INVALSI DATA SUPPORT IN POLICIES AND SCHOOLS

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

edited by
Patrizia Falzetti

FrancoAngeli

INVALSI

INVALSI PER LA RICERCA
STUDI E RICERCHE

INVALSI PER LA RICERCA

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

Copyright © 2024 by FrancoAngeli s.r.l., Milano, Italy & INVALSI. ISBN 9788835167938

INVALSI DATA SUPPORT IN POLICIES AND SCHOOLS

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

edited by
Patrizia Falzetti

FrancoAngeli

 ISBN 9788835167938

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.

Copyright © 2024 by FrancoAngeli s.r.l., Milano, Italy & INVALSI. ISBN 9788835167938

Index

 ISBN 9788835167938

Introduction

by Patrizia Falzetti

Education is a crucial issue for the economic, cultural and social development of all countries. International empirical evidence has actually highlighted the close correlation between school learning and the economic growth of a country, reaffirming the value of education as a central development lever. INVALSI data, therefore, in addition with other data sources – such as the ones of the Ministry of Education and Merit – may be precious in order to know, better understand, and deepen school processes. INVALSI tests, in particular, provide both educational institutions and policy makers with a wide range of useful and important information to respond to different school needs.

This volume collects some research papers presented at the sixth edition of the Seminar "INVALSI data: a tool for research and teaching" (held in Rome from 25th to 28th November 2021) aimed to explore this topic.

The first two chapters are about two experiences on policy evaluation which combine several data sources, highlighting the potential of this interaction. Specifically, the first chapter focuses on a new ISTAT system for the analysis of phenomena related to education and training that goes beyond the logic based on the use of individual sources to produce specific outputs, adopting instead a logic of integration. The second chapter analyses the mismatch between the quantity and the quality of data made available by ISTAT, INVALSI and the Ministry of Education and Merit and the actual knowledge needs of the local authorities responsible to plan the school networking and the educational offer.

In chapters three and four the focus is on the use of INVALSI data in teaching. In the third chapter, a cross-disciplinary research project involving experts in mathematics didactics, general didactics and evaluation is presented; its aim is to investigate primary school teachers' beliefs and attitudes

about the link between INVALSI mathematics tests and teaching-learning processes. While in the final chapter there is a research which, using INVALSI data, aims on the one hand to identify which assessment practices are most frequent by Italian teachers and on the other hand to understand whether assessment practices tend to influence students' learning and, thus, also their performance in INVALSI tests.

As the Statistical Service, we hope that these examples of the use of INVALSI data, either alone or integrated with other databases, will reinforce the belief that a proper use of data can make an important and crucial contribution to decision-making processes and be a determining factor in making worthy and strong decisions.

1. Towards an integrated system for the production of relevant statistical data on education and training

by Giovanna Brancato, Donatella Grassi

ISTAT proposes a revision of the current system of production of statistical indicators to analyse the phenomena of education and training, with the aim of ensuring its current statistical production of ISTAT and expanding its information potential, maintaining the quality of estimates and reducing costs and burdens. The new system will enable us to transcend the current logic of stovepipe processes, which relies on the use of one single source to produce specific outputs, and replace it with a data integration logic. Its core will be the Education and Training Register, which will be part of ISTAT Integrated System of Statistical Registers. The chapter describes the planned and ongoing activities and the methodological issues for the construction of the new integrated system of statistical production.

Per l'analisi dei fenomeni connessi all'istruzione e alla formazione, l'ISTAT propone la revisione dell'attuale sistema di produzione degli indicatori statistici, con la finalità di garantire la produzione corrente e ampliare le potenzialità informative, assicurando la qualità delle stime e la riduzione dei costi e del burden *statistico. Il nuovo sistema supererà la logica attuale di tipo* stovepipe *basata sull'utilizzo di singole fonti per produrre specifici output, adottando invece una logica di integrazione. Esso avrà come fulcro centrale il Registro dell'Istruzione e Formazione, che sarà parte del Sistema Integrato dei Registri Statistici dell'ISTAT. Il lavoro illustra le attività e gli aspetti metodologici pianificati e in corso per la costruzione del nuovo sistema integrato di produzione statistica.*

1. Introduction

All national and international agendas for sustainable development agree that qualified human capital and interventions against educational poverty are essential to a country's growth.

The fourth goal of the UN 2030 Agenda for Sustainable Development is in fact to «ensure inclusive and equitable quality education and promote lifelong learning opportunities for all»[1].

Along the same lines, the Italian National Recovery and Resilience Plan[2] (that is part of the Next Generation EU programme), acknowledges the significance of this sector for the harmonious and prosperous development of the Country. With its fourth mission ("Education and research") it allocates almost 20 billion euros for Component 1 ("Strengthening the supply of education services: from pre-primary schools to universities") through a system of financing and reforms that aim at:

- a qualitative improvement and expansion of education and training services;
- a reform of the legislation on teachers' careers, with particular reference to the recruitment and training processes;
- the improvement of skills – with particular reference to STEM (Scientific, Technological, Engineering, Mathematical) disciplines and multilingualism;
- the strengthening of infrastructures, in particular in terms of safety and energy efficiency;
- a reform of university degree programs and PhD courses.

In this context, official statistics should measure educational phenomena in order to highlight critical issues and identify areas and groups in need of improvement, namely to drive public policy.

Therefore, a new system of indicators should be designed, implemented and regularly updated. The system should support the monitoring of the phenomenon in all its dimensions. They include the participation in education, the supply and demand of educational services in different educational programs and levels, qualitative and quantitative skills achievements, the extent of early leavers, the student/staff ratio, the infrastructures conditions, possible gaps in education (by territory, gender, citizenship, etc.), jobs opportunities, skills mismatch, interactions between education and social exclusion risks.

[1] https://sdgs.un.org/goals/goal4.

[2] https://temi.camera.it/leg18/provvedimento/piano-nazionale-di-ripresa-e-resilienza.html.

The aim of this work is to show: i) how the Italian National Statistical Institute (ISTAT) has accomplished this task over time by using in parallel several sources for the analysis of this complex phenomenon and ii) the future developments of the system which improve the current one by integrating and increasing information capacity and efficiency.

2. Data sources

A wide range of information is available to analyse education and training phenomenon. Besides ISTAT surveys, data come from many institutions such as: the Ministry of education (ME), the Ministry of university and research (MUR), the National institute for documentation, innovation, educational research (INDIRE), the National institute for the evaluation of the education and training system (INVALSI), the National institute for analysis of public policies (INAPP) and the National agency for active labour policies (ANPAL)[3].

In some cases, the sources contain information suitable for the statistical purposes; in others, the variables may not be directly usable, due to the different nature (administrative vs. statistical), different concepts/definitions, different measuring tools. In such situations, it is necessary to evaluate the most suitable source for the statistical representation of the analysed phenomena. «Comparability is the extent to which differences between statistics from different geographical areas, non-geographical domains, or over time, can be attributed to differences between the true values of the statistics»[4]. It means that comparisons, both at national and international level, are possible once definitions and concepts underlying the variables are made comparable to all extent. In this context, the national classification on education and training programs and qualifications must be matched and harmonised with the levels and the fields of the international classifications ISCED 2011 (UNESCO, 2012) and ISCED-F 2013 (UNESCO, 2014) respectively, and with the CLA (EUROSTAT, 2016) type of learning activities, distinguishing formal education from non-formal and informal education programs.

Some sources are harmonised with the international classifications by design (it is the case of the European surveys conducted by ISTAT), while in

[3] The Annex contains a brief description of the different sources and their possible use in the production of statistical indicators.

[4] OECD Glossary of statistical terms https://stats.oecd.org/glossary/detail.asp?ID=395.

other cases the harmonisation takes place *a posteriori* (it is the case of the administrative sources). In this regard, ISTAT, in collaboration with the data providers, is completing a specific classification on the education and training programs and qualifications where each Italian formal education program is mapped to the international classification ISCED 2011 and ISCED-F 2013 (see paragraph 2.2.2) codes.

2.1. The system of indicators on education and training

For the multidimensional analysis of phenomena related to education and training, ISTAT has built over time a system of indicators based on surveys and administrative data available at individual and/or institution level (Tab. 1).

The resultant picture is almost complete and offers an accurate overview of all sectors of education and training both for analytic purposes and for the main dissemination scopes (national and international releases, data warehouses, microdata).

However, such a stovepipe approach, where sources are independent one each other, may implicate some limitations and disadvantages. Redundancies and possible discrepancies on some variables require a deep qualitative analysis to identify the most reliable source taking into account that often variables, definitions and concepts are not harmonised. Information gaps are also possible and for some sub-populations micro-data are still not available. Burden for respondents is high as well as survey costs. Finally, longitudinal analyses (e.g. school to university, education to work) are not possible.

