #168].

Perché nel lavoro di gruppo […] si parte da un contenuto comune, ma poi si incrociano e si intrecciano e si incontrano tante sensibilità diverse che vengono […]. da esperienze e peculiarità delle nostre scuole, ma anche della nostra persona. E quindi partecipare a un gruppo di adulti che partono da un punto comune di partenza è… vuol dire mettersi in gioco in maniera diversa. E quindi un po' si fatica, un po' ci si riorienta, un po' si fa memoria delle difficoltà, un po' si va a cercare magari piccoli pezzettini che si sono persi [ST #1121].

Sicuramente è un momento di confronto importante e anche sentire dirigenti, che magari hanno da più anni esperienza nella gestione, nella guida della scuola e si trovano in contesti anche molto diversi [SP #226].

Ho già fissato un momento dopo questo confronto con gli altri dirigenti per lavorare proprio sull'applicazione di alcuni di questi apprendimenti, di questi percorsi a quello che è la nostra situazione concreta, quindi l'applicazione del modello […] alla situazione particolare [SP #80].

Il valore aggiunto da quello che ho visto, e sul quale stanno lavorando […] passo dopo passo, ma secondo noi con evoluzioni significative nei docenti, è quello di una definizione cosciente dei livelli di priorità e di quello che è la priorità, perché poi alla fine è quello che ci interessa, capire bene quali sono le priorità della scuola [SP #23].

Cioè voglio dire che si è passati da un documento freddo come quello del RAV, che ci metteva per essere svolto quattro, cinque giorni a un processo che dovrebbe essere dell'istituto scolastico e comunque che accompagna nell'anno e ci aiuta a raccogliere più dati possibile [SP #154].

Stiamo acquisendo anche l'idea della necessità di dotarci di alcuni strumenti e della necessità di fermarsi ogni tanto a riflettere su dove sta andando la nostra scuola, a riflettere, a farci una fotografia della scuola, per vedere dove dobbiamo andare; e stiamo sperimentando alcuni strumenti di riflessione su alcune pratiche che già venivano effettuate [SP #399].

In occasione del visiting […] ho rivisto tutta la storia della scuola da quando abbiamo fatto il primo RAV ed è stato per me – parlo a livello personale, ma poi anche i docenti mi hanno confermato questa cosa – un'ottima occasione per riflettere sul cammino della scuola [SP #456].

Riflettere e sviluppare nuovi modi di analizzare i dati [ST #933].

Una griglia di analisi specifica ci ha permesso [...] di cambiare anche un po' il nostro modo di pensare, di rivedere i dati, di rivedere le nostre forme [...] di raccolta e monitoraggio [ST #1513].

Un mondo si è aperto per me perché abbiamo realizzato [...] quanto utile possa essere questo strumento se analizzato e osservato come ci è stato insegnato [ST #917].

In ultima analisi, in ogni corso lo abbiamo analizzato in modo leggermente più preciso e quindi ciò ci ha permesso di leggerlo e forse anche di provare a implementarlo in modo tale da superare i nostri punti critici [ST #2166].

Acquisizione di nuove conoscenze, abilità, specialmente per quanto riguarda la formulazione di nuovi indicatori [ST #823].

Tuttavia, alla fine, nessuno strumento può essere veramente efficace se non siamo in grado di leggerlo, non solo con competenze scientifiche, ma con l'atteggiamento giusto. [...] ovvero, prima di procedere al monitoraggio, è necessario capire cosa e come monitorarlo [ST #737].

Mi rendo conto che [...] sto puntando sempre di più sul discorso analitico, sulla domanda analitica, perché ho acquisito questa mentalità [ST#1875].

Attraverso questo corso, sono anche arrivato alla consapevolezza che partendo dalla stessa realtà scolastica, quindi dallo stesso RAV, si giunge a due interpretazioni diverse, due fotografie della stessa scuola, perché sono state scattate da insegnanti diversi [ST #1987].

La conoscenza vuol dire mettere a sistema una marea di informazioni […]. e ci costringe a una visione che è più chiara, no, anche più interrelata […].. Quando invece le informazioni io le metto a sistema in forma comparativa, cioè le ri-ragiono e poi le posiziono a livello storico, a seconda dell'ambiente di apprendimento, e poi le ricostruisco, e poi creo percorsi, e poi arrivo a una risposta, allora creo quel cambiamento che dà la competenza. Questo succede col percorso di formazione nel nostro interagire col gruppo, ma succede anche in classe, ovviamente. Saremmo schizofrenici se non fosse così, no. Cioè, e quindi ecco è bella questa cosa di avere un percorso di conoscenza che è data dai formatori, che sono molto competenti, però poi una costruzione delle competenze che noi svolgiamo nel gruppo e che ci rende poi più consapevoli – capaci non lo sapremo, i posteri… – ma sicuramente più capaci di creare competenza [ST #1140].

4. Small schools for self-evaluation

by Graziella Arazzi

The contribution intends to highlight the extent to which small schools (small school sites far from the centre, multi-class, mixed classes) – in their structural morphology and in the dynamics that unfold within them – contribute to strengthening attitudes and styles of self-evaluation within the comprehensive institutes in which they are placed. With the Small Schools Movement, INDIRE has stimulated research patterns and reflections on the nature of an organizational and management form (not just teaching) which highlights the possibility of developing cultures and practices of widespread self-evaluation, made evident by the following factors: planning for groups; the possession of active research methodologies on the territories; the use of discontinuous times in teaching; the construction of educational communities, fueled by informal networks with a focus on digital strategies; teacher training understood as peer research and experimentation with new professional contexts. In this context, however, to date, there has been almost an absence of research on the evaluation aspects and on the contribution that micro-schools (fluid structures, civic centers or training hubs) can provide to the institutions they belong to.

Il contributo intende mettere in evidenza in quale misura le piccole scuole (plessi ridotti e lontani dal centro, pluriclassi, classi miste) – nella loro morfologia strutturale e nelle dinamiche che si snodano al loro interno – contribuiscano a rafforzare atteggiamenti e stili di autovalutazione all'interno degli istituti comprensivi in cui sono inserite. INDIRE, con il Movimento delle Piccole Scuole, ha suscitato trame di ricerca e di riflessione sulla natura di una forma organizzativa e gestionale (non solo didattica) che evidenzia la possibilità di sviluppo di culture e pratiche dell'autovalutazione diffusa, rese evidenti dai seguenti fattori: la progettazione per gruppi; il possesso di

metodologie di ricerca attiva sui territori; l'uso di tempi discontinui nella docenza; la costruzione di comunità educative, alimentate da reti informali con focus sulle strategie digitali; la formazione dei docenti intesa come ricerca tra pari e sperimentazione di nuovi contesti professionali. In tale contesto, tuttavia, fino ad oggi è risultata quasi assente la ricerca sugli aspetti valutativi e sul contributo che le micro-realtà scolastiche (strutture fluide, centri civici o hub formativi) possono fornire agli istituti di appartenenza.

1. Introduction

Small schools (multi-classes, isolated centers with a small number of students, classes in internal areas) have been at the center of numerous investigations regarding the planning, documentation and professional development of teachers (propensity to change and epistemic agency) but they are poorly analyzed as engines of expansion of the self-evaluative culture. However, the fact that INDIRE formulated the role of tutors for small schools in 2019 – within the National Repertoire of Professionalism for Innovation – may signal the beginning of a new sector of research, in which the realities investigated create processes of improvement, not so much by overcoming critical factors as by recognizing the specificity of values and points of view with which self-evaluation is played. By encouraging the analysis of the impact and rooting of the innovations tested (multi-age, multi-level, combined classes) and by triggering relationships between isolated groups and other locations, the tutor – trainer and facilitator – «participates in the school's self-evaluation processes and commits to redefine its intervention in relation to improving the organization of the small school» (INDIRE programmatic document, 2019). Small groups of comprehensive schools find in the figure of the tutor, currently optional, help to trigger self-evaluation processes which could be critical in many situations in which the governance of the schools is entrusted to regent managers. The INDIRE repertoire also refers to «an active and reflective approach», with which the tutor guides colleagues in understanding the contribution of innovative methodologies and techniques in supporting situations of isolation. The online training created by INDIRE for teachers of small schools in the pandemic phase was subjected to an evaluation process aimed at providing «a measure of the ability to activate professional, individual and community development» (Mangione *et al.*, 2020, p. 151). The evaluation of online training – with the aim of understanding its effect on the genesis of specific skills but also on the maturation of new and more effective organizational-management

models – was based on the involvement of three categories: the beneficiary teachers; school managers (with whom the enhancement of training was investigated in the assignment of functional tasks to improve the mission of training institutions); the experts who had designed the training intervention, analyzed in terms of transfer, rooting and impact. From the evaluation of the laboratory model (contents, deliveries, training agreements between trainees and trainers) we proceeded with the analysis of factors and phases of strengthening the school system: loans of professional resources, cooperative forms of work, effects inside and outside the school, modeling of innovation processes across multiple complexes (central and peripheral) of very complex comprehensive schools. However, the link between the work of coordinators of multi-class/isolated classes and the growth of self-evaluative values still appears barely sketched.

2. Research object and hypothesis

The research takes place in the Ligurian territory and arises from the collaboration between the ten training hub schools and the regional education office. The research hypothesis is that where the NIV (Internal Evaluation Unit, from now on referred to in the text as NIV) involves the teachers of small schools or includes such teachers within it, the following dimensions can be easily achieved: climate of sharing and comparison; mutual exchange on critical issues; positive elements and lines of improvement; greater effectiveness in governing self-evaluative processes; construction of reflective practices common to the various school levels; expansion of critical thinking on the dynamics of the entire school[1]; strengthening the ability to read data and identifying new indicators; creation of tools to define priorities and goals in a more understandable and coherent way; perception of training needs on the communication of self-evaluation and the negotiation of its various phases.

The data used in the survey refer to two activities:

– a focus group – in telematic mode – dedicated to comprehensive schools including micro-realities (multi-classes, schools far from the center or in inland areas). The managers were asked to involve a member of the NIV and a teacher working in a small class in the discussion table. There were 27 participating institutes. In 8 cases the school managers themselves also intervened, indicating forms of distributed leadership and taking inspiration from the peer discussion table technique;

[1] https://www.invalsi.it/value/index.php.

– analysis of the responses to an in-depth questionnaire, administered – in a subsequent phase – to 27 teachers in service in the comprehensive schools involved in the online focus group. Respondents were identified by self-nomination, belonging to three specific categories: contact person or teacher in multi-classes, small schools, mountain schools; members of the NIV; both that is, members of the NIV and teachers in isolated groups. The questionnaire items were constructed following some indicators of the training courses of the *ValuE for Schools* project.

3. The online focus group: a self-evaluation table

The textual analysis – conducted on judgments and opinions expressed in the focus group by school managers, teachers of small or multi-class schools, members of the NIV – made it possible to examine and interpret the specificity of different points of view, emerging in the articulation of the proposed narrative and discussion topics. This was an exploratory qualitative research, conducted to intercept the central question of the investigation, namely whether the presence of small schools within comprehensive institutes in Liguria actually contributes to strengthening self-evaluation skills. The reflections were conducted on the corpus deriving from the transcription of the statements (statements, expressions, points of view) that emerged within the focus group which involved representatives of 27 comprehensive institutes (4 from SP, 3 from SV, 10 from IM and 10 from GE), with a sample varied by Institute and globally distinct: 23 members of the NIV (of which 8 school directors), 15 teachers working in the complexes with small schools, 3 teachers in charge of multi-classes and also belonging to the NIV of the relevant place of service. A total of 41 subjects participated in the focus group. During the focus group, 8 interventions by school managers and 72 interventions by teachers were recorded. Each intervention represented a case, which was associated with a series of background variables descriptive of the sample, including the professional role. The sample was saturated, as it adequately allowed the exploration of the thematic categories under study.

The online focus group was fueled by a wide series of probe questions, explained below:

– How does the school self-evaluation take small schools into account?
– To what extent do small schools promote assessment skills for school staff?
– In small schools, what approaches, methods and tools exist for strengthening self-evaluation skills?

– Constraints and opportunities, internal and external to the school: is shared reading possible?;
– Role of networks in various territories (Giampietro, Poliandri and Romiti, 2020);
– New professional skills that are necessary in small schools;
– New training models for teachers and operators of isolated centers;
– Evaluation and improvement: spaces for collaborative study, with the aim of supporting reflective processes (for example, through peer to peer, virtual learning environments, the creation of alliances with local entities and agencies).

During the investigation it was found that further meetings would not have led to an enrichment of knowledge of the phenomenon being analyzed. The interventions were recorded and transcribed with a pen/paper coding process. A process of decomposition and segmentation of the focus expressions was implemented, obtaining fragments, to which codes were attributed that explain their meaning. We then proceeded by connecting the semantically close codes and gradually generating general concepts. Using an inductive approach, from individual judgements, opinions and representations we moved on to construct interpretative categories. Since this is exploratory research, it is underlined that the interpretative categories thus constructed do not allow the formulation of generalizations and transfers to similar populations or contexts.

The focus group's stimulus questions were constructed in correspondence with some first level categories, set out below:
– critical issues and levers of development of comprehensive institutes with small school sites/multi-classes/isolated locations;
– professional profiles;
– territorial alliances;
– communicative approaches for self-evaluation;
– evaluation training;
– types of collaboration between NIV and representatives of small schools.