These are the reasons why a new approach for the use of all the available sources is being developed. Its core is represented by the education and training register (ETR). Through a preliminary phase of harmonisation, rationalisation and treatment, ETR will integrate the information coming from the different sources and, as described in the following paragraphs, it will be part of the broader ISTAT system of registers.

Tab. 1 – Italian indicators system on education and training

Data type	*Source*	*Main observation units*	*ISCED*	*CLA**	*Main indicators*
Administrative data	National register of school students (ME)	Students, Schools	0.2-3	F	Enrolled and graduates by individual characteristics, school grade and program of education
	School staff database (ME)	Staff individuals	0.2-3	F	Staff by individual characteristics, contract type, and private/public administration
	School building register (ME)	State school buildings	0.2-3	F	Infrastructure and building status by geographical area
	Higher Technical Institutes database (INDIRE)	Students, Institutions	5 (554)	F	Enrolled and graduates by individual characteristics, school grade and program of education
	National register of students and graduates (MUR)	Students, Universities	6-8	F	Enrolled and graduates by individual characteristics, level and program of education
	University staff database (MUR)	Staff individuals	6-8	F	Staff by gender, public/private university and program of education
Survey data	European Adult Education Survey (ISTAT)	Individuals		F NF I	Comparable-across-countries indicators on adult education participation by individuals characteristics
	European Labour Force Survey (ISTAT)	Households, Individuals		F NF	Comparable-across-countries indicators on education participation and work opportunities by individuals characteristics
	Survey-system on the education-to-work-transition (ISTAT)	Graduates	3-8	F	Education-to-work-transition by level of education and other individuals' characteristics
	Survey on school services for early childhood (0-2) (ISTAT)	Service providers	0.1	F	Services for early childhood on population 0-2, number of enrolled
	Household Budget Survey (ISTAT)	Households	0-8	F NF	Expenditure for education goods and services by level of education and household characteristics, poverty indicators by education level
	European Income and Living Conditions Survey (ISTAT)	Households			Income and Social exclusion indicators by education level, household difficulties in affording education costs
	Multipurpose Survey: aspects of daily life (ISTAT)	Household, Individuals			School services accessibility, participation and quality of online lessons during Covid-19 pandemic
	Surveys on schools (ME)	Schools	0.2-3	F	Additional data on enrolled and graduates by individual characteristics, students per class ratio
	Survey on AFAM institutions (MUR)	Institutions	6-8	F	Enrolled and graduates by individual characteristics, grade and program of education, staff characteristics
	National Surveys on student learning (INVALSI)	Students, Class	1-3	F	School context, students' characteristics, scores obtained in the standardized test
	Survey on IeFP (INAPP)	Courses	3 (353)	F	Courses, enrolled and graduates by individual characteristics

* In accordance with the CLA: F = Formal education; NF = Non-Formal education; I = Informal education

Consequently, the available information will be exploited at its best and the new integrated system will enable to:

– guarantee ISTAT current statistical production, assuring coherence of the estimates with those from the data providers (e.g. with data disseminated by ME, MUR and INVALSI) and timeliness of the results;
– widen the potentialities of analyses, by optimising the complementarity of the sources, further expanding the detail of information, improving territorial coverage, allowing the production of derived variables and generalised indicators;
– reduce the respondents' burden and the production costs;
– allow to easily correlate different phenomena, primarily education and occupation, both at cross-sectional and longitudinal level;
– support other ISTAT production processes.

3. The methodological development of the new statistical production system

The use of administrative data organised into statistical registers has a long tradition and is highly spread in several northern European statistical institutes[5]. In 2016, ISTAT introduced a modernisation program, where one of the pillars was the shift towards register-based-statistics (ISTAT, 2016). Following what proposed in the literature (Wallgren and Wallgren, 2014), ISTAT Integrated System of Statistical Registers (ISSR) is based on the development of two types of registers: i) *base registers* on relevant statistical populations (individuals, economic units, places and activities); ii) *satellite registers*, categorised into *extended registers* (which extend with additional variables the information available for a specific population of a *base register*) or *thematic registers* (which offer a detailed picture of a statistical domain).

Such a system entails a relevant methodological investment concerning for example the conceptual and physical integration among the registers, the coherence of the estimates produced by different registers, stock-flows data harmonisation and specific architectural solutions for privacy and confidentiality concerns (Alleva *et al.*, 2019; Radini *et al.*, 2018).

As mentioned, the new production system of education and training statistics hinges around the ETR. It will be one of the *thematic registers* of the

[5] See Seminar on Register Statistics – 21-23 May 2007, https://www.stat.fi/registerseminar/index_en.html.

ISSR. In order to meet the information needs that the sole register cannot satisfy, the ETR will be complemented by a set of surveys.

Indeed, once the ETR will be developed and the information potentialities coming from the ISSR fully comprehended, the whole system of ISTAT current surveys on education and training, presented in section 2 and in the Annex, will be rethought in order to exploit the information provided by the registers and integrate it with more qualitative information. In the new context, the role of the surveys in the system will be twofold. On the one hand, they will use register's data to support the design and estimation phases (pre-filling of questionnaires, definition of sampling frames, editing and imputation procedures, weighting systems). On the other hand, they will provide information to the register to enrich it with additional contents.

3.1. Content of the ETR

ETR will concern several statistical populations and relative variables. It will identify individuals currently attending any program of the Italian education and training system. Furthermore, it will allow the assignment of the education level to individuals. For the latter purpose, in order to have as much data as possible historical information from other ISTAT sources, data from the population and households' censuses will be retrieved. Concerning the feeding of the ETR with the administrative sources, it is planned to use data starting from the academic/school year 2010-2011.

The main statistical unit will be represented by the composite unit "educational position", obtained by the relationship, in a given period of time, among the following three elementary units: i) individual, ii) education and training institution and iii) education and training program. Besides the above listed units, other units, such as internships, school buildings, staff, etc. will be in the ETR.

Each unit will be supplied by a set of specific variables. For the educational position the main variables will be starting and ending times, events of changes and completion of the program, events of movements from an institution to another. They will be obtained mainly by ME, MUR, INDIRE.

Concerning the individuals and institutions, a subset of variables is obtained by linking within the ISSR. In particular, demographic characteristics of the individuals will be drawn from the relative *base register*. The legal form of the education and training institutions will be obtained from the *base register* on economic units. Where applicable, e.g. for residence of the indi-

viduals, localisation of universities and schools, the assignment of a unique address code will be derived from the *base register* of the places.

In addition, ETR will record information on the quality of the learning skills provided by the individual results to the INVALSI tests. Variables concerning all the attainments (e.g. date, performance, effective duration of completed programs) and the internships (e.g. period, hours, partners, economic activity) will be recorded. The register will also contain context information concerning the education and training institutions, e.g. university dimensional class, equipment and conditions of the schools' buildings, average performance indicators at class level (from the INVALSI tests). Data on staff working in the education and training institutions will also be considered, taking into account the information already available in other ISTAT *thematic registers*.

Concerning the education and training program, it will be linked to the new version of an enriched classification on education programs and qualifications, as illustrated in the next paragraph.

3.2. Metadata and quality indicators

The development of such a new production system requires the parallel development of the supporting systems managing the metadata and the quality indicators. Statistical metadata are information on statistical data. It is common to distinguish between structural and reference metadata (Androvitsaneas *et al.*, 2006). The former are metadata that act as identifiers of the structure of the data and the associated information. The latter are those describing the content, methodologies and quality of the statistical data. Statistical classifications are included in the structural metadata while quality indicators are managed together with the reference metadata.

ISTAT can rely on a unified system for the management of structural and reference metadata, described in Signore *et al.* (2015). Standard quality indicators are collected at central level and regularly analysed in order to provide objective evidence for ISTAT quality policy (Brancato *et al.*, 2006).

The definitions of the underlying concepts, as well as the statistical and methodological activities concerning the new ETR, are being developed ensuring coherence and harmonisation with the above-mentioned metadata standards used at ISTAT. The identification of the concepts and the conceptual relationships among sources is supported by the adoption of an ontology-based data management approach (Radini *et al.*, 2018).

Concerning the development of a system of quality controls and indicators to support the register and the produced statistics, a quality framework,

developed within a European collaborative project, will be adopted. This approach defines quality considering several perspectives: input-process-output quality (Ascari *et al.*, 2020).

As part of the study of the quality of the sources and the pre-treatment and treatment procedures, the first specific quality controls are being identified. They will provide the quality indicators useful for the monitoring of the annual statistical process used to update the ETR, as well as the quality assessment of the final products.