At a detailed level, the following micro-categories subsequently emerged from the focus group, also highlighted by empirical elements present in the schools involved in the participatory research activity:
– self-evaluative reflexivity of the NIV;
– involvement of colleagues through differentiated procedures of widespread self-evaluation;
– definition of improvement paths, following the self-assessment;
– role of the working commissions that revolve around the NIV;
– design of strategies to connect the school sites;

– increased communication of self-evaluation in comprehensive institutes with small schools;
– how to transfer self-assessed sensitivity to other colleagues (especially new hires);
– consideration of the RAV as a tool for promoting continuity between levels of education, also through the introduction of new indicators;
– interpretation of self-assessment as teamwork;
– experimentation with innovative methods of self-evaluation in small schools;
– focus on empathy and comparison as necessary elements to translate self-assessment into training;
– collaboration in self-evaluation processes;
– shared and participatory organization, with an increase in the number of staff figures;
– continuous process between organization, management and evaluation of the training structure;
– ideas and values for self-evaluation that come from decentralized schools and which enrich comprehensive schools;
– articulation of the local curriculum as a lever for the development of self-evaluation.

If we consider the first level categories, the analysis of the focus group coding segments – from a comparative perspective between school managers and teachers – does not highlight many differences between the various professional profiles regarding the formulations of the group interview/discussion. However, the differences between NIV members, teachers in charge of small schools and managers become apparent if we consider the judgments and comments expressed with respect to the coding subcategories set out above. Proceeding with a comparison between the opinions expressed by the members of the various NIVs and by the teachers in charge of small groups, interesting remarks emerge, capable of capturing dissonances and diversified opinions:
– many NIV teachers declare that they have improved their self-evaluation process, thanks to comparisons with colleagues who serve in multi-classes and isolated schools;
– on the other hand, some teachers in the latter situations claim a lack of involvement in self-evaluation processes;
– all teachers (both in central classes and in small schools) highlight the need for peer training on self-evaluation, capable of guaranteeing professional comparison and co-evolution and of allowing relationships between planning and evaluation of outcomes and processes. The recurring

expression in the focus group scenario is significant: «ask those who have more experience than you»;
- all participants underline the urgency of developing training chains, with the support of scientifically accredited research bodies and working groups;
- many teachers of small schools demonstrate the difficulty of their work in peripheral areas, with the emergence of critical factors for self-evaluation. Consequently, they highlight the need to create coordination figures between isolated groups and central nuclei for the diffusion of analysis processes, the transfer of materials, the development of empowerment paths of priority definition practices and improvement goals (Mangione, 2022);
- the teachers of small schools – compared to the NIV and the managers – highlight "a large amount of work" in the self-evaluation discourse which, however, instead of constituting an obstacle, can nourish the pedagogical and organizational heritage of the schools;
- they point out that the NIV often does not operate concretely, limiting the process to little feedback with colleagues from the decentralized school sites;
- some teachers believe that an objective of the Improvement Plan in small schools is the implementation of laboratory activities, with a view to facilitating active learning for both students and operators;
- for NIV and other teachers, the evaluation of the school depends on the planning of constant and continuous communication flows, favored by the provision of digital support tools and participatory reflection paths (e.g. video club meetings).

Analyzing the observations of the school managers participating in the focus group, it is noted that – compared to the teachers of the NIV or small schools – those involved in school governance do not highlight particular critical issues in the context dimensions. The opinions formulated by the managers are in line with the opinions expressed by teachers in isolated or multi-class groups, demonstrating awareness of the importance of generating territorial networks and signaling, on the contrary, a large detachment from the members of the NIV, for whom the inter-institutional alliances they still remain in the shadows or are configured only in embryo. Distance from the coast, isolation, infrastructural difficulties therefore do not represent for managers a deterrent to the rooting of the teaching staff or the cause of an often highlighted turnover. They underline how many newly hired teachers, in exodus from small schools, then tend to return, developing mature professionalism and awareness of management processes. They

highlight the importance of distributed governance between central and peripheral groups, with the establishment of formalized working groups on self-evaluation and the implementation of exchanges between teachers/staff/system figures during the drafting of the RAV. Furthermore, they agree with multi-class teachers on the fact that self-evaluation is a fruitful experiment, capable of bringing schools, bodies and institutions together in promoting self-efficacy. In the statements of the managers there are no indications on training hypotheses, unlike what is looming for NIV and teachers in small schools who highlight the importance of introducing the evaluation culture in the probationary year of newly hired teachers, especially with the predisposition of visiting itineraries in innovative structures. For those who exercise governance, it is considered essential that teachers must possess excellent organizational skills and master, in addition to teaching methodologies, the innovation of learning disciplines in order to deal with emergency situations, which are more frequent in isolated schools than in standard classes. The various actors involved in the discussion believe that small schools perform a social reorientation function for pupils (open classes) and adults (families but also companies, libraries, municipalities). For everyone, it is necessary to strengthen self-analysis in teams and dedicate energy to building networks for self-evaluation between schools of different levels (RAV longitudinal process). Furthermore, the urgency of creating hybrid networks with associations and Third Sector organizations is highlighted, especially in areas where first-generation foreign users are strong. In many cases, exploiting the potential of already existing socio-cultural ties, the construction of systemic social reporting processes is hypothesized. Teachers and managers indicate the need to increase the evaluative eye also on the PNRR measures that the school develops and which are included in the RAV, with particular reference to the actions developed to combat school dropout (Decree of the Ministry of Education no. 170 of 24 June 2022 – "Actions to prevent and combat school dropout") and to the activation of School 4.0 training courses (Decree of the Ministry of Education no. 161 of 14 June 2022).

4. Projections from the focus group

Many teachers of small schools, also belonging to the NIV, point out that the shared reflection between central and peripheral groups is undoubtedly an opportunity for improvement. They require specific training on the following dimensions:

– data literacy;
– ability to reflect and argue judgments better;
– ability to identify priorities and improvement paths;
– ability to analyze data and trace connections between the school's weak points and action priorities.

Generally, in the new 2022-2025 cycle, some teachers of multi-classes or isolated schools were involved in the drafting of the RAV and also of the PTOF. Possessing a good knowledge of the constraints and opportunities of the territory, they were delegated to construct perception questionnaires, aimed at families, local authorities or associations. The presence of many of them in commissions preparatory to self-evaluation made it possible to refine the interpretation of the data and focus on improvement plans. According to some members of the NIV – not teachers in small schools – the involvement of the latter in the paths of reflection for change has made it possible to intercept unprecedented evaluation needs, leading to reflection on the experimentation of strategies to achieve improvement goals. The creation of face-to-face or remote discussion groups has led to an effective sharing of critical points and development hubs. Active listening skills and proactive collaboration have improved, with a significant decrease in competitive or conflictual dynamics. On the other hand, the peer review process in self-evaluation needs to be consolidated. Generally, the moment of drafting the RAV was seen as the generation of a powerful link of cohesion between the various school sites. Only in some cases was the presence of a self-assessment network found, which includes schools, local companies and libraries. Self-assessment is perceived as "getting involved, asking those who have more vision and experience" but at the same time it outlines a space in which multi-classes find recognition for their constant mediation work, in overcoming unexpected events, risks and obstacles. Among the training methods on self-evaluation, small school teachers report: collaboration between different grades; peer learning and visiting; mobile and variable groups with intersection spaces; moments of co-planning with stakeholders (beyond the logic of accountability).

5. Emergencies from the questionnaire: perception of the evaluation skills of the school community and the role of small schools in strengthening self-evaluation

The questionnaire was available at the following link: https://docs.google.com/forms/d/e/1FAIpQLSfhkxoWs1NnkIN2JmsNGncDgMWFaFxXFixPZEPHlta7AshkNg/viewform.

It is considered appropriate to examine the responses to the various items, both globally and by subsequently disaggregating the results by three categories of respondents defined as follows:

– referent or teacher in multi-classes, small complexes, mountain schools;
– members of the NIV;
– both.

The question "What are the evaluation skills that your institute possesses?" involved the possibility of several choice options, to be made from eight items:

– reading and data analysis INVALSI tests, school questionnaire, other sources;
– construction of new self-assessment tools;
– inclusion of new school evaluation indicators;
– reflective and interpretative skills;
– ability to motivate self-assigned levels for outcomes and processes in the RAV;
– ability to identify priorities and paths for improvement;
– ability to identify strengths and weaknesses;
– ability to find links between weaknesses and priorities for action.

If we consider the choices of the entire audience, the following map of recognized evaluation skills emerges and reported in the following list, in order of the number of occurrences expressed (in round brackets):

– ability to identify strengths and weaknesses (23);
– reading and analysis of data INVALSI tests, school questionnaire, other sources (24);
– ability to identify priorities and paths for improvement (24);
– ability to find links between weaknesses and priorities for action (18);
– ability to motivate self-assigned levels for outcomes and processes in the RAV (12);
– reflective and interpretative skills (12);
– inclusion of new school evaluation indicators (6);
– construction of new self-assessment tools (4).

Proceeding with the analysis by individual category of respondents, the dimensions most perceived by teachers belonging to the NIV and at the same

time responsible for isolated complexes are the following (in parentheses the various occurrences):
– ability to identify priorities and paths for improvement (8);
– reading and data analysis INVALSI tests, school questionnaire, other sources (6);
– ability to find links between weaknesses and priorities for action (6);
– ability to identify strengths and weaknesses (6);
– ability to motivate self-assigned levels for Outcomes and processes in the RAV (5);
– reflective and interpretative skills (4);
– inclusion of new school evaluation indicators (2);
– construction of new self-assessment tools (2).

The members who belong exclusively to the NIV, by projecting their data literacy activities on the school context, in a sort of global vision, deviate from the previous framework of competences, recognizing the school in which they operate with the following evaluation skills:
– ability to identify priorities and paths for improvement (10);
– reading and analysis of data INVALSI tests, school questionnaire, other sources (10);
– ability to find links between weaknesses and priorities for action (8);
– ability to identify strengths and weaknesses (8);
– ability to motivate self-assigned levels for Outcomes and processes in the RAV (7);
– reflective and interpretative skills (3);
– inclusion of new school evaluation indicators (2);
– construction of new self-assessment tools (1).

Indirectly, the number of choice options reveals strong self-efficacy but also the conviction of playing a decisive role in the self-evaluation of the school, without the need to open up to the collaboration of the teachers of the isolated complexes and to generate and develop a peer evaluation of the school. Finally, the referents or teachers in multi-classes, small complexes, mountain schools, illustrate a map of the school's evaluation skills that appears distributed, without excessive overlapping of choice, on seven of the eight categories of evaluation competence defined in the survey. From the survey of occurrences, the "Ability to identify strengths and weaknesses" (9) is more attributed to the school institution. At the school, the teachers of the isolated complexes recognize a varied set of evaluation skills, expressed in the following table:
– reading and analysis of data INVALSI tests, school questionnaire, other sources (8);

– ability to identify priorities and paths for improvement (6);
– reflective and interpretative skills (5);
– ability to find links between weaknesses and priorities for action (4);
– inclusion of new evaluation indicators (2);
– construction of new self-assessment tools (1).

Compared to the other two categories of respondents, teachers of small schools believe that they do not attribute to the school the «Ability to motivate the levels self-assigned for outcomes and processes in the RAV». It is therefore necessary to find the reasons for this perception and for this two different reasons are highlighted:
– the fact that the teachers/contact persons of small schools are still developing a self-evaluation path, which is changing and needs system reorganization;
– the fact that these teachers are not completely involved in the self-evaluation by the NIV or by school principals.

The second motivation appears to be consistent with the judgments and opinions expressed within the focus groups. The data of the online questionnaire shows that both teachers working in small schools and NIV members value two dimensions:
– the need for training chains with the accompaniment of research structures (INVALSI is mentioned as a priority);
– the importance of peer training that allows for shared communication and documentation of processes and materials.

However, while the NIV approach tends to be merely formal or in line with standard training, administered by institutions and universities (in a predominantly frontal and standardized manner), that of teachers of small schools refers to innovative training, contextualized and rooted in the complexity of operating in situations of isolation and emergency and articulated in the following phases: laboratories, workshops, action-research, peer review, communities of practice, social networking, skills mapping. In a balanced synthesis of face-to-face activities, personal study, reflection and documentation, networking, and feedback to the entire comprehensive institute, teachers of small schools appear aligned with the indications of the "National Teacher Training Plan – PNFD" inaugurated in the three-year period 2026-2019.

In addition, teachers in small schools – compared to the other two groups of interviewees – stand out for a greater number of occurrences for the items relating to:
– ability to identify strengths and weaknesses;
– reflective and interpretative skills.

Tab. 1 – The point of view on the evaluation skills possessed by the school in which he works

Skills assessed – multiple choices	*NIV choices*	*Teachers of small schools choices*	*NIV members also teachers of small schools choices*	*Total occurrences of assessments for each skill*
Identify priorities and improvement paths	10	6	8	24
Reading and analysis of INVALSI test data, school questionnaires, and other sources	10	8	6	24
Finding links between weaknesses and action priorities	8	4	6	18
Point out strengths and weaknesses	8	9	6	23
Ability to motivate self-assigned levels for Outcomes and processes in the RAV	7	0	5	12
Reflective and interpretative capacity	3	5	4	12
Introduction of new school evaluation indicators	2	2	2	6
Construction of new self-assessment tools	1	1	2	4
Total occurrences expressed by the 3 groups of teachers	49	35	39	123

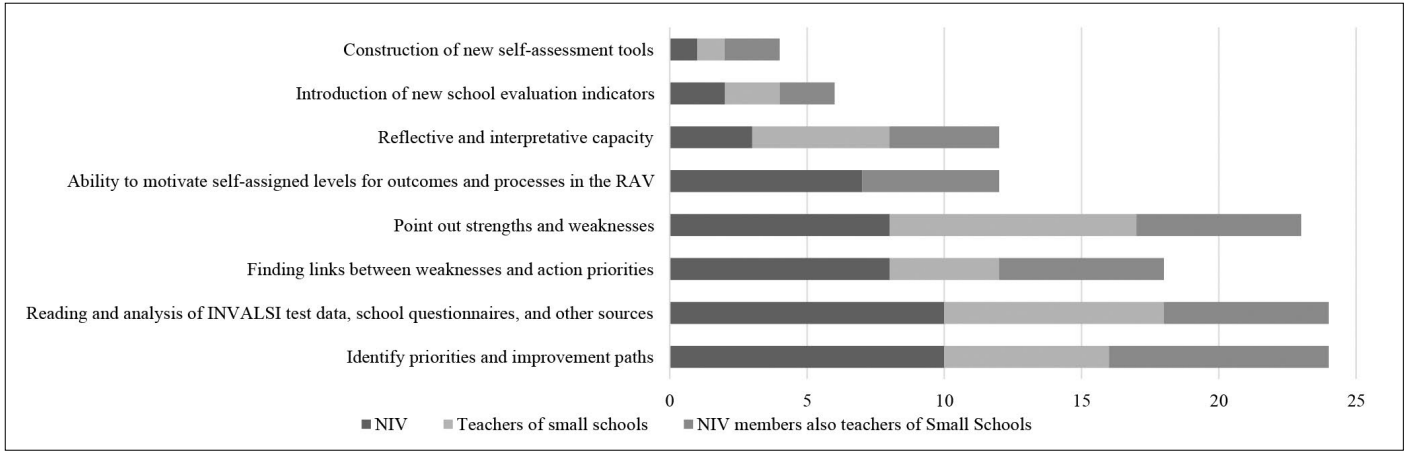


Fig. 1 – The point of view on the evaluation skills possessed by the school in which he works

Also for teachers of multi-class classes – compared to the composite audience (NIV members who also work in isolated complexes) – there is a greater number of occurrences for the category Reading and data analysis INVALSI tests, school questionnaire, other sources.

For all three categories of respondents, the need for more in-depth self-assessment training was highlighted, both with regards to the identification of school-specific indicators and the development of new self-analysis tools.

Regarding the question "To what extent do teachers of small schools contribute to the self-evaluation of the entire institution?" (Likert scale from 1 to 5 steps. 1 = little; 5 = very much), overall the audience of 27 teachers focuses on the following context: 11 teachers are positioned on the value 3; 6 teachers are placed on the value 5; the value 4 is chosen by 5 teachers; for the value 2 3 teachers opt and for the lowest value (1) 2 teachers. Proceeding to an examination in disaggregated form, NIV members are positioned more on the value 5; the contacts or teachers in multi-classes, small groups, mountain schools are placed on the value 3; those who are part of the NIV and are also teachers in multi-classes, small schools, mountain schools give emphasis to the value 4.

Calculating the scores for the three types of social actors separately, the following context emerges:

Tab. 2 – Calculation of the scores expressed by the members of the NIV (10 teachers) – Population 1

Rating	*Number of voters*
5	1
4	1
3	5
2	2
1	1
Total scores	29
Percentage of maximum value expressed for the role of teachers in isolated or multi-class complexes in the self-assessment	58%
Average: 2.9	2,9
Standard deviation: 3.96	3,96
Effect size: 0,53	0,53

Tab. 3 – Calculation of the scores expressed by the reference teachers or teachers in multi-classes, small complexes, mountain schools (9 teachers) – Population 2

Rating	*Number of voters*
5	2
4	3
3	3
2	1
1	0
Total scores	33
Percentage of maximum value expressed for the role of teachers in isolated or multi-class complexes in the self-assessment	73, 33% – rounded down: 73%
Average	3,66
Standard deviation	4,15
Effect size	0,61

Tab. 4 – Calculation of the scores expressed by the members of the NIV who are also referent teachers or teachers in multi-classes, small complexes, mountain schools (8 teachers) – Population 3

Rating	*Number of voters*
5	3
4	1
3	3
2	0
1	1
Total scores	29
Percentage of maximum value expressed for the role of teachers in isolated or multi-class complexes in the self-assessment	72,5%
Average	3,62
Standard deviation	4,69
Effect size	0,53

Interviewed by telephone interview, the 27 school principals expressed values of recognition of the contribution of teachers at small schools to self-evaluation, summarized in the following table.

Tab. 5 – Calculation of the scores expressed by school principals (27 principals) – Population 4

Rating	*Number of voters*
5	4
4	20
3	3
2	0
1	0
Total scores	109
Percentage of maximum value expressed for the role of teachers in isolated or multi-class complexes in the self-assessment	80, 74% – rounded up: 81%
Average	4,03
Standard deviation	5,63
Effect size	2,01

The analysis of the data brings out a paradox. Many NIV teachers declare that they have improved the self-assessment process, thanks to the comparison with colleagues who serve in multi-classes and isolated complexes. However, for NIV members, the average percentage of the highest recognition of the contribution to self-assessment by colleagues from isolated realities is lower than the average percentage of the value attributed to small schools in the self-assessment process by the other two audiences investigated, namely: a) the group of teachers in small schools who are also members of the NIV, within a school institution that enhances the functional versatility of professional resources; b) the set of teachers or referents in small schools/isolated complexes who, starting from initial training, have been oriented towards building competency balances, consequently improving the self-perception of the role.

In addition, a contradictory situation occurs where the various NIVs, for the purpose of self-assessment, consider constant and continuous communicative flexions between all the complexes important and emphasize the participatory reflection of the institute but do not favor or activate connection actions (also through digital tools) with the isolated complexes of the comprehensive institute, signaling a poor alliance between their group, the coordinating teachers of small schools and the teachers who are instrumental functions of support to the school curriculum and the training/research dimension.

School manager, while differing from multi-class teachers, in that they believe that small schools do not present excessive criticalities or in any case possess socio-cultural resources capable of limiting or counteracting teacher

turnover, are nevertheless in tune with these teachers on the following elements of development/improvement:

– need to generate mentoring and tutoring interactions between all complexes;
– the need to consider self-evaluation as a strategy to solve problems of each complex, particularly felt by teachers in multi-class;
– the need to link self-assessment to the development of proactive social reporting towards families, public bodies, associations, with a view to hybrid training networks that respond to the needs and vocations of the territories.

In addition, the managers emphasize how self-evaluation coincides with the training of all school staff on organizational-management issues. They also perceive how isolated complexes, multi-classes and micro-educational realities can anticipate the physiognomy of a new school, projected into a future that provides for co-teaching of teachers, destructuring of classes by age of learners, definition of longitudinal and transversal centers of interest (learning nuclei, composed of pupils of various ages and focused on hybrid disciplinary epistemologies, in relation to local cultures), differentiation of teaching and learning through school/territory integration and the correlation between formal and non-formal education.

Teachers/referents in isolated complexes – compared to NIV members – denote a greater perception of the self-evaluation of the institute as a factor of cohesion and a moment of peer tutoring among colleagues.

6. Emergencies from the questionnaire: training for the development of evaluation capacity: methods, strategies and teacher involvement

To the question "For which section of the RAV did you provide your contribution to the self-evaluation?" (open field response), overall the responses lead to the following codings, with occurrences from greatest to least: Reading and analysis of INVALSI test data; analysis of school questionnaire data and other sources; identification of strengths and weaknesses; choice of priorities and goals; identification of links between weak points and the improvement plan; inclusion of new evaluation indicators; motivation of self-assigned levels for outcomes and processes.

Analyzing the results by professional categories, differences emerge in the contribution of the social actors involved. In fact, it transpires that most of the members of the various NIVs participated in the drafting of all the sections of the Self-Assessment Report (RAV). This denotes a strong determination to be protagonists of the entire process but also the conviction of

being able to explore all the articulations of the system, without involving other teachers, including those who work in multi-classes.

Only a few NIV members focused exclusively on one of the following areas: Context – Outcomes – Processes.

In addition, they all contributed to defining the framework of the priorities, goals and process objectives.

The representatives of the multi-class/small schools worked more on the analysis – in order – of the Context data and the data relating to the Outcomes of the students, leaving the definition of the Priorities in the shade. Their activity suggests a priority attention to the concreteness of the socio-educational dimension and to the supervision of emergencies but also the lack of propensity to define themselves as real experts in school evaluation.

In fact, an attitude of delegation to the NIV prevails when it comes to establishing priorities, goals, process objectives, with a view to a systematic improvement plan.

The task of school principals is twofold: on the one hand, to strengthen the self-efficacy and confidence of these teachers in the specific evaluation skills possessed; on the other hand, to promote the development of collaborative professionalism among teachers.

Those who belong to both the NIV and the team of teachers of isolated complexes report that they have focused their gaze equally on Outcomes, priorities, process areas. In this case, a situation of balance appears between the ability to read and interpret data, the ability to ask questions to carry out comparative evaluations between classes, complexes, geographical areas on a micro and macro-territorial scale, the use of the various qualitative-quantitative elements (including the school effect) to establish decisions (improvement plan) and act both in everyday life and in the innovation of teaching-organizational practices.

The question, relating to the preferred and/or desired training (by the individual schools or the Pole Schools for the training of the 10 Ligurian Areas), stated: "Which one would you choose among the following training paths?", requiring you to indicate only one option, among the following:

– examine and evaluate how schools work;
– communicate and manage relationships in the evaluation process;
– organize self-evaluation at school.

The global answers to the question allow us to draw a scenario that is still predictable:

– organize self-evaluation at school (15 choices);
– communicate and manage relationships in the evaluation process (8 choices);
– examine and evaluate how schools work (4 choices).

If we analyze the disaggregated responses, focusing on the three categories of participants in the questionnaire.

We note a more articulated situation:

1) for NIVs, the same order of training priorities prevails as for the general audience:
   – satisfaction with self-evaluation training emphasizes members' sense of belonging;
   – of the NIV to a technical community that holds the exclusivity of the evaluation functions.

   In this context, an appreciable element is the intention to entrench the culture of evaluation, while a critical element is the self-isolation of the NIV and the poor sharing of processes with other teachers;
2) teachers in decentralized areas mostly want training on the course "Examining and evaluating how schools work", confirm and attest to the participatory dimension of their interventions and the propensity to collaborate with the entire school community and with bodies outside the school;
3) for the category "Both, i.e. members of the NIV and also teachers in isolated complexes", the choices are addressed exclusively to two training sectors, indicated in order of priority:
   – organize self-evaluation at school;
   – communicate and manage relationships in the evaluation process.

From these choices it is possible to deduce that the neglected educational axis (examination and evaluation of the school) is perceived as carried out at the level of self-training, in the dimension of daily practices. Compared to the various NIVs, the audience investigated denotes a strong interest in managing the communication system, disseminating information, disseminating results, socializing technical knowledge on evaluation (Evaluation Capacity Building – ECB).

In all likelihood, it will be possible for school principals to build a reflective and participatory leadership taking into account the training expectations of the various teachers, in order to prepare a development plan for professional resources that enhances common content but also specific requests of the various profiles.

To the question "Which methods are effective in self-evaluation?" – multiple choices possible among the following options:

– comparison between peers;
– video analysis;
– collective analysis of concrete cases in small businesses;
– work in networks with other schools and the local area;

– exchange of documentation;
– virtual learning environment.

The general audience speaks with the definition of the following occurrences:
– comparison between peers – 22;
– collective analysis of concrete cases in small businesses – 22;
– work in networks with other schools and the territory – 19;
– exchange of documentation – 19;
– virtual learning environment – 3;
– video analysis – 2.

By directing the analysis to the three categories of respondents, we note that the NIVs attribute importance, in order, to: comparison between peers; collective analysis of concrete cases in small businesses; exchange of documentation; networking with other schools and the local area.

The teachers in isolated groups enhance, in order: comparison between peers; exchange of documentation; collective analysis of concrete cases in small settings; working in networks with other schools and the territory.

For the category "Both i.e. NIV and multi-class referents" the priorities are: collective analysis of concrete cases in small realities; working in networks with other schools and the territory; comparison between peers; exchange of documentation.

In short for everyone there is a strong lack of attention to the use of new technologies and the preparation of innovative environments in self-assessment.

It is interesting to analyze the answers to the open-ended question "What strategic skills should someone who evaluates their school possess?".

The answers – unambiguously for the three audiences investigated – lead to the following codings, with occurrences from highest to lowest number:
– analytical skills;
– organizational skills;
– communication skills;
– management skills of environments (including virtual) suitable to encourage interaction between peers,
– between teachers and families;
– planning and vision skills;
– mediation and relationship skills;
– autonomy and flexibility;
– critical thinking/decision making/problem solving;
– diagnosis and self-organization skills;
– awareness of the co-evolution between school and territories.

The profile of the teacher-evaluator, as defined by teachers, still appears to lack certain fundamental characteristics: critical thinking, the ability to di-

agnose problems and make decisions, organizational flexibility, and a sense of the networked perspective of the school context.