3.3. The new classification of education and training programs and qualifications

A pillar of the new education and training statistical system is the specific classification on the education and training programs and qualifications (Cascioli and Grassi, 2021). It also represents an essential tool supporting the ETR.

The new classification must meet the following requirements: i) enable the assigning of unambiguous codes to all current and past programs in statistical surveys; ii) ensure that the information in current and past administrative data is properly mapped; iii) ensure comparability over time and alignment with international standards; iv) be flexible to changes in the education and training system; v) be supported by a mechanism for its maintenance over time.

As already mentioned, the statistical production on education and training is supported by the international classifications: ISCED 2011, ISCED-F 2013, CLA 2016.

At the national level, ISTAT released the Classification of the Italian Qualifications in 2003[6]. Since then, many important reforms of the education system have been endorsed, changing substantially the landscape of the education and training offer in Italy.

Consequently, there is the need to update the Italian classification, assuring harmonisation with the international ones.

ISTAT, together with the national relevant institutions[7], is committed in the update and enrichment of the national classification. A preliminary ver-

[6] The classification is available at the following link: https://www.ISTAT.it/en/archivio/6640.

[7] ME, MUR, Ministry of Labour, ANPAL, INAPP, Regions and other territorial Institutions.

sion of the classification has been prepared. It has a hierarchical structure, from the more aggregated information to the more detailed one. It allows direct mapping with last versions of ISCED and ISCED-F and it takes into account all the normative reforms that have taken place over the time. It considers both the education programs and the qualifications, reporting for each of them the theoretical duration.

3.4. Methods for the design and implementation of the new statistical production system on education and training

In this paragraph, the main activities for the design and implementation of the new system are illustrated.

– *Information needs.* The starting point for the development of the system is the review of the actual and potential statistical information needs. To this aim, all the tables published in ISTAT dissemination channels have been analysed and disentangled in order to formally identify the underlying units, variables and classifications. Moreover, consultations have been conducted to take into account for other ISTAT producers of statistics on education and training, for example: Labour Force statistics, Education-to-work transition surveys, Permanent census of population and housing, Sustainable Development Goal 4, Equitable and sustainable well-being indicators. A reflection on potential information needs has been carried out, although in a less structured form.
– *Architecture of the register.* As for the other ISTAT statistical registers, the ETR data architecture will be organised into two main areas: a "working area", in which the yearly data are stored, and a "longitudinal area" where the historical data are structured. This implies to consider the statistical procedures and quality activities for each production step, from the acquisition of the input sources, to the structuring of the data in the working area, to the organisation of the information in the longitudinal area.
– *Integration in the ISSR.* To increase the information potentialities within the ISSR, it is deemed necessary to establish the relationships between the statistical units in the ETR and those in the other *base registers*. Indeed, this is a precondition for the linkability of the information. It is first necessary to compare the definitions of the units in the various registers, to establish the relationships at a conceptual level also with the support of the ontology models, then to build the physical connections. It has to be mentioned that in the ISSR the management of the variables follows structured governance rules. For each variable, it is established which

register is in charge of its management and the same variable cannot be in different registers. The access and the modification/update of the variables is allowed based on the profile of the user and the related permissions. Such a system avoids redundancies and facilitates coherence (ISTAT, 2016).

– *Data sources.* Another important activity has concerned the review and analysis of the input sources. They have been deeply analysed with respect to the metadata and the data quality. As mentioned, ISTAT has been acquiring the majority of the administrative sources since many years. However, as already stated, so far, their content has not been fully exploited for statistical purposes. Therefore, for each administrative dataset, the definitions of the variables included are being further investigated and quality controls on the extent of missing values and internal consistency are being planned.

  We intend to gradually increase the information and its detail (both at unit and variable levels) to be integrated in the new system. Besides INVALSI data, it is planned to study the possibility to integrate other administrative datasets, e.g. the dataset of the students enrolled in the Universities of higher education, research and high-level training with a special status; the results of the tests performed for the access to some universities' courses and the open data on the school building.

  Finally, it is worth to mention that ISTAT has already produced a preliminary integrated database on Education and Qualifications, named BIT (Runci *et al.*, 2017). BIT collects, checks and integrates data from different sources, on a yearly basis. It produces data on the attained level of education and the attendance to a course of students currently in the education system. The BIT database is used in a prediction procedure to estimate the attained level of education for the individuals in the relative *base register* (Di Zio *et al.*, 2019). For the development of the ETR, the BIT is being deeply studied. It represents a relevant experience for the planning of the ETR and a source of data for feeding part of the new register data, especially with historical information on the attainments.

– *Statistical procedures.* The next step consists in selecting the variables from the sources and defining the procedures to apply to the variables. The selected variables are assigned different roles: some of them are selected as direct information, others for quality checks. In general, statistical procedures are aimed at data acquisition, transformation, integration and estimation. As mentioned before, they are defined for the passages from the input sources to the data in the working area and from the latter to the longitudinal area. Along with the statistical procedures, quality

controls are identified following a quality framework. They can concern a single variable in a dataset, or its relationship with other variables in the same or in other datasets. For each source, input quality indicators are defined, for example a check between the age of the student and the school grade attended. Once the data are integrated, indicators on consistency among related variables are defined, e.g. in a given year, an individual cannot attend different grades in different schools or in the longitudinal data, propaedeutic conditions have to be met. Indicators on output quality will be produced and in particular, indicators on the coherence between statistics disseminated using ISTAT register and those from the data providers (e.g., ME, MUR and INVALSI).

– *Additional information from thematic registers of the ISSR and survey data.* One the advantages of the inclusion of the ETR in the ISSR is the possibility to relate the information on education and training with information available from other *thematic registers*, under the privacy and confidentiality constraints. It is intended to study the education phenomenon in relation with the work opportunities, thus relating the ETR with the Labour Register (LR) and the Income Register (IR).
Finally, the new system requires a re-design of some current surveys in the light of the availability of the ETR and its links with other *thematic* registers. In particular, ISTAT surveys on the education-to-work transitions in upper secondary, university and PhD graduates will be re-considered. It is planned to partially substitute these surveys with data available from the ISSR and maintain only reduced surveys more focused on the qualitative information, not available from the administrative sources and the ETR. This issue will be further explored, by means of an example, in the next paragraph.

4. Case studies

In this paragraph, two different planned case studies are presented. The first one is a simulation showing how the information coming from the ETR and the other registers of the ISSR can be exploited to reduce redundancies and survey burden. The second one shows the benefits deriving from the integration of the register data with auxiliary information that better describes the context and the quality of the individual education and training history.

4.1. Feasibility analysis for a new approach in the education-to-work transition study

The first case study aims to explore the feasibility of using ISSR data to derive, at least partially, the information coming from the three surveys on study-to-work transition conducted on upper secondary, university and PhD graduates[8]. These surveys collect information on the type of studies conducted, the occupational status a few years after graduation, the type of job and the activity sector, the income, the job satisfaction and skill mismatches.

The idea is not completely new, as ISTAT has already explored the information potentialities of micro-data integration using the data from the 2014 Survey on PhD graduates and data on occupation (Fraboni *et al.*, 2019).

Since then, employment data have been further structured and the *thematic* LR is now in a more advanced development phase. It provides longitudinal information on the population of employed and self-employed workers.

With a LEED structure (employer-employee linked data), the main statistical unit of the LR is the job position (more than 25 million job positions), defined as the work relationship between an employer and an employee with a starting date and the type of employment (dependent, not dependent and so on). From this unit, other statistical units can be derived, such as the individual worker or the employer. In addition, the register contains information related to the characteristics of the work (type of occupation, duration, contractual qualification, type of contract, seniority on the job position), and variables related to wages, labour costs, gross labour income, paid/worked hours (Baldi *et al.*, 2018).

The goal of the application is to utilise administrative data on PhD graduates provided by MUR, simulating they are a prototype of the ETR on a specific sub-population.

The data will be linked with ISTAT *base register* on individuals, with the LR and with other relevant administrative sources in order to derive part of the information provided by ISTAT 2018 Survey on PhD graduates (for the two cohorts of 2012 and 2014).

The link with the individual *base register* enables the identification of graduates who leave the cohort due to demographic events (such as death or migration), as well as information on the parents. MUR sources are used to find signs of re-enrolment in the university system, participation to research projects, fellowships, and other related activities. LR provides information

[8] See https://www.ISTAT.it/it/archivio/224302 and https://www.ISTAT.it/it/archivio/190692 for published data on PhD, upper secondary and university graduates.

on study-to-work transition and the working condition 1-3 years after PhD graduation. It is clear that the subjective information gathered in the survey, which includes all the questions on satisfaction issues, cannot be derived from the administrative sources and the registers. Observing this information and integrating it into the register data may be possible in the future through more focused and less burdensome surveys.