The scenario outlined by the exploratory survey indicates that the process is underway to enable teachers to achieve the competency standard "Possession and exercise of orientation and research, documentation, and evaluation skills" (Ministry of Education Decree No. 226/2022 – "Provisions concerning the training and probationary process for newly hired teaching and educational staff").

Evaluation capacity building can be encouraged and guaranteed by the systemic (pedagogical, organizational, and managerial) activity of school principals who, together with all social actors in the educational context, develop a reflective, concrete leadership style capable of embracing change and transformation.

7. Lessons learned

The exploratory investigation led managers and teachers to intuit and consequently build forms of horizontal collaboration to solve the most significant system problems. The same path guided managers to define models of distributed leadership (Paletta, Greco and Santolaya, 2022) and transformative research, generating widespread communication strategies between teachers and top management, between staff and complexes; promoting the sharing of knowledge and the dissemination of practices; encouraging and intensifying initiatives for the professional development of teachers, with particular emphasis also on self-training and co-planning in working groups. In addition, in order to ensure consistency and cohesion in the self-assessment action, the managers have prepared tools to oversee and monitor the lines of innovation (not only digital) of the school.

Priority was given to the construction and management of:

- quality team;
- mentoring groups;
- communication departments;
- junction principals for the management of multi-classes and the connection between the different school complexes (Chipa, Mangione and Cannella, 2022).

In search of a balance between leadership and management, the executives implemented the following lines of activity:

- creation of communication support for teachers in dealing with challenges, risks and problems;

– assumption of a leading role also for the territory, especially with regard to unprecedented forms of inclusion and social cohesion;
– reshaping the strategic vision of development;
– attention to the standards of teaching professionalism defined in the EU context;
– use of new forms of enhancement of the staff of the entire institution;
– intuition of the transformative research potential present in the institution.

The research into the role of small schools in improving evaluation processes has resulted in some important lessons both for the reorganization of the leadership of institutions with small schools (Greco and Paletta, 2022) and for the deepening of the relationship between school and territory.

New models of school organization have made it possible to:
– expanding the number and varying the type of staff figures (implemented with multi-class teachers);
– create support tools for sustainable educational interventions in individual complexes and on the network;
– generate exchanges of information and knowledge between teachers and experts;
– promote the use of new technologies not only at the communicative level but also in the construction of knowledge common to all professional figures.

In particular, the following have been prepared and activated:
– rotation paths of figures in the staff, with openings to the referents/coordinators of the multi-classes;
– modular paths of NIV reorientation/modification/implementation;
– mutual listening desks between teachers in sunny complexes and teachers in standard classes;
– models of authentic training, with the aim of confronting teachers with problems and inducing them to find possible solutions;
– "mirroring" and peer tutoring workshops between multi-class teachers and standard class teachers, through exchanges of professionalism and visiting opportunities.

There are also important empowerment levers of valuation capital, identified by the managers among the following elements:
– adoption and development of forms of "teaching contract" between teachers and experts in the area;
– actions to implement data analysis capabilities to be used in all the strategic documents of the cchool (RAV/Three-year plan of the educational offer/social reporting);
– use of spaces outside the school through a pedagogical codification of the same;

– elaboration of network actions between teachers of nearby schools to compare critical issues and lines of development of the self-evaluation of individual schools;
– development of monothematic/project-based connections between teachers working in geographically distant rural schools;
– interpretative exercises of the Student outcomes and the School effect, comparing standard and multi-class classes on the basis of indicators/descriptors of the self-assessment rubric of the National Evaluation System/ INVALSI;
– "Technical colleges" to support schools in the implementation of improvement plans and in solving system problems (objective-focus: to go beyond the bureaucratic compliance of some members of the NIV and the empirical dimension of some representatives of small schools);
– sharing of professional resources and strategic ways of reorganizing the school to deal with dropout, implicit dispersion, educational poverty;
– transfer of management methods of the time/space of teaching implemented in small complexes to urban schools, with attention to the following dimensions: use of multi-class as a form of peer learning; identification of meta-classes that arise from the relationship between isolated complexes and coastal classes; personalization of educational interventions; organizational contextualization; reshaping of the school curriculum by centers of interest.

The role of isolated and decentralized complexes in the rooting of system evaluation cultures, as revealed by the survey, has oriented the managers to create an organization of the school that – in parallel and similarly to the structure of the territory – appears reticular, multicentric and with topological geometry.

In a context of continuous transformations, homogeneous networks (between comprehensive schools) have federated with local bodies and associations, generating hybrid constellations, focused on the following dimensions, which are not very evident and scarcely interceptable in standard educational institutions:
– definition of the professional portfolio of teachers, differentiated into two subsets (class teachers and teachers who are experts in the organization and evaluation of the school's strategic processes);
– review of models and procedures in the construction of improvement plans and training plans;
– use of problem posing as well as problem solving in daily practice;
– exercise of critical thinking within collegial bodies;
– development of digital cultures that allow and make sustainable both social creativity and pedagogical networks;

– elaboration of a new culture of relationship with families;
– identification between the school's vision and organizational models.

Important management innovations are noteworthy, including:

– use of inter-professionalism (teachers within the school who acted as educators for the territory; external experts who collaborated in the implementation of the PNRR measures, with a focus on Mission 4 – Component 1 – "Strengthening the Offering of Education Services: From Nursery Schools to Universities");
– construction of agreements with Public Administrations to increase and entrench processes of change in the organizational/evaluation system, allowing answers to be found to the following question: how can the characteristics of an extended and proximity school (multi-classes, isolated complexes) represent a solution to the critical issues of pupils' learning and teacher training, integrating skills and well-being status of all social actors?
– development of multiple spaces for education and continuing education, in synergy with volunteer workshops, civil service, international exchanges.

Taking the principle of small schools as the engine for the development of social capital in schools, some teaching and organizational practices were examined and evaluated, involving a differentiated audience of social actors (teachers, parents but also local administrators) and promoting opportunities for authentic social accountability.

In particular, some working indications developed by school leaders are highlighted:

– produce and maintain levels of self-efficacy by institutions with small complexes;
– evaluate the duplicity of a teacher who is at the same time "teacher and context leader", in synergy with the school principal;
– identify multigrade classes as the innovative focus of the school;
– experimenting with dossiers of professional skills, contextualized and based on place-based pedagogical constructs;
– develop the traits of an "embodied educational leadership": awareness, responsibility, care (Gomez Paloma, Di Tore and Mangione, 2025);
– nurture strategies and practices to reduce or counteract constraints and critical issues (territorial isolation, digital divide, scarcity of students and social exchanges);
– enhance forms of widespread and informal self-evaluation (through digital narratives of processes and approaches);
– orient the entire community to provide policy makers with specific indications for the sustainable development of education and training, consid-

ering the urgency of governing interdependent systems, subject to fluctuations and changes as well as constant risks;
– prevent networks and collaboration agreements with the territory from translating into mere obligations (Poliandri *et al.*, 2023).

A further indication of policy can come from the particular type of school structure that is declared by the managers. It is necessary to use the classification of schools involved in self-evaluation processes since the birth of the National Evaluation System (Presidential Decree 80 of 2013), which leads educational institutions back to three priority categories (perceived by the protagonists of self-evaluation themselves):
– open structure with organizational leadership;
– unstable learning organization;
– uncoordinated structure with widespread leadership (Quadrelli *et al.*, 2014).

The 27 managers involved in the exploratory research – unanimously – affirm that the educational institution they govern is in the transition phase from dimension c) to dimension a). The effort to be made is to integrate the strengths already possessed (valuable didactic planning, presence of interdisciplinary working groups, care of relationships and school climate) with still residual lines of development (distributed governance, openness to the territory, self-assessment activities).

The transformations triggered by participatory research offer a glimpse of the development – with varying phases in different regions – of the alliance between collaborative learning and sustainable school evaluation (Salvadori, 2024).

References

Chipa S., Mangione G.R.J., Cannella G. (2022), "Leadership models and alliances in 'small European schools'. An interpretative analysis on an international basis", *IUL Research*, 3, 5, pp. 63-86, retrieved on January 20, 2026, from https://doi.org/10.57568/iulres.v3i5.353.

Giampietro L., Poliandri D., Romiti S. (2020), "Networking among schools in Italy: innovations, constraints and real opportunities", *Local Autonomies and Social Services*, 43 (2) pp. 381-400.

Gomez Paloma F., Poliandri D., Giampietro L. (2020), "The Value for Schools project: Pedagogical Research and Learning Analytics for the self-evaluation of schools", *Training & Teaching*, 18 (1) pp. 294-307.

Gomez Paloma F., Di Tore P.A., Mangione G.R.J. (2025), *Teacher Training and Student Learning – Past Values, Present Uncertainties and Future Prospects*, In-

techOpen, London, retrieved on January 7, 2025, from https://www.intechopen.com/books/1004355.
Mangione G.R.J., Cannella G., Parigi L., Bartolini R. (2020), *Communities of memory, communities of the future. The value of the small school*, Carocci, Roma.
Mangione G.R.J. (2022), *Online classes. Analyze opportunities for change in small schools*, Think MultiMedia Editore, Lecce.
Paletta A., Greco S., Santolaya E.M. (eds.) (2022), "Leadership, innovation and organizational change. Foster professional learning communities. Leadership, innovation and organizational change. Promoting professional learning communities", *IUL Research*, 3, 5, pp. 1-5, retrieved on January 20, 2026, from https://doi.org/10.57568/iulres.v3i5.361.
Poliandri D., Nerli Ballati E., Litteri A., Giampietro L. (2023), "Network collaboration between schools. A case study in support of self-evaluation processes", *Democratic School. Learning for Democracy*, 3, pp. 543-566.
Quadrelli I., Poliandri D., Perazzolo M., Giampietro L, Muzzioli P. (2014), *Self-evaluation in the schools of the Vales Project*, speech at the seminar Towards the National Evaluation System: indications from the VALES project, Rome 25 June 2014, retrieved on January 20, 2026, from https://www.invalsi.it/invalsi/ri/vales/doc.php website.
Salvadori I. (2024), "Supporting school evaluation processes through collaborative professionalism", *RIV – Italian Evaluation Review*, XXVIII, 88-89, pp. 55-74.

5. What makes a good reader? Evidence from PIRLS 2021

by Valeria F. Tortora, Paola Giangiacomo, Michele Cardone

PIRLS 2021 is an international survey that assesses the reading skills of students in the fourth year of schooling in more than 50 countries. The survey is repeated every five years and has provided comparable data on reading skills since 2001. The framework defining reading skills is divided into two purposes: literary experience and acquisition and use of information. The survey used paper and digital tests until the 2016 edition, since 2021 there has been a transition to digital tests that aim to measure the ability to read and comprehend texts online. The survey also collects information on the family, school and educational context of reading to identify factors that influence learning. The survey is coordinated at international level by Boston College and at national level by INVALSI since the first cycle. The last survey took place in 2021, despite considerable difficulties caused by the Covid-19 pandemic, and involved approximately 300,000 students worldwide. The results of the survey were published in 2023 and show differences in performance between countries, trends over time and relationships between reading and context. The survey thus provides an overview of the reading skills and abilities of students in grade 4 and the challenges and opportunities for improving education and training. The results of this IEA international survey were chosen because reading competence is cross-cutting and fundamental to the acquisition of other school skills and the improvement of learning levels. Therefore, the aim of this article is to analyse the factors influencing the performance of readers, both at individual and contextual level with a focus on Italy. The paper presents the main results of the IEA PIRLS 2021 survey for Italy and compares them with those of the other participating countries, highlighting the strengths and areas for improvement of the Italian educational system with particular reference to the reading skills of students in the fourth year of schooling. The characteristics of students, families, schools and ed-

ucational systems that are positively or negatively associated with reading achievement are examined. Good practices and critical issues that emerged from the survey are highlighted, with suggestions for improving the quality of reading teaching and learning. Implications of the results for educational policies and future research are discussed.

PIRLS 2021 è un'indagine internazionale che valuta le competenze in lettura degli studenti del quarto anno di scuola in più di 50 Paesi. L'indagine viene ripetuta ogni cinque anni e dal 2001 fornisce dati comparabili sulle competenze in lettura. Il quadro che definisce le competenze di lettura è suddiviso in due scopi: l'esperienza letteraria e l'acquisizione e l'uso di informazioni. L'indagine ha utilizzato test cartacei e digitali fino all'edizione del 2016, mentre dal 2021 si è passati a test digitali che mirano a misurare la capacità di leggere e comprendere testi online. L'indagine raccoglie anche informazioni sul contesto familiare, scolastico ed educativo della lettura per identificare i fattori che influenzano l'apprendimento. L'indagine è coordinata a livello internazionale dal Boston College e a livello nazionale dall'INVALSI, fin dal primo ciclo. L'ultima indagine si è svolta nel 2021, nonostante le notevoli difficoltà causate dalla pandemia Covid-19, la rilevazione ha coinvolto circa 300.000 studenti in tutto il mondo. I risultati dell'indagine sono stati pubblicati nel 2023 e mostrano le differenze di rendimento tra i Paesi, le tendenze nel tempo e le relazioni tra lettura e contesto. L'indagine fornisce quindi una panoramica delle competenze e delle abilità di lettura degli studenti di quarta classe e delle sfide e opportunità per migliorare l'istruzione e la formazione. I risultati di questa indagine internazionale dell'IEA sono stati scelti perché la competenza in lettura è trasversale e fondamentale per l'acquisizione di altre competenze scolastiche e per il miglioramento dei livelli di apprendimento. Pertanto, l'obiettivo di questo articolo è analizzare i fattori che influenzano le prestazioni degli studenti, sia a livello individuale che contestuale, con un focus sull'Italia. L'articolo presenta i principali risultati dell'indagine IEA PIRLS 2021 per l'Italia, evidenziando i punti di forza e le aree di miglioramento del sistema educativo italiano con particolare riferimento alle competenze di lettura degli studenti del quarto anno di scolarità. Vengono evidenziate le buone pratiche e le criticità emerse dall'indagine, con suggerimenti per migliorare la qualità dell'insegnamento e dell'apprendimento della lettura dell'insegnamento e dell'apprendimento della lettura. Vengono discusse le implicazioni dei risultati per le politiche educative e la ricerca futura.