4.2. Education and training pathways and quality of learning in ETR

The second case study focuses on the plan for integrating INVALSI data into the ETR. As previously stated, the INVALSI tests are an external, objective, and impartial assessment based on the measurement of the quality of learning abilities in Italian, Mathematics and English (reading and listening). INVALSI methodology is highly developed and takes into account the pupils' background (social, cultural, family factors summarised in a socio-economic context indicator) and for cheating probability, by means of introducing a correction factor[9].

These data are highly interesting for the register as, contrary to the majority of the available administrative sources, they represent information on the quality of the education, tailored for the age of the pupils.

The information on INVALSI can be linked in ETR to the educational position unit and its elementary units (individual, institution and program).

For the individual (pupil), levels and scores of the different performed INVALSI tests will be integrated in the system. In addition, the individual socio-economic status indicator will be taken.

At institution (school and class) level, it is planned to integrate in the ETR average scores and distributions of the levels obtained for the various tests. These indicators will contribute to outline a profile of the schools together with other data on the availability of facilities (canteen, gym, equipment against physical barriers for disablers) and equipment of IT facilities coming from the National Register on School Buildings.

Similarly, at program level, average scores and levels' distributions will be linked to different pathways, for example general vs. vocational program.

INVALSI data will represent relevant information to be usefully included as explicative variables in statistical models exploring the various factors influencing the education and training phenomenon.

[9] https://www.invalsiopen.it/.

5. State of the art and conclusions

This chapter details the main plans and actions being taken to modernise, enrich, and improve the ISTAT system of statistical production in education and training.

An in-depth analysis and rationalisation of the administrative sources is currently under way.

With the involvement of thematic experts, a comprehensive and enhanced national classification of education and training programs and qualifications is being validated.

The ETR is in its design phase. The required output statistical information, in terms of units, variables and classifications, has been identified. The definitions concerning the main units of the register have been developed. The units and variables to be selected by different sources have been identified together with the definition of the preliminary quality controls. The conceptual relationships of the ETR within the ISSR are being formalised. Simulations on possible information integration among registers are being designed.

In conclusion, it seems clear that the understanding of complex phenomena such as education and training, and all its interactions with other phenomena, requires structured and multi-perspective data systems. In order to respond adequately to this challenge, it will be important to create synergies across the different institutions and to be able to exploit efficiently all the available information assets. Investments will also be needed to improve the availability and detail of data for some segments of education and training.

References

Alleva G., Falorsi P.D., Luzi O., Scannapieco M. (2019), *Building the Italian Integrated System of Statistical Registers: Methodological and Architectural Solutions ESS Workshop on the use of administrative data and social statistics*, Valencia, 4-5 June 2019, retrieved on January 25, 2023, from https://ec.europa.eu/EUROSTAT/cros/system/files/building-italia-integrated-system_ISTAT_0.pdf.

Androvitsaneas C., Sundgren B., Thygesen L. (2006), *Towards an SDMX User Guide: Exchange of statistical data and metadata between different systems, national and international*, OECD Expert Group on Statistical Data and Metadata Exchange, Geneva, 6-7 April 2006.

Ascari G., Blix K., Brancato G., Burg T., McCourt A., van Delden A., Krapavickaitė D., Ploug N., Scholtus S., Stoltze P., de Waal T., Zhang L.C. (2020), "Quality of Multisource Statistics – the KOMUSO Projec", *The Survey Statistician*, 81, pp. 36-51.

Baldi C., Ceccarelli C., Gigante S., Pacini S., Rossetti F. (2018), "The labour register in Italy: the new heart of the system of labour statistics", *Rivista Italiana di Economia Demografia e Statistica*, 52 (2), pp. 113-124.

Brancato G., Pellegrini C., Simeoni G., Signore M. (2004), *Standardising, Evaluating and Documenting quality: the implementation of ISTAT information system for survey documentation – SIDI*, Proceedings of the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, 24-26 May 2004.

Cascioli R., Grassi D. (2021), *Revisione e aggiornamento della classificazione dei percorsi e dei titoli di studio*, XIV Conferenza Nazionale di Statistica, ISTAT, 30 novembre-1° dicembre 2021.

Di Zio M., Filippini R., Rocchetti G. (2019), "An imputation procedure for the Italian attained level of education in the register of individuals based on administrative and survey data", *Rivista di Statistica Ufficiale*, 2-3, retrieved on January 25, 2023, from https://www.ISTAT.it/it/files/2021/03/RSU_2-3_2019_Article-4.pdf.

EUROSTAT (2016), *Classification of learning activities (CLA): manual: 2016 edition*, Publications Office, retrieved on January 25, 2023, from https://data.europa.eu/doi/10.2785/874604.

Fraboni R., Gallo F., Palmieri S., Siciliani I. (2019), *Potenzialità e limiti nell'uso dei dati amministrativi per l'analisi dell'inserimento lavorativo dei dottori di ricerca*, ISTAT Working Papers, 12/2019, retrieved on January 25, 2023, from https://www.ISTAT.it/it/files//2019/10/IWP-12-2019.pdf.

ISTAT (2016), *ISTAT's modernization program*, retrieved on January 25, 2023, from https://www.ISTAT.it/it/files/2011/04/IstatsModernistionProgramme_EN.pdf.

Radini R., Scannapieco M., Tosco L. (2018), *The Italian Integrated System of Statistical Registers: Design and Implementation of an Ontology-based Data Integration Architecture*, retrieved on January 25, 2023, from https://www.ISTAT.it/it/files/2018/11/Scannapieco_original-paper.pdf.

Runci M.C., Di Bella G., Cuppone F. (2017), "Integrated Education Microdata to Support Statistics Production", in N.C. Lauro, E. Amaturo, M.G. Grassia, B. Aragona, M. Marino (*eds.*), *Data Science and Social Research. Epistemology, Methods, Technology and Applications*, Springer International Publishing, Studies in Classification, Data Analysis, and Knowledge Organization, Heidelberg.

Signore M., Scanu M., Brancato G. (2015), "Statistical Metadata: a Unified Approach to Management and Dissemination", *Journal of Official Statistics*, 31 (2), pp. 325-347.

UNESCO (2012), *International Standard Classification of Education*. ISCED 2011, retrieved on January 25, 2023, from http://uis.UNESCO.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf.

UNESCO (2014), *ISCED Fields of education and training classification*, ISCED-F 2013, retrieved on January 25, 2023, from http://uis.UNESCO.org/sites/default/files/documents/isced-fields-of-education-and-training-2013-en.pdf.

Wallgren A., Wallgren B. (2014), *Register Based Statistical methods for Administrative Data*, Wiley, New York.

Annex

National register of school students

The National register of school students[10], managed by ME, was established by law in 2005 (and implemented in 2010) for the first and second cycle of education. Then, the register has been enriched with additional contents (pre-primary school, adult education centres and data on the disabled people).

This source constitutes an indispensable information asset for the fulfilment of the institutional tasks of the Ministry and to measure participation in education from pre-primary to the upper-secondary education. According to the ISCED 2011 the register covers all education levels 0.2-3 (0.1 at the moment is excluded).

Each school belonging to the national education and training system (public and accredited private), provides, for each student, individual data relating to the entire education and training career (identification code; personal data; attendance; transfer; interruption; course typology and characteristics; final exams of each cycle and qualification exams).

Data relating to students who choose parental education are also included in the register.

Surveys on schools

The statistical surveys on schools, managed by ISTAT since the end of the '90s and then by ME, are surveys carried out annually on all public and accredited private schools, to integrate the information in the National register of school students with additional general and aggregated data specific for each school[11] (i.e. number of classes, pupils by grade, disabled pupils, full-time/part-time pupils, education outcomes, average number of pupils per class).

The surveys contribute to the creation of an essential information base for monitoring the school education system and for defining school policies (levels 0.2-3 of ISCED 2011).

[10] https://www.miur.gov.it/anagrafe-nazionale-studenti.

[11] https://miur.gov.it/rilevazioni-scuole.

School buildings register

The School buildings register, managed by ME, was established by law in 1996[12] in order to evaluate the amount, status and functionality of the school building stock.

Each state school provides for each school building i) identification code and main general characteristics; ii) data on certifications and documentation; iii) dimensional data; iv) data on qualitative subjective evaluations about building itself, support infrastructure, environmental conditions.

A large set of data is available online and offers an overview on building status of State schools, but also on infrastructure and environments nearby schools (levels 0.2-3 of ISCED 2011).