1. Introduction

Reading is a fundamental competence for the cognitive, social and emotional development of boys and girls. It enables them to access knowledge, understand the world, express their identity and participate in democratic life. However, reading is not a natural skill, but the result of a complex and articulated learning process, which requires the intervention of various individual, family, school and social factors (Mullis *et al.*, 2023).

Contextual factors have been identified as variables that significantly influence students' reading performance. However, the weight of several factors simultaneously on students' reading performance is still clearly defined and deserves to be explored further. On the basis of Walberg's educational productivity theory (Walberg, 1984) and Bronfenbrenner's ecological system theory (Bronfenbrenner, 1979), it was decided to analyse the PIRLS data in order to try to understand which factors contribute most to high levels of reading proficiency.

Bronfenbrenner's model emphasises how a child's environment influences the way he or she grows and develops (Bronfenbrenner, 1979). Proficiency in reading is a basic skill useful for acquiring knowledge and achieving important learning goals, which are essential for students' high achievement (Graham *et al.*, 2017).

Taking a cue from Walberg's educational productivity model, it is considered important to emphasise the influence of contextual factors on the complex learning process; Walberg considers nine factors categorised into learner attitude, education and environmental factors (Walberg, 1984).

Among the variables that most influence reading achievement, gender certainly shows a great variance in results: females achieve higher levels of reading competence than males (INVALSI, 2023; Mullis, 2022).

In the same way, the geographical divide shows how crucial it can be for academic success in reading (INVALSI, 2023; Mullis, 2022).

Another variable determinant is the family socio-economic background: a recent study investigated the role of mediating factors of parental education (e.g. informal reading activities at home and the child's early reading ability) on reading ability, concluding that parental education has a substantial effect on it (Chen, 2021).

2. Conceptual and methodological framework of the PIRLS survey

The aim of the PIRLS survey is to provide internationally comparable data on the performance and contextual factors influencing the learning and teaching of reading in fourth grade students. Furthermore, the survey assesses four comprehension processes within each of the two purposes: a) focusing and retrieving explicitly stated information; b) making simple inferences; c) interpreting and integrating ideas and information; d) evaluating and critiquing content and textual elements. The survey uses proven tools and methodologies validated through extensive cooperation of international research organisations, under the supervision of the International PIRLS Centre and TIMSS of Boston College. The survey involves more than 60 educational systems (countries and benchmarking entities) worldwide.

The survey uses both paper and digital tests, the latter to measure the ability to read and understand online texts. The survey also collects information on the family, school and educational context of reading to identify factors that influence learning.

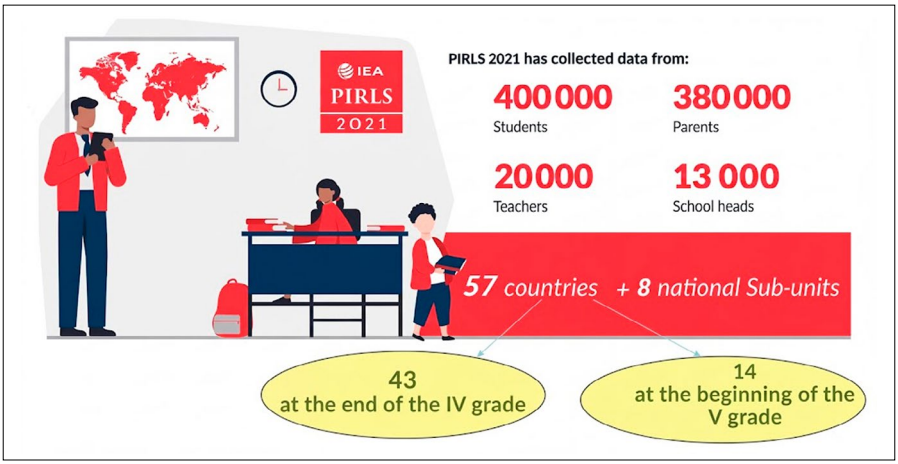


Fig. 1 – International participation in PIRLS 2021

Source: IEA PIRLS, 2023

The PIRLS 2021 survey framework defines reading as a constructive and interactive process, which depends on the context of the reading situation, the different text types and the social context. The framework identifies two main purposes of reading: for literary experience and for acquiring and using information. The first purpose concerns the reading of narrative texts, which stimu-

late the readers' imagination, emotionality and creativity. The second purpose concerns the reading of informative texts, which provides data, facts and arguments on various topics. Furthermore, the PIRLS survey, in addition to measuring pupils' achievement levels, allows us to collect and analyse, via context questionnaires, a great deal of information about the family environment and numerous factors concerning the student, which are closely correlated, in literature, with learning to read. Children, in fact, learn to read in different contexts and through a wide variety of activities and experiences, and each environment (home, school and national context) can support the other, creating a connection between home and school that is fundamental to learning.

3. Purpose and hypothesis of the research

Reading is a key competence for the cognitive, social and emotional development of children, preparing them to face the challenges of the contemporary world. However, reading is not an innate skill, but the result of a long and complex learning process that requires the support of various individual, family, school and social factors.

The research hypothesis is that the survey results offer food for thought and action for the improvement and development of Italian students' language skills, taking into account the challenges and opportunities offered by the family, school and social context. It is also hypothesised that the digital assessment mode of the survey has introduced novelties and difficulties in the assessment of online reading, which requires specific skills different from those required for paper-based reading. Finally, it is hypothesised that the Covid-19 pandemic has had negative effects on students' performance, both in terms of the interruptions and changes in teaching activities and the psychological and emotional consequences it has caused.

4. Object and research hypothesis

The aim is to provide internationally comparable data on the reading literacy levels of boys and girls after four years of primary school and to survey the variables associated with reading achievement. This article will present the main results of the PIRLS 2021 survey, with particular attention to the positioning of Italy compared to other European countries. We will analyse the differences between geographical areas and the variables that most influence students' reading ability. Finally, the educational and didactic impli-

cations of the survey results will be discussed, proposing possible actions to improve and develop students' linguistic skills. We will compare these data with the previous 2016 survey to analyse the impact that the pandemic may have had on the results achieved.

This article describes and analyses information relating to the family context and the characteristics and attitudes of students towards reading; this was done both by comparing the results relating to the indices and the average return associated with them in a comparative perspective with respect to the other countries participating in the survey, and by investigating in more detail the situation of our country, at the level of the various geographical areas, analysing the indices and variables contained within the context questionnaires.

5. Method

The method used for the article is based on the analysis of data from the IEA PIRLS 2021 survey, which assessed the reading skills of students in the fourth year of schooling in more than 50 countries. The survey took place between 2020 and 2021, with data collection facing the challenges posed by the Covid-19 pandemic. To ensure the safety of the students, teachers and staff involved, prevention and protection measures adapted to national regulations and local conditions were taken. In some countries, including Italy, the possibility was offered to administer the survey digitally (ePIRLS), with interactive materials and online project simulations.

The Italian sample consisted of 7,419 grade 4 students from 222 primary schools distributed throughout the country (in 2021, there was a transition to CBT, so we had two samples: 5,440 took the CBT tests and 1,979 students took the paper-based tests). The participation rate of the schools was 93%, while that of the students was 95%. 51% of the students took the survey in digital mode, while 49% took the survey in paper mode. The digital mode made it possible to include the ePIRLS tests, which assess the ability to read and understand online texts, using a simulated Internet interface. The digital assessment mode also made it possible to implement an adaptive group design, which assigns tests to students based on their estimated ability level. The paper-based assessment mode ensured continuity with previous cycles and comparability of results.

In addition to the reading test, the students completed a questionnaire on their reading habits and preferences, their school and family experiences and their opinions on reading. The students' parents also completed a questionnaire on the socio-economic characteristics of the family, their reading activities with their children and the resources available at home for reading. The

teachers of the sample classes completed a questionnaire on their professional characteristics, teaching practices, classroom climate and reading curriculum. Finally, the school leaders completed a questionnaire on their school characteristics, available resources, school climate and health emergency management.

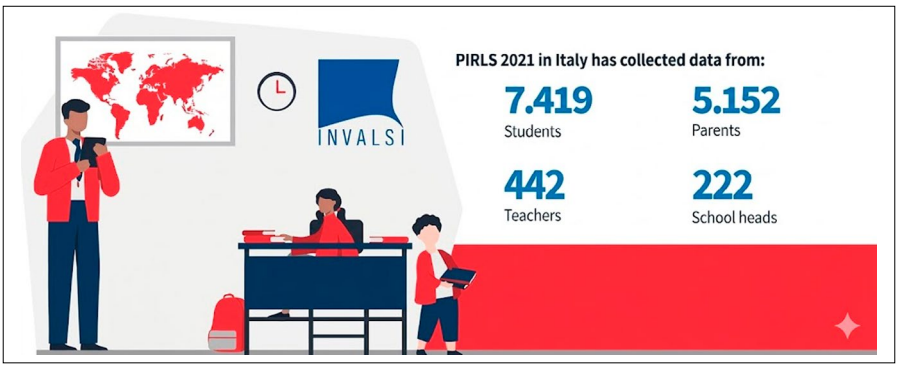


Fig. 2 – National participation in PIRLS 2021

Source: INVALSI IEA PIRLS, 2023

The data collected were processed and analysed using appropriate statistical techniques, taking into account the complexity of the sample design and the hierarchical nature of the data. The data were weighted to correct for any bias due to non-response or sample selection. The data were also standardised to obtain a common scoring scale for all countries, with an international mean of 500 and a standard deviation of 100. Finally, the data were disaggregated by variables of interest, such as gender, geographical area, school type, socio-economic and cultural background, reading practices, school climate and reading curriculum.

The data analysis aimed to answer the following research questions:

– What is the reading competence level of Italian students in their fourth year of schooling and how does it compare with other European and non-European countries?
– What are the differences in performance between Italian geographical areas and what are the variables that explain them?
– What are the contextual factors that influence Italian students' learning to read and how do they relate to performance?
– What is the impact of the Covid-19 pandemic on the survey results and what are the differences compared to the 2016 survey?
– What are the educational and teaching implications of the survey results and what are the possible actions to improve and develop the language skills of Italian students?

The IEA PIRLS 2021 survey took place between 2020 and 2021, with data collection facing the challenges posed by the Covid-19 pandemic. To ensure the safety of the students, teachers and staff involved, prevention and protection measures have been adopted that are appropriate to national regulations and local conditions. In some countries, including Italy, the possibility has been offered to administer the survey in digital mode (ePIRLS), with interactive materials and online project simulations.

The Italian sample is made up of 5440 grade 4 students from 164 primary schools distributed throughout the country. The participation rate of the schools was 93%, while that of the students was 95%. 51% of students took the survey digitally, while 49% took the survey on paper.

In addition to the reading test, students completed a questionnaire about their reading habits and preferences, their school and family experiences, and their opinions about reading. The parents of the students also filled out a questionnaire on the socio-economic characteristics of the family, on reading activities with their children and on the resources available at home for reading. The teachers of the sample classes filled in a questionnaire on their professional characteristics, teaching practices, class climate and reading curriculum. Finally, the school principals filled out a questionnaire on the characteristics of the school, the resources available, the school climate and the management of the health emergency.

6. Results

The 2021 edition of the IEA PIRLS survey was particularly challenging due to the Covid-19 pandemic, which imposed restrictions and changes in the way tests and questionnaires were administered. Despite the difficulties, most participating countries completed the survey, ensuring the quality and comparability of the data collected. Italy participated in the survey with a sample of approximately 5,440 fourth grade students from 164 primary schools distributed throughout the country.

The results of the IEA PIRLS 2021 survey were made public on 10 March 2024 with the release of the international and national reports. The results show the performance of Italian students in terms of average score, position in international and European rankings, and distribution of proficiency levels. In addition, the results highlight differences in performance between students in relation to contextual variables, such as gender, socio-economic background, school background and family background. Finally, the results make it possible to compare the performance of Italian students with that of

students in other participating countries, and to analyse trends of change with respect to previous editions of the survey.