National surveys on students learning

INVALSI carries out every year a set of national surveys on Italian, Mathematics and English skills[13]. These surveys are carried out in the second and fifth grade of primary school, in the third class of lower secondary school (eighth grade), in the second and fifth class of upper secondary school (tenth and thirteenth grade).

They are based on a centralised scoring systems to make the data directly comparable. A standardised test (same test, same time available, same correction protocol) is submitted to all students of an age cohort in order to assess some fundamental skills. They represent one of many components of a school self-assessment and the individual skills certification does not replace the teachers' assessment.

The database contains information from various sources, including administrative data from schools, students' skills from the standardised tests, and information requested directly from students through questionnaires.

This information completes the framework on school attendance and performance for the first and second cycle of education with objective information on the quality of learning (level 1-3 of ISCED 2011).

[12] https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1996;23; https://www.istruzione.it/edilizia_scolastica/anagrafe.shtml.

[13] https://INVALSI-areaprove.cineca.it/index.php?get=static&pag=rapporti_INVALSI; https://www.gazzettaufficiale.it/eli/id/2017/05/16/17G00070/sg; https://INVALSI-areaprove.cineca.it/docs/2018/INVALSI_tests_according_to_INVALSI.pdf.

Education and vocational training (IeFP)

Education and vocational training (IeFP) represents one of the channels, in addition to the upper secondary school and apprenticeship, valid to complete the compulsory education cycle (level 353 of ISCED 2011). IeFP courses are offered by vocational upper secondary schools and institutions accredited by Regions.

For the vocational upper secondary education without direct access to first tertiary programs, data are available only partially in the National Register of school students (only courses provided by school institutions). Other information (for courses provided by other institutions) are available only at aggregated level and provided by INAPP[14].

In fact, every year, INAPP carries out a survey on IeFP courses to obtain quantitative and qualitative information, both for three and four-year courses, on training offer (courses), participation (enrolment), results (qualification and graduation) by type of program, professional figure, age, gender, citizenship and other individual characteristics.

Moreover, ANPAL is developing an information system for professional training in order to collect from the providers all the information relating to courses and participants.

Higher Technical Institutes database

INDIRE, appointed by ME, has created and manages the database of the Higher Technical Institutes (ITS) and carries out annual monitoring of the ITS performance[15] (level 554 of ISCED 2011). The database, available since 2015, provides, for all offered programmes, since their institution in 2010, information on participants and their performance.

Yearly information on job opportunities is also provided.

National register of students and graduates

The national register of students and graduates[16], managed by MUR, is an administrative database created with the aim of recording all university

[14] https://INAPP.org/it/dati/iefp.

[15] https://www.INDIRE.it/progetto/its-istituti-tecnici-superiori/monitoraggio-nazionale/.

[16] https://www.miur.gov.it/anagrafe-nazionale-degli-studenti-e-dei-laureati.

students enrolled in degree and post-degree programs (levels 6-8 of ISCED 2011). It was established by law in 2003 and implemented from 2004.

For the first time a database on the complete careers of students was created in order to support national and international monitoring and evaluation of the tertiary education system.

Through electronic transmission to MUR, all career events, from enrolment to graduation, including contributions paid and any course change, are stored in the register. The database offers a complete overview of students for all Italian public and private universities (included online universities).

Survey on higher education in fine arts, drama, dance and music institutions (AFAM)

The AFAM system is defined as tertiary level higher specialised education and is parallel to the university system[17] (levels 6-8 of ISCED 2011).

Data on AFAM system are collected yearly by MUR with a statistical census survey conducted on all AFAM institutions. The survey collects information on training programs, students enrolled and graduated, international mobility of students and teachers, contributions paid and transfers.

School, university and AFAM staff databases

ME also provides administrative data on School staff (teaching staff, school principals and non-teaching staff). A large set of variables is available for each group (i.e. individual characteristics, type of school, teaching subject or tasks, contract type, start/end date of current job).

In the same way, MUR provides information on permanent and fixed-term staff in state and non-state universities.

Open source data are available on ME and MUR website[18].

For AFAM Institutions data on staff are derived from the survey already described.

[17] Law 21/12/1999 n. 508.

[18] https://dati.istruzione.it/opendata/opendata/catalogo/#Scuola; http://ustat.miur.it/opendata/.

Further insights

Additional information on participation and achievements on education is derived from the large number of surveys conducted by ISTAT, which allow further analysing the phenomenon broadening the perspective to non-formal and informal education. In fact, the surveys cover aspects not derivable from administrative sources and allow joint analysis of variables collected separately by different sources.

European Adult Education Survey (AES)

AES covers adult (18-74) participation in formal, non-formal and informal education and training according to CLA. It is one of the main data sources for EU lifelong learning statistics[19]. The survey was carried out in 2007, 2011 and 2016 (2017 in Italy) and provides comparable-across-countries data: characteristics of the learning activities; possible obstacles to participation; access to information on learning possibilities; financing and costs of learning; self-reported language skills.

European Labour Force Survey (LFS)

LFS is a large household sample survey providing quarterly results on labour participation of people aged 15 and over as well as on people outside the labour force living in private households[20].

The survey also collects a large set of information on education and training, functional to the analysis of labour outcomes (educational attainment, work experience at workplace as part of the educational program, participation in formal e non-formal education and training)[21].

[19] https://ec.europa.eu/EUROSTAT/web/microdata/adult-education-survey; https://www.ISTAT.it/it/files//2018/12/Report_Partecipazione_adulti.pdf.

[20] https://ec.europa.eu/EUROSTAT/web/microdata/european-union-labour-force-survey.

[21] https://www.ISTAT.it/it/archivio/8263.

Survey-system on the education-to-work-transition

The survey-system on the education-to-work-transition completes the picture of sources useful to analyse the performance of different levels and fields of education attainments on labour market and, in general, to evaluate the quality of the higher education system (levels 3-8 of ISCED 2011).

It includes three different surveys: i) the Survey on upper secondary school graduates' transition to university and labour market[22]; ii) the Survey on education-to-work-transition of university graduates[23] and iii) the Survey on the education-to-work-transition of PhD students[24].

These three surveys provide a great extent of information: educational experience; access to the labour market; skill mismatch; job search; family situation, referring both to the family of origin and to the current one at the time of the interview.

Survey on school services for early childhood (0-2)

The Survey on school services for early childhood (0-2) provides information on a segment of population not covered by administrative data (levels 0.1 of ISCED 2011)[25]. The survey is carried out by ISTAT in collaboration with the State general accounting office, the Ministry of economy and finance and territorial Institutions.

The analysis units are the Municipalities and the other Institutions or private associations that provide childhood services.

The questionnaire provides information on the side of the public and private supply of services: number of structures; number of available places; Municipalities, association and households contribution to expenditures; number of users in the educational year; type of administration.

[22] https://www.ISTAT.it/it/archivio/8392.

[23] https://www.ISTAT.it/it/archivio/8338.

[24] https://www.ISTAT.it/it/archivio/8555.

[25] https://www.ISTAT.it/it/files/2020/10/REPORT_ASILI-NIDO-2018-19.pdf; https://www.ISTAT.it/it/archivio/7566.

Household Budget Survey (HBS) and European Income and Living Conditions Survey (Eu-silc)

HBS[26] and Eu-silc[27] allow to widen the information landscape with data on households' expenditure and income and to investigate the relationship between education and the risk of poverty and social exclusion.

Specific information is also available in HBS on expenditure for formal education by ISCED 2011 level (tuition fees, textbooks, stationery and drawing school materials, canteens in schools, universities and other educational establishments, school journeys). Expenditures for non-formal education and training are also covered.

On the other side, Eu-Silc investigates the household difficulties in affording education costs.

Multipurpose Survey: Aspects of daily life

The Multipurpose Survey: Aspects of daily life[28] provides indicators on school services accessibility, participation in non-formal education and the reasons for not attending a pre-primary school in early childhood. Starting from 2020 it also includes a module on online lessons during restrictions due to Covid-19 pandemic.

[26] https://ec.europa.eu/EUROSTAT/web/microdata/household-budget-survey; https://www.ISTAT.it/it/archivio/71980.

[27] https://ec.europa.eu/EUROSTAT/web/microdata/european-union-statistics-on-income-and-living-conditions; https://www.ISTAT.it/it/archivio/5663.

[28] https://www.ISTAT.it/it/archivio/91926.

2. A new alliance: matching data and needs of local authorities

by Lorenzo Maraviglia

Many crucial decisions about the articulation and concrete composition of educational supply are taken by local authorities (municipalities, conferences of municipalities, provinces).

However, these institutions often lack the necessary data to make informed choices. This poses a sensible issue, since decisions taken at the local level are largely the result of a negotiation process involving a multitude of heterogeneous interests. Such problems are likely to exacerbate with the decline in number of students that is expected over the next few years, as schools will find themselves competing fiercely for a diminishing amount of resources.

In light of these considerations, the contribution analyzes in detail the mismatch between the quantity and quality of the data made available by ISTAT, INVALSI and MIUR and the actual cognitive needs expressed by local authorities.

The issue is illustrated through the presentation of concrete cases, drawn from the multi-year experience of support for school planning carried out by the Statistics Office of the Province of Lucca. In particular, the solutions adopted by the Office through the integration of ISTAT microdata, INVALSI microdata and aggregated MIUR data are illustrated and discussed.

On the basis of the analysis, operational proposals are finally made for a better integration of the available data sources and for their greater adherence to the cognitive needs expressed by local institutions.

Nell'ambito della cornice delimitata dalle norme nazionali e regionali, molte decisioni cruciali in ordine all'articolazione ed alla composizione dell'offerta scolastica sono assunte da autorità locali (comuni, conferenze di comuni, province).

Tali istituzioni, tuttavia, non dispongono dei dati necessari per compiere scelte informate e consapevoli su vari aspetti. Ciò pone una questione particolarmente delicata, dal momento che le decisioni prese a livello locale sono in larga misura il risultato di un processo negoziale che coinvolge una moltitudine di interessi eterogenei. Questi problemi sono destinati ad acuirsi a fronte della diminuzione degli studenti attesa per i prossimi anni, quando le scuole si troveranno a competere aspramente per una quantità decrescente di risorse.

Alla luce di tali considerazioni, il contributo analizza in dettaglio il mismatch fra quantità e qualità dei dati resi disponibili da ISTAT, INVALSI e MIUR ed effettive esigenze conoscitive degli enti locali preposti alla programmazione di dettaglio della rete scolastica e dell'offerta formativa.

La problematica in oggetto viene illustrata attraverso la presentazione di casi concreti, tratti dall'esperienza pluriennale di supporto alla programmazione scolastica svolta dall'Ufficio di Statistica della Provincia di Lucca. In particolare, sono illustrate e discusse le soluzioni adottate dall'Ufficio attraverso l'integrazione di microdati ISTAT, microdati INVALSI e dati aggregati MIUR.

Sulla base dell'analisi sono infine avanzate proposte operative per una migliore integrazione delle fonti di dati disponibili e per una loro maggior aderenza ai bisogni conoscitivi espressi dalle istituzioni locali.

1. Introduction

Many crucial decisions concerning the provision of educational services are taken by local authorities. For example, municipalities are required to fix the amount of full-time classes in infant and primary schools; provincial authorities[1], on their part, supervise the introduction of new curricula in upper secondary schools. Just like the central government, nowadays the local level has to face some fundamental challenges that affect the very essence of the Italian education system. First of all, there is the expected decrease in the number of students following the collapse in births of the last decade. Secondly, schools must face the radical change in the cultural composition and expectations of their users that derives from the increase in the share of students with a migrant background; it is important to underline that immigration is a multiplier of complexity, since it introduces elements

[1] The functions of provincial authorities (province) have been redesigned by the law 56/2014 ("Legge Delrio"). For an updated analysis, see UPI (2021).

that are not only different from those of the indigenous component but also profoundly heterogeneous among themselves. Thirdly, there are the transformations induced by the digital transition, which require a modernization and an adjustment of the educational provision that struggle to make headway. Furthermore, these challenges emerge in a context in which, due to the aging of the population and the high spending expectations of the adult, there are tendencies to limit the resources allocated to schools to the advantage, for example, of those allocated to pensions and healthcare.

In order to face these challenges local authorities need both material and immaterial resources, such as knowledge and data. The latter are particularly scarce, due to limitations in the supply of territorial statistics and insufficient development of integrated and accessible databases (ISTAT, 2010). This is a stimulating challenge for (the few) statisticians and data analysts working within, or on behalf of, local organizations. As we will show, some kind of "patchworking" attitude is needed to produce a picture that can support plausible planning efforts – where "plausible" means accounting for future problems in a realistic manner.

2. The basic issue: how many students, and what kind, in the next future

At the essence, planning deals with trying to anticipate some aspects of the future. In the context of education, the basic forecast is about the number and the type of students who will attend schools in the next years. Local authorities tent to adopt a short-to-medium time horizon, despite some decisions – for example, whether a new school ought to be built – have long-lasting consequences. Nonetheless, explicit reference to demographic projections is almost completely absent from local plans.

Insight about the demographic future of a local areas such as "province" (counties) or "comuni" (municipalities) can be drawn from population forecasts released by ISTAT (http://demo.ISTAT.it).

Fig. 1 – ISTAT Demographic web page (content licensed under CC-BY standard; link to demographic projections in the bottom-right corner)

ISTAT official projections don't go below the regional (NUTS2) level (ISTAT, 2021). This is a relevant drawback since in a country like Italy sub-regional variation is potentially wide; in this regard, it is possible that the unavailability of more detailed geographical forecasts plays a role in the underrepresentation of demographic issues within the local planning debate.

In our opinion, however, rather than territorial coverage the flaws of available population projections deal more with other aspects, such us the lack of breakdown by citizenship (foreign students are a growing share of Italian school population) and an overly "optimistic" flavour.

Unfortunately, demographic trends have proved much worse than expected since year 2018, when ISTAT projections were released[2]. Leaving aside the rise in mortality caused by the Covid-19 pandemics, births have fallen in all territories. To make an example, in the province of Lucca[3] the amount of yearly births has shrinked of almost 40% compared to recent (2006-2008) peak years (Fig. 2).

[2] Revisions are under way and should be released in the next months.

[3] In the following, we refer to the case of the province of Lucca to illustrate our argument; Lucca is a medium-sized (according to Italian standards) province in the Center of Italy.

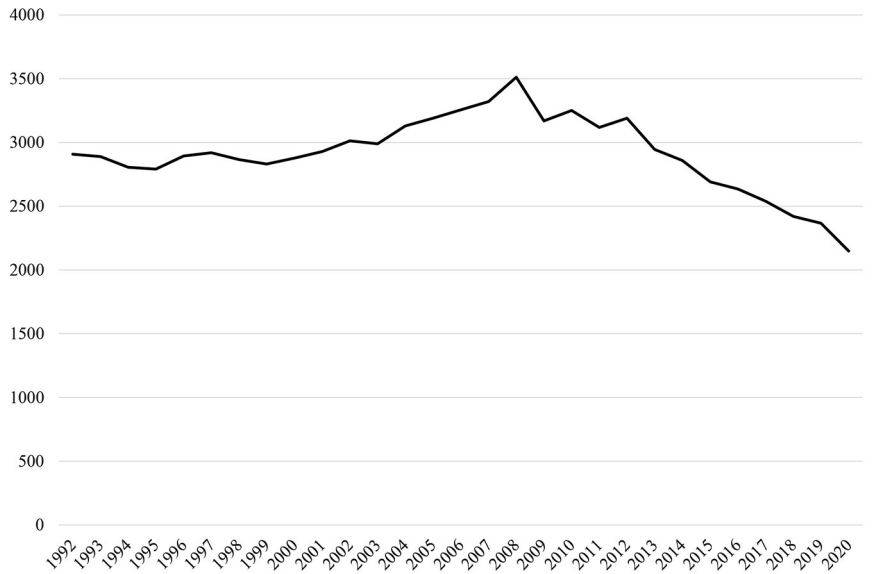


Fig 2 – Time series of number of births in the province of Lucca (1992-2020)

Source: our elaboration on ISTAT data

On the other side migration has cooled down, especially the kind of family immigration that had boosted births during the first decade of the 21st century.

Data about both births and migration flows for every single municipality can be extracted from ISTAT online demographic data base. More detailed information can be achieved by local administrations statistical officers through the SISTAN (Sistema Statistico Nazionale) protocol, which allows direct access to ISTAT elementary (micro-) data.

Drawing on this source, we have estimated the migration balances by five-year age group for the population of the province of Lucca (fig. 3). It can be noted that in recent years the age balance for the younger classes (light grey bars on the left) tends to zero.

Neutrality of migration implies that, in the next few years, the size of population attending local schools will depend mainly on previous birth trends – which are known, both for the Italian and for the foreign component.