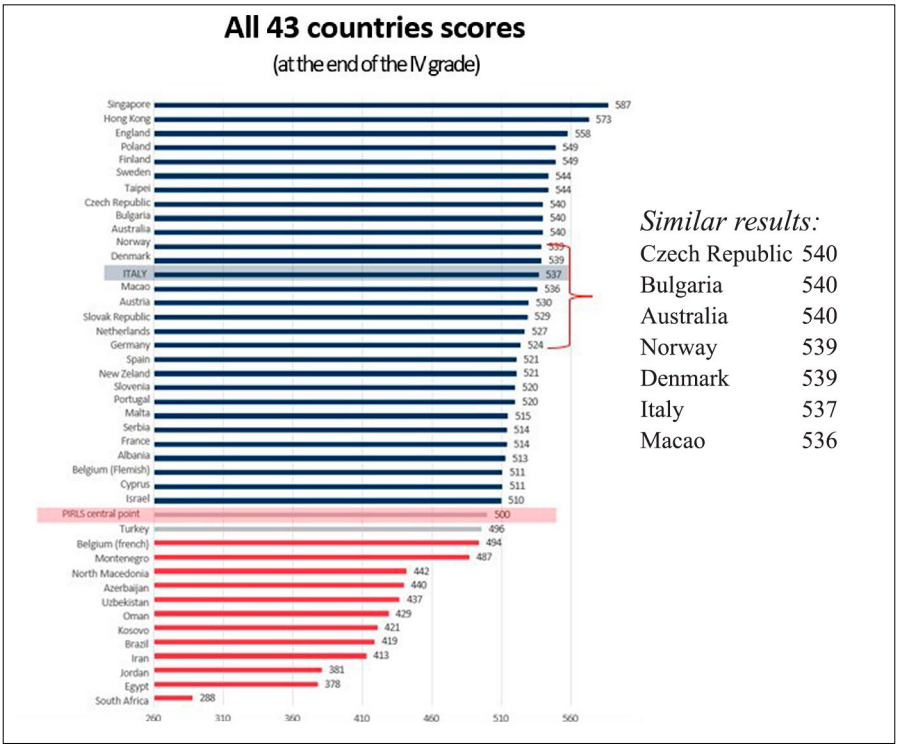


Fig. 3 – International average Reading achievement

Source: IEA PIRLS, 2023

Italian results in general are good, these are our "similar" countries (in terms of average results). The results indicate that, on the whole, these 43 countries achieved relatively high results in reading in the fourth grade: students in almost three quarters of the countries (30 out of 43) scored above the midpoint of the scale, i.e. 500 – a stable point across assessment cycles.

The national average result can also be appreciated if we refer only to the EU participants: Italy is 10 points above the average. Another good news is that our results are more concentrated around the average (so less excellence but also less poor results).

In this section, we will describe the main results of the IEA PIRLS 2021 survey for Italian students, delving into some relevant aspects. The text will be organised in four sections: the first section will present the aver-

age score and ranking of Italian students; the second section will examine the differences in performance among Italian students by reading purpose and process; the third section will analyse the differences in performance among Italian students by gender; the fourth section will discuss the differences in performance among Italian students by socio-economic, school and family context. The text will conclude with a brief summary and some final considerations.

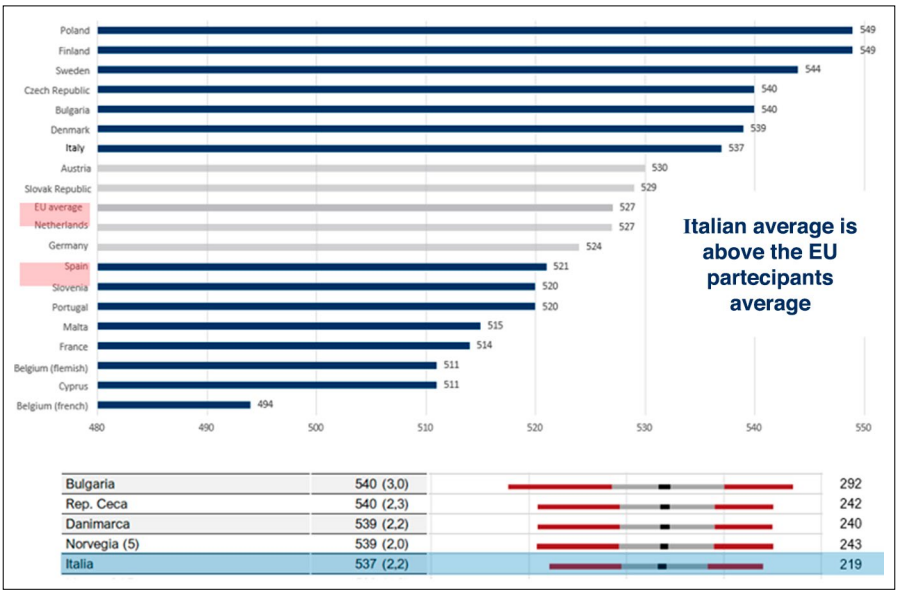


Fig. 4 – Average Reading achievement in European participants

Source: IEA PIRLS, 2023

The average score of Italian students in the IEA PIRLS 2021 survey is 525, down 11 points from 2016. This result places Italy 23rd out of 57 participating countries, with a statistically significant difference to the international average (500). Among the European Union countries, Italy ranks 14th out of 24, with a non-significant difference from the European average (524). In Italy, as in most of the participating countries, students obtain a significantly lower average result in 2021, in particular, by 11 points compared to that recorded 5 years earlier, bringing the results of Italian students back in line with those of 20 years ago (PIRLS 2001) and 10 years ago (PIRLS 2011).

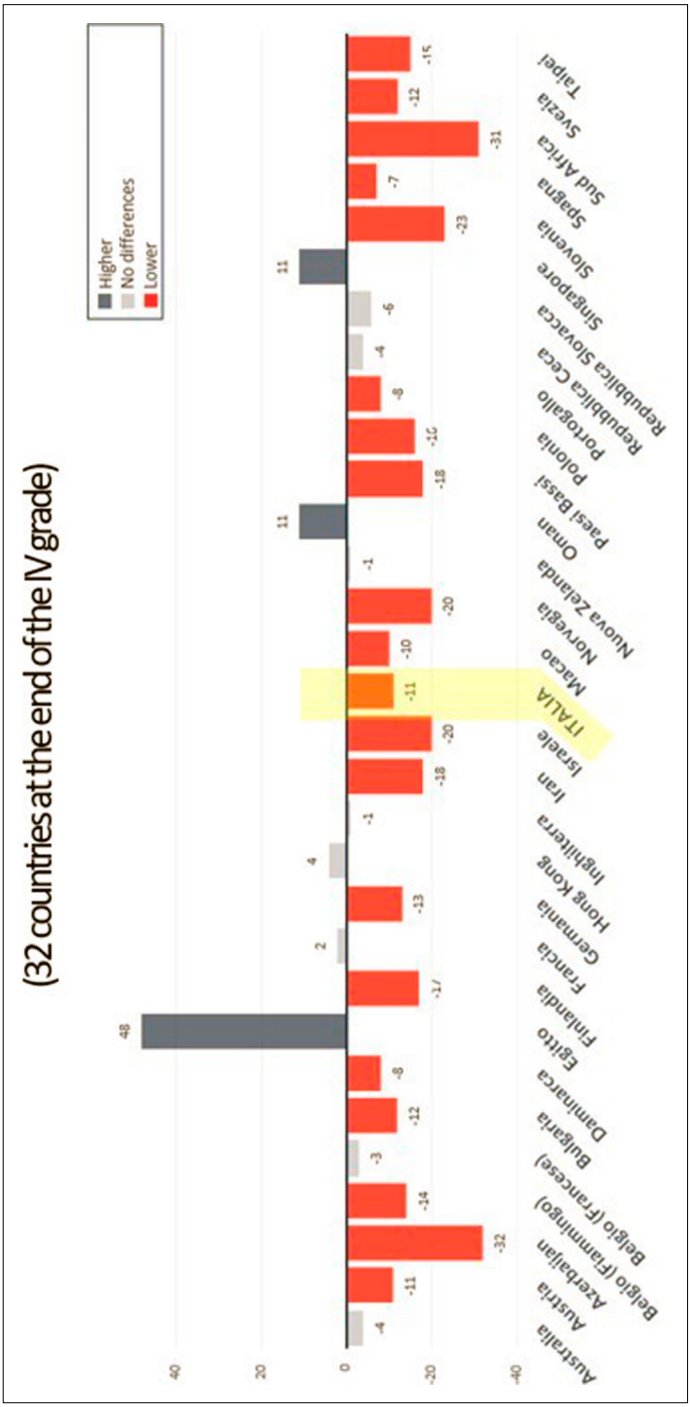


Fig. 5 –Trend in Reading achievement PIRLS 2021 versus 2016

Source: IEA PIRLS, 2023

The average score of Italian students is lower than that of many European countries, such as Ireland (582), Finland (578), Poland (577), Sweden (570), Norway (569), Denmark (565), Lithuania (562), the Netherlands (559), England (558), Portugal (555), Slovenia (552), Hungary (551) and the Czech Republic (549). The average score of Italian students is higher than that of some European countries, such as Spain (522), France (520), Belgium (519), Malta (518), Germany (517), Austria (516), Slovakia (513), Romania (510), Bulgaria (505) and Cyprus (499). The average score of Italian students is similar to that of two European countries, namely Croatia (526) and Greece (525).

The Fig. 6 compares the results in Italy between PIRLS 2016 and PIRLS 2021, showing the differences by geographical areas.

Overall, the results indicate a general decline in reading performance across the area. The national average dropped by 11 points compared to 2016; the North-East experienced the largest decrease, with a loss of 15 points and the North-West also showed a strong decline of 12 points. The Center recorded a smaller decrease (-5 points), suggesting relatively more stability in this area; the South and the South/Islands registered declines of 11 and 12 points, respectively, in line with the national trend.

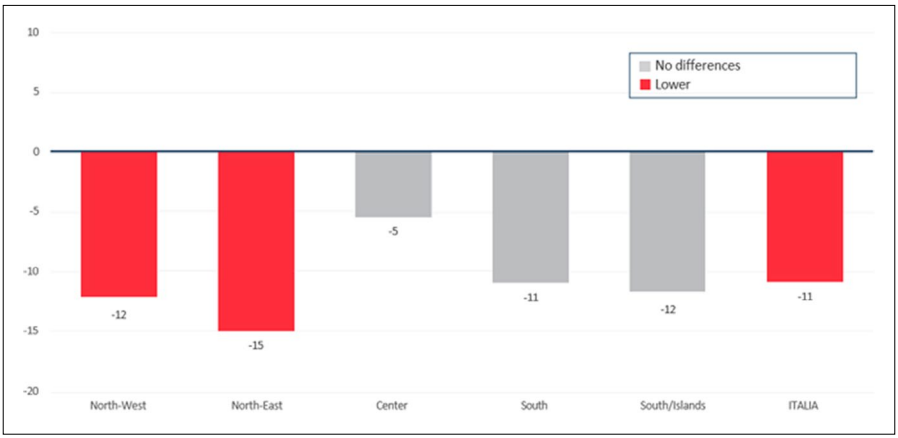


Fig. 6 –Trend in Reading achievement for Italy and Italian area PIRLS 2021 versus 2016

Source: INVALSI IEA PIRLS, 2023

The data suggest that the decline in reading performance affected all regions, though with varying intensity, and the North-East appears to be the most negatively impacted. The distribution of Italian students' proficiency

levels shows in Fig. 7 that 12% of students reached the advanced level, 35% the high level, 36% the intermediate level, 14% the low level and 3% the very low level. This distribution is similar to that of the international average, but different from that of the European average, which has a higher percentage of students in the advanced levels, and a lower percentage of students in the low and very low levels.

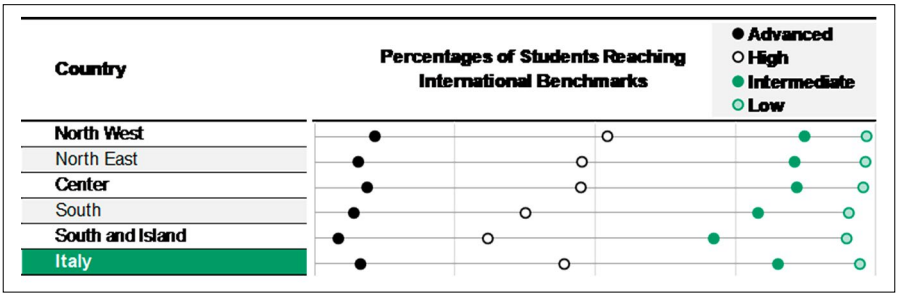


Fig. 7 – Percentages of students reaching benchmarks for Italian area

Source: INVALSI IEA PIRLS, 2023

	Average (s.e.)	*North West*	*North East*	*Centre*	*South*	*South Islands*
2021						
North-West	550 (3,3)				↑	↑
North-East	542 (3,5)				↑	↑
Centre	543 (2,9)				↑	↑
South	527 (6,8)	↓	↓	↓		
South/Islands	513 (7,1)	↓	↓	↓		
2016						
North-West	562 (2,5)				↑	↑
North-East	557 (3,5)				↑	↑
Centre	549 (4,9)				↑	↑
South	538 (5,0)	↓	↓	↓		↑
South/Islands	525 (6,6)	↓	↓	↓	↓	

Fig. 8 – Differences in average reading scores between areas for Italian area

Source: INVALSI IEA PIRLS, 2023

The data in the Fig. 8 highlight consistent and significant differences in average scores between geographic areas in Italy. Over both 2016 and 2021, Northern regions (North-West, North-East, and Centre) clearly outperform the Southern regions (South and South/Islands).

2021 Trends:

– Northern regions lead:
  • North-West: 550 ↑
  • North-East: 542 ↑
  • Centre: 543 ↑
– Southern regions lag:
  • South: 527 ↓
  • South/Islands: 513 ↓

The upward (↑) and downward (↓) arrows emphasize a North-South performance divide.

2016 Trends:

– The performance gap was already evident:
  • North-West: 562 ↑
  • North-East: 557 ↑
  • Centre: 549 ↑
  • South: 538 ↓
  • South/Islands: 525 ↓

Although overall scores were higher in 2016, the regional hierarchy remains unchanged.

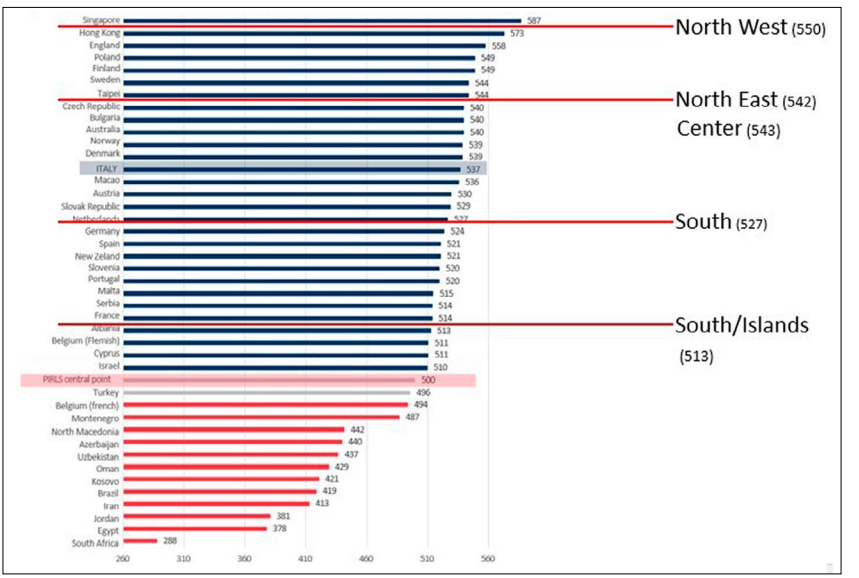


Fig. 9 – Differences in average reading scores between areas for Italian area and all countries

Source: IEA PIRLS, 2023

The chart in Fig. 9 presents the average scores of 43 countries, including data from different geographical areas within Italy, offering a broader perspective on Italy's educational performance globally.

Top performers:

– Singapore, Hong Kong, and England lead the ranking with average scores above 550;
– these countries represent the highest-performing education systems in the comparison.

Italy's position:

– Italy (national average) is placed in the middle of the international ranking;
– its average is comparable to countries like Austria and Norway, suggesting a moderate overall performance.

Regional breakdown within Italy:

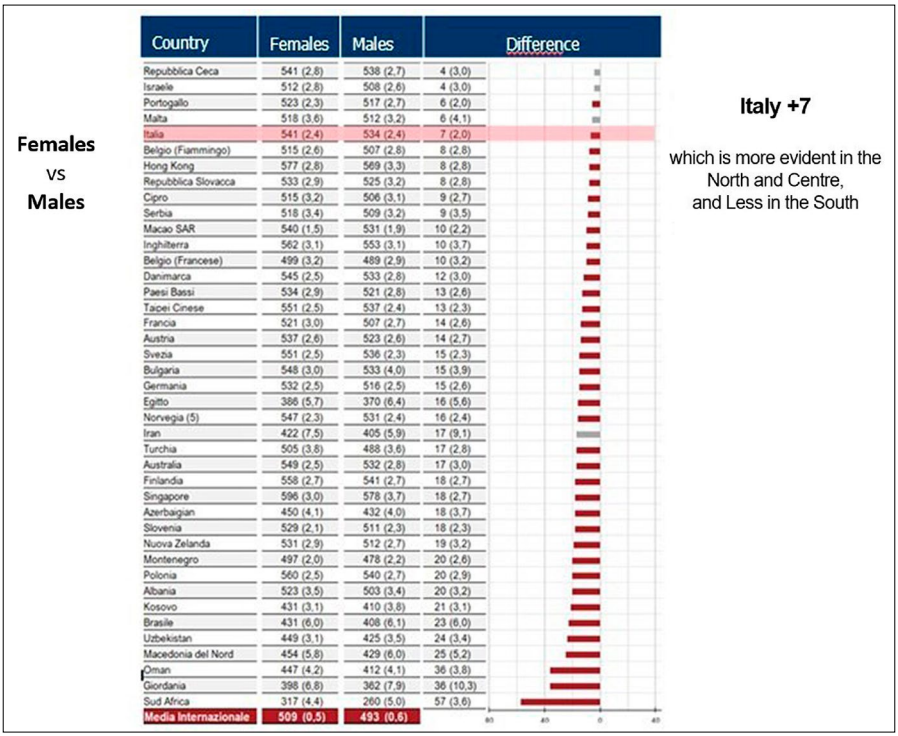
– North-West, North-East, and Centre:
  • these regions score above the Italian national average;
  • their performance aligns them with countries in the upper third of the international list, showing stronger educational outcomes.
– South and South/Islands:
  • these areas score significantly lower;
  • they are positioned closer to the bottom of the international ranking, just above Turkey and Montenegro.

The decline observed in all macro-areas could be due, at least in part, to the negative effects of the Covid-19 pandemic on student learning. The trend in the change in the average score of Italian students is similar to that of some European countries, such as France, Germany, and Spain, which recorded a decline in 2021 after remaining stable in previous editions. In contrast, other European countries, such as Ireland, Finland, and Poland, showed steady growth or stability at a high level throughout the editions of the survey.

Analysis of the results by gender shows that Italian female students have a 20-point advantage over their male peers (541 vs. 534), in line with the international average. This gender gap is present in all reading purposes and processes, but is widest in reading for literary experience (24 points) and in evaluating and criticizing content and textual elements (23 points). This means that Italian female students are better at reading narrative texts that require empathy and sensitivity, and at reading texts critically and reflectively, than their male peers.

The difference is for all countries in favour of females. Italy has one of the lowest differences which is higher in the North-West and Centre. Figure 10 shows the average reading scores broken down by gender in the 43 countries

involved. The countries are ordered according to the size of the gender gap in the results: ranging from no or very small differences between girls and boys to a significant difference in favour of girls.



Country	Females	Males	Difference
Repubblica Ceca	541 (2,8)	538 (2,7)	4 (3,0)
Israele	512 (2,8)	508 (2,6)	4 (3,0)
Portogallo	523 (2,3)	517 (2,7)	6 (2,0)
Malta	518 (3,6)	512 (3,2)	6 (4,1)
Italia	541 (2,4)	534 (2,4)	7 (2,0)
Belgio (Fiammingo)	515 (2,6)	507 (2,8)	8 (2,8)
Hong Kong	577 (2,8)	569 (3,3)	8 (2,8)
Repubblica Slovacca	533 (2,9)	525 (3,2)	8 (2,8)
Cipro	515 (3,2)	506 (3,1)	9 (2,7)
Serbia	518 (3,4)	509 (3,2)	9 (3,5)
Macao SAR	540 (1,5)	531 (1,9)	10 (2,2)
Inghilterra	562 (3,1)	553 (3,1)	10 (3,7)
Belgio (Francese)	499 (3,2)	489 (2,9)	10 (3,2)
Danimarca	545 (2,5)	533 (2,8)	12 (3,0)
Paesi Bassi	534 (2,9)	521 (2,8)	13 (2,6)
Taipei Cinese	551 (2,5)	537 (2,4)	13 (2,3)
Francia	521 (3,0)	507 (2,7)	14 (2,6)
Austria	537 (2,6)	523 (2,6)	14 (2,7)
Svezia	551 (2,5)	536 (2,3)	15 (2,3)
Bulgaria	548 (3,0)	533 (4,0)	15 (3,9)
Germania	532 (2,5)	516 (2,5)	15 (2,6)
Egitto	386 (5,7)	370 (6,4)	16 (5,6)
Norvegia (5)	547 (2,3)	531 (2,4)	16 (2,4)
Iran	422 (7,5)	405 (5,9)	17 (9,1)
Turchia	505 (3,8)	488 (3,6)	17 (2,8)
Australia	549 (2,5)	532 (2,8)	17 (3,0)
Finlandia	558 (2,7)	541 (2,7)	18 (2,7)
Singapore	596 (3,0)	578 (3,7)	18 (2,7)
Azerbaigian	450 (4,1)	432 (4,0)	18 (3,7)
Slovenia	529 (2,1)	511 (2,3)	18 (2,3)
Nuova Zelanda	531 (2,9)	512 (2,7)	19 (3,2)
Montenegro	497 (2,0)	478 (2,2)	20 (2,6)
Polonia	560 (2,5)	540 (2,7)	20 (2,9)
Albania	523 (3,5)	503 (3,4)	20 (3,2)
Kosovo	431 (3,1)	410 (3,8)	21 (3,1)
Brasile	431 (6,0)	408 (6,1)	23 (6,0)
Uzbekistan	449 (3,1)	425 (3,5)	24 (3,4)
Macedonia del Nord	454 (5,8)	429 (6,0)	25 (5,2)
Oman	447 (4,2)	412 (4,1)	36 (3,8)
Giordania	398 (6,8)	362 (7,9)	36 (10,3)
Sud Africa	317 (4,4)	260 (5,0)	57 (3,6)
Media Internazionale	509 (0,5)	493 (0,6)	

Fig. 10 – Differences in average reading scores by gender

Source: IEA PIRLS, 2023

Comparison with previous editions of the survey shows that the gender gap in favour of Italian female students remained substantially stable from 2001 to 2016, fluctuating between 18 and 21 points, and then increased in 2021, reaching 20 points. This increase could be due, at least in part, to the negative effects of the Covid-19 pandemic on boys' reading habits and motivations, which may have been more affected than those of girls.

Comparison with other participating countries shows that the gender gap in favour of female students is present in almost all countries, with a few exceptions, such as Chile, Colombia and Morocco, where there are no significant differences between the two genders, and Oman, where boys have a 7-point advantage over girls. The gender gap in favour of female students

is wider in some European countries, such as Finland (35 points), Ireland (32 points), Sweden (31 points) and Norway (30 points), and in some non-European countries, such as Hong Kong SAR (33 points), Taiwan (32 points) and Canada (31 points).

The differences in performance among Italian students are also interesting considering the socio-economic family background.

In order to collect the family background variables, PIRLS administers a family questionnaire, with questions concerning the possession of certain assets that are considered indicative of the level of economic well-being and the availability of educational resources: the number of books owned, the presence in the home of significant assets for study, such as an Internet connection and a room of one's own, the educational qualification and occupation of the parents. Considering the information collected, it was possible to construct a socio-economic and cultural index (SES).

As Figure 11 shows, the socio-economic and cultural index discriminates between the two categories of students: international results show that there is a 64-point difference in reading achievement between students with "high" and "low" SES (504 vs. 568).

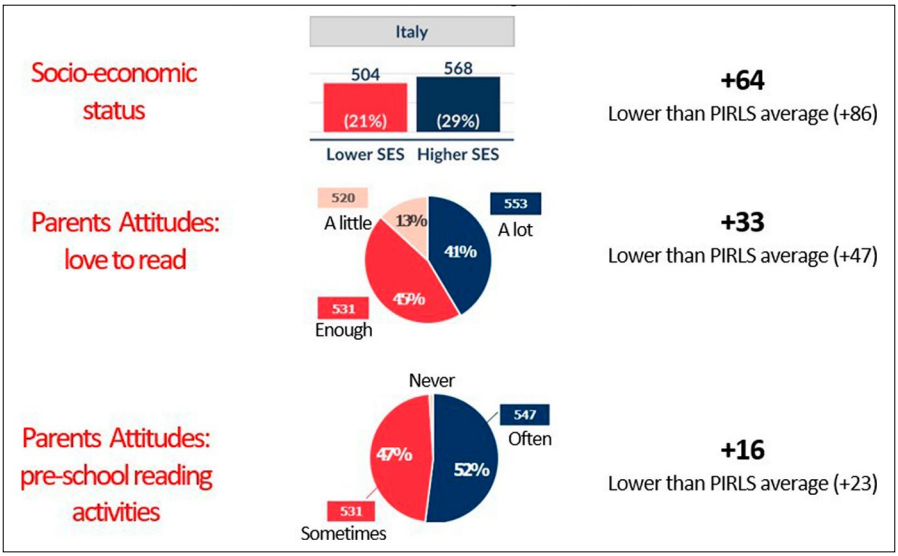


Fig. 11 – Parents' socio-economic status and Reading attitudes on children's performance

Source: IEA PIRLS, 2023

These results indicate that the family socio-economic context has a strong impact on the reading performance of Italian students, and that social inequalities are reflected in educational inequalities. These results are in line with those of other international surveys, such as PISA[1] and TIMSS[2], which show that Italy is one of the countries with the highest correlation between socio-economic background and students' skills.

Analysis of results by school context shows that Italian students attending schools with a positive climate have an advantage of 42 points over students attending schools with a negative climate (539 vs. 497). This difference is lower than the international average (51 points) and the European average (53 points). In addition, Italian students attending schools with good EMS have a 28 points advantage over students attending schools with bad EMS (532 vs. 504). This difference is similar to the international average (27 points) and the European average (28 points).