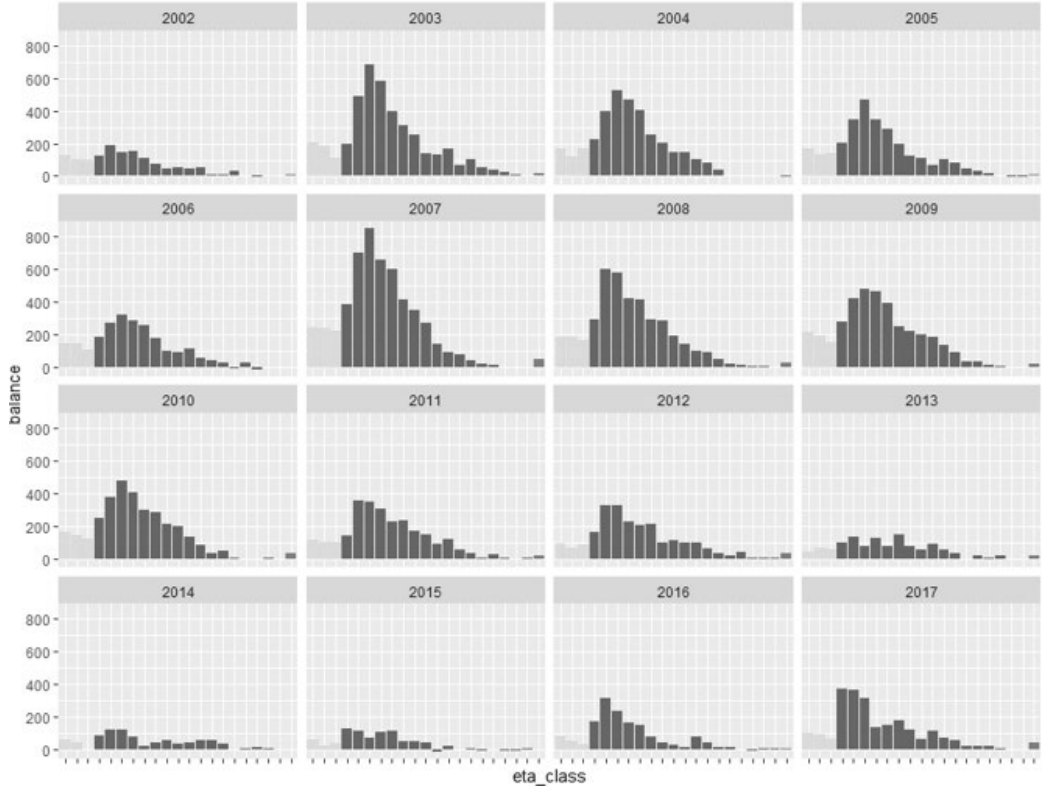


Fig. 3 – Time series of migration balances in provincia di Lucca (2002-2017)

Source: our elaboration on ISTAT data

This allows quite accurate projections, that can be used to plan future levels of service provision.

Tab. 1 – Projections for population aged 14-19 living in province of Lucca

		2021		*2026*		*2031*		*2035*	
		N.	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*
Piana	Italian	6,976	89.6	6,840	88.0	5,863	84.1	4,835	82.4
	Foreign	806	10.4	936	12.0	1,106	15.9	1,032	17.6
	Total	7,782	100.0	7,776	100.0	6,969	100.0	5,867	100.0
Versilia	Italian	6,393	94.4	6,216	93.4	5,104	91.4	4,293	89.5
	Foreign	378	5.6	436	6.6	482	8.6	502	10.5
	Total	6,771	100.0	6,652	100.0	5,586	100.0	4,795	100.0
Valle del Serchio	Italian	2,080	92.3	2060	92.2	1,676	89.6	1,287	87.0
	Foreign	173	7.7	175	7.8	194	10.4	193	13.0
	Total	2,253	100.0	2,235	100.0	1,870	100.0	1,480	100.0
Provincia	Italian	15,449	91.9	15,116	90.7	12,643	87.6	10,415	85.8
	Foreign	1,357	8.1	1,547	9.3	1,782	12.4	1,727	14.2
	Total	16,806	100.0	16,663	100.0	14,425	100.0	12,142	100.0

		2026-2021 (difference)		*2031-2021 (difference)*		*2035-2021 (difference)*	
		N.	*%*	*N.*	*%*	*N.*	*%*
Piana	Italian	-136	-1.9	-1,113	-16.0	-2,141	-30.7
	Foreign	+130	+16.1	+300	+37.2	+226	+28.0
	Total	-6	-0.1	-813	-10.4	-1,915	-24.6
Versilia	Italian	-177	-2.8	-1,289	-20.2	-2,100	-32.8
	Foreign	+58	+15.3	+104	+27.5	+124	+32.8
	Total	-119	-1.8	-1,185	-17.9	-1,976	-29.2
Valle del Serchio	Italian	-20	-2.2	-404	-19.4	-793	-38.1
	Foreign	+2	+1.2	+21	+12.1	+20	+11.6
	Total	-18	-0.8	-383	-17.0	-773	-34.3
Provincia	Italian	-333	-2.2	-2,806	-18.2	-5,034	-32.6
	Foreign	+190	+14.0	+425	+31.3	+370	+27.3
	Total	-143	-0.9	-2.381	-14.2	-4,664	-27.8

Source: our elaboration

In Table 1 we show our projections for upper-secondary school-aged[4] population living in the province of Lucca in the next five, ten and fifteen years. In order to get a more informative picture, we break down overall

[4] Population aged between 14 and 18 years old.

projections by sub-provincial area (Piana, Versilia, Valle del Serchio) and citizenship (Italians, Foreign).

Even taking uncertainty into account[5] one cannot help but being struck by the amount of projected decline (about 28%) by the year 2035[6].

Furthermore, what is particularly worrying is the timing of the expected fall; in fact, the number of students will remain quite stable for the next five years, only to collapse very quickly thereafter (Tab. 1). These are bad news because when things are (apparently) stable, hard decisions tent to be delayed; with the risk of having no time left, in the next future, to tackle the challenges posed by the change.

3. Going deeper: citizenship composition of projected school population

Our projections suggest that while the overall number of students will shrink, both amount and share of foreign students is expected to increase. Obviously planning choices about the provision of school services must take into account this scenario. Foreign students are different under many respects from their Italian mates. Furthermore, they are not distributed evenly among schools.

Table 2 shows data on foreign students' attendance in upper secondary schools in Lucca province. These data have been downloaded from Italian Ministery of Education (MI) website and are supplied with an open licence for all territories (province, comuni)[7]. The share of foreign students ranges from 2% to almost 20%. Vocational and technical schools are ranked high in this ordering, whereas "licei", especially "classici" and "scientifici", tend to occupy the bottom positions. This apparent segregation is an issue in itself. Foreign students are highly under-represented in certain type of secondary schools[8]. But heterogeneities emerge even at lower educational levels.

[5] Uncertainty is not reported in the table; as a general rule, uncertainty in projections increases with time but for the reasons mentioned above (decreasing impact of migration flows) is limited in this context.

[6] Which soars up to about 33% for Italian students.

[7] This is important because allows comparisons among territories.

[8] A similar pattern of segregation affects students with disabilities for which, however, no open data are supplied by MIUR.

Tab. 2 – Percentage of foreign students in upper secundary school in Lucca province, year 2020

School name	*Foreign (n)*	*Italian (n)*	*All (n)*	*Foreign (%)*
Omitted	46	201	247	18.6
Omitted	75	332	407	18.4
Omitted	90	433	523	17.2
Omitted	15	77	92	16.3
Omitted	54	288	342	15.8
Omitted	53	329	382	13.9
…	…	…	…	…
Omitted	22	705	727	3.03
Omitted	6	203	209	2.87
Omitted	7	244	251	2.79
Omitted	4	148	152	2.63
Omitted	9	364	373	2.41
Omitted	8	333	341	2.35

Source: our elaboration in MIUR data

In Table 3, foreign students' percentages are supplied for all primary schools of a single municipality of the province of di Lucca[9].

Tab. 3 – Percentage of foreign students in primary schools in the municipality of Lucca, year 2020

School name	*Foreign (n)*	*Italian (n)*	*All (n)*	*Foreign (%)*
Omitted	46	201	247	22.0
Omitted	75	332	407	14.6
Omitted	90	433	523	12.9
Omitted	15	77	92	12.3
Omitted	54	288	342	11.2
Omitted	53	329	382	10.8
Omitted	71	445	516	4.0

Source: our elaboration on MIUR data

Unbalance is even more striking, since we are before the tracking process which takes place when students move to secondary education. Here disparities depend mainly on residential choices of foreign families and on the attitudes of certain communities to locate in some specific areas.

[9] We omit the name of the municipality for confidentiality reasons.

However this poses challenges that are bound to intensify in a scenario or quick demographic, social and economical change.