These results indicate that the school environment has a moderate impact on the reading performance of Italian students, and that schools that offer a safe, welcoming and organised environment promote student learning. These results are in line with those of other international surveys, which show that school climate and emergency management are important factors in the quality of education.

Analysis of the results by family context shows that Italian students living in families with a positive climate have a 44 points advantage over students living in families with a negative climate (537 vs. 493). This difference is lower than the international average (54 points) and the European average (56 points). Furthermore, Italian students living in families with a good availability of reading resources have a 63 points advantage over students living in families with a poor availability of reading resources (551 vs. 488). This difference is higher than the international average (46 points) and the European average (47 points).

[1] PISA is a three-yearly survey sponsored by the OECD that surveys 15-year-old students around the world on the degree to which they have acquired the knowledge and skills essential for full participation in social and economic life. PISA assessments are not limited to whether students at the end of compulsory schooling are able to reproduce what they have learnt, but also examine the level of preparedness of students and how well they are able to extrapolate from what they have learnt and apply their knowledge in unfamiliar contexts, both inside and outside school.

[2] The TIMSS survey is an IEA-sponsored survey and, every four years, measures the performance of students in primary school grade IV (grade 4) and secondary school grade III (grade 8) and monitors the implementation of school curricula in the countries participating in the survey.

These results indicate that family context has a significant impact on the reading performance of Italian students, and that families that provide an affective, stimulating environment and rich reading opportunities promote student learning. These findings are in line with those of other international surveys, which show that family climate and the availability of reading resources are key determinants of the development of reading skills.

7. Summary and final considerations

The analysis conducted in this study used an exploratory approach of the main results of PIRLS 2021, no assumptions were made, but a data-driven model was developed based on the theoretical framework of PIRLS.

Italy shows a good average result in reading competences comparing to other EU/non EU countries. Almost all our IV grade students them reach the basic benchmark, but there's much territorial variability when reaching Intermediate and High benchmarks.

Socio-economic context and parents' attitudes towards reading appear to be the most influencing factors studies have shown that children who are avid readers and who know the purpose of reading are more likely to have successful careers and a healthy lifestyle.

Agenda 2030 use this data to monitor progress.

These results highlight the need for action in the Italian education system to improve the quality of teaching and learning to read, and to reduce inequalities among students in relation to gender, socio-economic, school and family background.

Some possible actions to be taken are:

– promoting reading as an enjoyable, meaningful and cross-curricular activity, valuing the diversity of texts, genres and formats, and encouraging individual and shared reading, in and out of school;
– supporting the development of all reading purposes and processes, with a focus on reading to acquire and use information, and on evaluating and critiquing content and textual elements, which require skills of analysis, synthesis and evaluation, fundamental to critical and creative thinking;
– countering the gender gap in reading by providing boys and girls with opportunities and stimuli tailored to their preferences and motivations, and fostering an inclusive and equal reading culture that overcomes stereotypes and social expectations;
– compensating the socio-economic gap in reading, ensuring that all students have access to quality resources and services, and supporting the

most disadvantaged families and schools with targeted and customised interventions that take into account the needs and requirements of each individual;
– improving the school climate and management, creating a safe, welcoming and organised learning environment that fosters the well-being and participation of all those involved, and that is able to cope with the challenges and emergencies that arise, such as the Covid-19-related health emergency.

These actions require a shared and coordinated commitment on the part of all the parties responsible for the education and training of students, such as the Ministry of Education, the Regions, the Provinces, the Municipalities, the schools, teachers, school managers, parents, associations, libraries, publishing houses, the media and civil society. Only through effective collaboration and collective responsibility will it be possible to guarantee all Italian students the right to quality reading that helps them grow as aware and competent citizens.

References

Bowers J.S. (2023), "There is still little or no evidence that systematic phonics is more effective than common alternative methods of reading instruction: Response to Brooks", *Review of Education*, 11 (3), e3432.

Chen J., Zhang Y., Hu J. (2021), "Synergistic effects of instruction and affect factors on high- and low-ability disparities in elementary students' reading literacy", *Reading and Writing: An Interdisciplinary Journal*, 34 (1), pp. 199-230.

Graham S., Liu X., Aitken A., Ng C., Bartlett B., Harris K.R., Holzapfel J. (2017), "Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis", *Reading Research Quarterly*, 53, pp. 279-304.

Hu J., Dong X., Peng Y. (2022), "Discovery of the key contextual factors relevant to the reading performance of elementary school students from 61 countries/regions: insights from a machine learning-based approach", *Reading and Writing*, 35 (1), pp. 93-127.

INVALSI (2023), *INVALSI Report 2023*, retrieved on December 5, 2024, from https://invalsi-areaprove.cineca.it/docs/2023/Rilevazioni_Nazionali/Rapporto/Rapport o%20Prove%20INVALSI%202023.pdf.

Leino K., Nissinen K., Sirén M. (2022), "Associations between teacher quality, instructional quality and student reading outcomes in Nordic PIRLS 2016 data", *Large-Scale Assessments in Education*, 10 (1), 25.

Mullis I.V.S., von Davier V.S., Foy M., Fishbein P., Reynolds B., Katherine A. Wry, Erin (2023), *PIRLS 2021 international results in reading*, Boston College, retrieved on December 5, 2024, from https://www.iea.nl/publications/study-reports/international-reports-iea- studies/pirls-2021-international-results.

Mullis I.V.S., Martin M.O., Foy P., Hooper M. (2017), *PIRLS 2016 International Results in Reading*, TIMSS & PIRLS International Study Center, Boston College.
Mullis I.V.S., Martin M.O., Gonzalez E.J., Chrostowski S.J. (2004), *Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*, TIMSS & PIRLS International Study Center, Boston College.
Oberleiter S., Fries J., Schock L.S., Steininger B., Pietschnig J. (2023), "Predicting cross-national sex differences in large-scale assessments of William students' reading literacy, mathematics, and science achievement: Evidence from PIRLS and TIMSS", *Intelligence*, 100, 101784.
OECD (2023), *PISA 2022 Results*, vol. I: *The State of Learning and Equity in Education, PISA*, OECD Publishing, Paris, retrieved on December 5, 2024, from https://doi.org/10.1787/53f23881-en.
Reilly D. (2015), "Gender differences in reading from a cross-cultural perspective: The contribution of gender equality", *Proceedings of the International Convention of Psychological Science*, Amsterdam.
Scheeren L. (2022), "The Differential Impact of Educational Tracking on SES Gaps in Educational Achievement for Boys and Girls", *European Sociological Review*, 38 (6), pp. 942-958.
William H. R., Timothy R.B., William D.N. (2009), "Effective reading instruction for struggling readers: The role of direct/explicit teaching", *Reading & Writing Quarterly*, 25 (2-3), pp. 125-138.
Yan J., Cai Y. (2022), "Teachers' Instruction of Reading Strategies and Primary School Students' Reading Literacy: An Approach of Multilevel Structural Equation Modelling", *Reading and Writing Quarterly*, 38 (2), pp. 139-155.
Walberg H.J. (1984), "Improving the productivity of America's schools", *Educational Leadership*, 41, pp. 19-27.

 ISBN 9788835191087

The authors

Graziella Arazzi, Professor of Philosophy and History at the Liceo Human Sciences and Art of Imperia Collaborates with INVALSI (Area 5 Innovation and development). Latest publications: "School and territories: plots of coevolution to reinvent the world", *IUL Research*, 3, 5, 2022, pp. 230-239.

Michele Cardone holds a degree in Statistics for demographic and social sciences and a Master (I level) in Statistics for the management of information systems (University of Roma "La Sapienza"). Working for INVALSI since 2004, member of the Statistical Service since 2010, mainly involved in the analysis of the school tests data and in the management of the annual data return to schools.

Letizia Giampietro is a Senior Researcher at INVALSI, in the Innovation and Development Area. Her research interests range from educational policies to participatory evaluation models, with particular attention to the development of evaluation skills among school staff, collaboration between schools and local communities, and processes of school and social inclusion. On the applied side, she has taken part in the pilot phases for the launch of the National Evaluation System (SNV), as well as in PON, PRIN, and Erasmus+ projects managed by the research area.

Paola Giangiacomo is a Senior Researcher at the INVALSI Statistical Service, she serves as the Scientific Coordinator of the DIGCOMP research line, where she oversees the measurement of digital competences in upper secondary education. She also holds the position of National Data Manager for OECD-promoted international surveys. Her primary activities include the revision and adaptation of assessment instruments, the design of sampling plans, and the statistical analysis of quantitative and qualitative data

from large-scale national and international databases. Additionally, she is involved in drafting technical and scientific reports and conducting training activities in data analysis.

Marco Giganti is an Assistant Professor (in tenure track) at the Department of Human and Social Sciences, University of Bergamo. His research and teaching activities focus on educational assessment (formative assessment in Emergency Remote Teaching contexts; teachers' beliefs and practices; standardized assessment of digital competences; assessment through AI-based tools), pre-service and in-service teacher education, and empirical research methodology in education.

Emanuele Marcora has been Headmaster of the Istituto Omnicomprensivo Europeo of Arconate and Buscate (MI) since the 2019/2020 school year. With a degree in History, he has taught since the year 2001/2002 and is passionate about educational innovation in all its forms, particularly the construction of digital learning environments. He was invited as Keynote Speaker to the official closing event of the open public consultation for the New Digital Education Action Plan promoted by the European Commission – Directorate General for Education, Youth, Sport and Culture.

Sara Mori is a researcher at INDIRE. She has a PhD in Evaluation of Educational Systems and Processes; she is a psychologist and a cognitive-behavioural psychotherapist. She has worked at INDIRE since 2011 and has been an adjunct lecturer at IUL since 2013. Her main research interests are the assessment and development of transversal skills and the development of students' motivation and well-being, with a focus on methodologies for the personalisation of teaching.

Jessica Niewint-Gori is the Head of the Research Department for "Teaching, Didactics and Competences" at INDIRE (Italy's National Institute for Documentation, Innovation and Educational Research). She is principal researcher in European and national projects focused on STEM education, AI literacy, and data literacy, including Erasmus+ initiatives on generative AI in education. She collaborates with UNESCO, the European Commission, and the Council of Europe on policy development and teacher training frameworks for AI integration in schools.

Giuseppe C. Pillera, PhD, Researcher at University of Catania. He has been researcher at INVALSI, research fellow at the University of Catania,

lecturer at the universities of Catania and Messina and visiting researcher at the Universidad de Sevilla. He is focused on evaluative issues and methodologies, principally related to e-learning, teacher training, intercultural education.

Donatella Poliandri, PhD, is Senior Researcher and Head of the Innovation and Development Area at INVALSI. She contributed to the design of the Italian National School Evaluation System and leads national and international research projects. Her research focuses on social research methodology and the evaluation of educational policies, with particular attention to building and supporting schools' evaluation capacity through the design of tools and structured processes for self-evaluation and improvement.

Francesca Storai is a researcher at INDIRE. She has dedicated herself to the study of innovative training models for teachers and school leaders. She has conducted in-depth research on methods and tools for developing improvement plans and on educational innovation.

Valeria F. Tortora is a researcher at INVALSI, where she is National Manager for the International Association for the Evaluation of Educational Achievement (IEA). She is PhD in Comparative Education, the most recent research concerns the study of social inequalities, the variables connected to educational performance of students.

Questo

LIBRO

ti è piaciuto?

Comunicaci il tuo giudizio su:

www.francoangeli.it/opinione

VUOI RICEVERE GLI AGGIORNAMENTI
SULLE NOSTRE NOVITÀ
NELLE AREE CHE TI INTERESSANO?

ISCRIVITI ALLE NOSTRE NEWSLETTER SEGUICI SU:

FrancoAngeli
La passione per le conoscenze

Vi aspettiamo su:

www.francoangeli.it

per scaricare (gratuitamente) i cataloghi delle nostre pubblicazioni

DIVISI PER ARGOMENTI E CENTINAIA DI VOCI: PER FACILITARE
LE VOSTRE RICERCHE.

Management, finanza, marketing, operations, HR
Psicologia e psicoterapia: teorie e tecniche
Didattica, scienze della formazione
Economia, economia aziendale
Sociologia
Antropologia
Comunicazione e media
Medicina, sanità

Architettura, design, arte, territorio
Informatica, ingegneria
Scienze
Filosofia, letteratura, linguistica, storia
Politica, diritto
Psicologia, benessere, autoaiuto
Efficacia personale
Politiche e servizi sociali

FrancoAngeli

La passione per le conoscenze

Recent changes to the school system have made assessment an increasingly important tool for reflection and improvement. Moving away from a purely measurement-based approach means recognising the value of assessment as a shared process that can inform educational and organisational decisions. With this in mind, we have collected contributions presented at the VIII Seminar, "INVALSI data: a tool for teaching and scientific research", held in Rome from 23 to 26 November 2023. This publication contributes to the ongoing dialogue surrounding the evolution of education systems, with a specific focus on evaluation, self-evaluation, and teaching innovation processes within Italian and European schools. The experiences and research presented all focus on the role of assessment and reflective practices in supporting the professional development of teachers and school leaders, improving teaching and learning processes, and strengthening the governance of educational institutions.
The volume therefore aims to facilitate dialogue between research, practice and policy for those engaged in rethinking schools as organisations capable of learning, evaluating themselves and innovating in a responsible and sustainable manner.

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