Such issues and risks are hardly perceived at local level; or, better, decision makers can have an intuition, but they lack the details and a clear comparative picture that allow them to plan appropriate policies. That's why data are so important.

4. Delving into nationality: the INVALSI data

As we have shown, many relevant aspects of the picture that we are trying to sketch can be obtained from accessible sources such as ISTAT or MIUR open data. In the remaining we will discuss the possibility to access micro-data, since existing aggregated figures are not always consistent with needs of local decision makers.

Data are not relevant in themselves; they matter to the extent that contain the type of information that is needed to take proper decisions.

Let's take the challenges of a future school where the number of students shrink while, at the same time, the share and heterogeneity of foreign students is set to increase.

At the pragmatic level at which problems must be tackled, being aware of the amount and distribution of "foreign" students is not enough. On average, a Romanian student is different from a Moroccan or Chinese one. Asians don't speak Italian at home. Some Romanian or Albanian families do it. This has huge implications.

Tab. 4 – Languages spoken by students: sample from INVALSI micro-data

School	*Italian*	*Albanian*	*Arab*	*...*	*Else*	*Foreign Lang (rate)*
Omitted	29	0	1	…	1	14.3
Omitted	36	1	0	…	1	13.0
Omitted	89	3	1	…	2	9.2
Omitted	10	0	0	…	1	13.3
Omitted	27	0	0	…	0	7.0
Omitted	59	1	0	…	1	6.0
Omitted	87	1	9	…	5	17.9
Omitted	42	6	0	…	2	17.6
Omitted	52	6	1	…	0	17.6
…	…	…	…	…	…	…

Source: our elaboration on INVALSI data

In order to make informed choices, local decision makers should know "who" are the foreign students attending the schools of their territories. Obviously, nationality is not the only source of heterogeneity among foreign students; but it is a powerful vector of information; in addition, data about nationality are ordinarily recorded and can be easily extracted by institutional databases.

INVALSI micro-data are one of the few sources that local authorities can draw on to obtain information about the national composition of students attending schools in a specific area. Most importantly, in INVALSI micro-data information on the national background[10] is recorded together with that on skills and academic performance, on educational attainment and professional status of parents, on availability of books and other learning resources. This is the kind of insight that can take local school policies to an upper level, allowing decision makers to make effective allocation choices.

In order to make this happen, cooperation between statistical data providers, such as ISTAT and INVALSI, and local authorities must be fostered. For example, the level of disaggregation at which data are made available should be flexibly identified within an assessment involving both providers and users and not set once for all on an a priori basis (by providers alone).

In many circumstances, local authorities should be granted access to granular information, such as the name of schools which are omitted in Table 4. At this level, the discussion on confidentiality limits to data disclosure should be reset so as to include the real effectiveness issues with which local policies are confronted daily; by local policy we mean an action that must necessarily address targets taken in their individual complexity – a policy that cannot limit itself to the administration of standardized measures[11].

5. Discussion

We have sketched a picture aimed at identifying data from different sources that can be used to support the delicate decisions that, in the next future, local administrations will have to take regarding the offer of school services.

[10] Nationality is not analytically recorded in INVALSI micro-data (with regard to citizenship, students are classified as Italians, first-generation foreigners and second-generation foreigners); however, detailed information is provided about which language is spoken at home.

[11] We are not saying that, for a well-determined local authority, it is impossible to obtain such data from ISTAT or INVALSI: sometimes it is possible, maybe as the result of a tiring and time-consuming confrontation; we want to highlight the more general features of the issue, the need for a cooperation which is informed by a transparent comprehension of interests and point of views of all actors involved.

For each source we have discussed the potential content and availability issues.

There are a few points that we would like to further bring to attention. First one is the absolute need for a comparative picture: every territory must be entitled to access data of all the others. The reason is that, in order to appropriately understand what's going on, comparisons are needed. One cannot learn from itself only: planning is not a matter of introspection, but rather involves a hard knowledge effort aimed at picking general and deep changes in our society – that obviously must be scaled at a local level but involve a wider context. On the other side, local decision makers must be able to connect the relevant information – the one that is required to achieve efficiency and avoid "unnecessary errors" (Arrow, 1974) – to concrete actors; they must be allowed to identify the targets of their intervention – which school will receive what – without the burden and costs to re-collect data (for example about nationality of students) that have already been collected by other authorities; but are not made available due to a restrictive interpretation of privacy rules. This is not to deny the relevance of privacy issues: they are obviously important and worth of maximum protection. But privacy must be understood in the context of its implications for equally vital goals, moreover pursued by other public authorities. To this regard, our impression is that challenges and needs faced by local authorities in the field of planning of educational services are not totally understood by those who have to take decisions about what data must be made available and for what reasons. With these reflections we try to make a contribution to the discussion.

References

Arrow J.K. (1974), *The Limits of Organization*, Norton & Co., New York.

ISTAT (2010), *Navigando fra le fonti demografiche e sociali*, Roma.

ISTAT (2021), *Previsioni della popolazione residente e delle famiglie*, retrieved on Juanary 25, 2023, from https://www.ISTAT.it/it/archivio/263995.

UPI (2021), *Le priorità delle province per il nuovo governo*, retrieved on Juanary 25, 2023, from https://www.upinet.it.

3. The INVALSI tests for teaching: teachers' beliefs about the validity of INVALSI tests and teaching practices implemented

by Elisa Truffelli, Barbara Balconi, Daniela Maccario*

This chapter originates from an interdisciplinary research project involving experts in the didactics of mathematics, general didactics and assessment, with the aim of investigating the beliefs and attitudes of primary school teachers on the connection between the INVALSI mathematics tests and teaching-learning processes.

The survey was administered in the 2019/2020 school year in the form of a semi-structured questionnaire aimed at investigating: teachers' awareness of the learning objectives detected by the INVALSI tests, their conceptions of errors in mathematics, the use of the tests in daily didactics, misconceptions about standardised tests and the concept of assessment. The tool was administered to a non-representative national sample consisting of 526 fifth-grade mathematics teachers at primary schools.

The analyses presented focused on the one hand on the link between teachers' beliefs on the type of knowledge and skills detected by the INVALSI tests and on the validity of the tests themselves, and on the other hand on the convictions and statements regarding the teaching practices usually implemented by the teachers. The interpretation of the links that emerged on a correlational basis, between the two blocks of variables, showed how teachers' beliefs influenced the direction of their didactic choices. Teachers who score high on the considered indices and factors perceive themselves as professionals and believe that learning can and should be constructed. Recognising the validity of the INVALSI tests in relation to the skills they demand of students on the one hand, and their usefulness for analysing, reflecting and planning didactics on the other, affect the type of prevalent use they make

* The chapter was conceived and designed jointly by the authors. Paragraph 1 and 7 are to be attributed to Daniela Maccario; Barbara Balconi wrote the paragraphs 2 and 3, Elisa Truffelli wrote the paragraphs 4, 5 and 6.

of the test in the classroom context: a use aimed mainly at developing skills rather than at teaching-to-test and aimed at improving the didactic proposal to students to help them strengthen their skills.

Il presente contributo nasce all'interno di un progetto di ricerca interdisciplinare, che ha coinvolto esperti di didattica della matematica, di didattica generale e valutazione, con l'obiettivo di indagare convinzioni e atteggiamenti degli insegnanti di scuola primaria sul legame tra i test INVALSI di matematica e i processi di insegnamento-apprendimento.

La rilevazione è avvenuta nell'anno scolastico 2019/20 attraverso un questionario semi strutturato volto ad indagare: la consapevolezza dei docenti in merito agli obiettivi di apprendimento rilevati dai test INVALSI, le loro concezioni sugli errori in matematica, l'uso dei test nella didattica quotidiana, le misconcezioni in merito alle prove standardizzate e l'idea di valutazione. Lo strumento è stato somministrato ad un campione nazionale non rappresentativo costituito da 526 docenti di matematica di classe quinta di scuola primaria.

Le analisi presentate si sono focalizzate da un lato sul legame tra convinzioni dei docenti sul tipo di conoscenze e abilità rilevate dai test INVALSI e sulla validità dei test stessi e dall'altro sulle convinzioni e dichiarazioni in merito alle pratiche didattiche messe in atto abitualmente dagli insegnanti. L'interpretazione dei legami emersi su base correlazionale tra i due blocchi di variabili ha evidenziato come le convinzioni degli insegnanti abbiano influenzato la direzione delle loro scelte didattiche. Gli insegnanti che hanno punteggi alti sugli indici e fattori considerati, si percepiscono come professionisti e credono che l'apprendimento possa e debba essere costruito. Riconoscere la validità dei test INVALSI in relazione alle competenze che sollecitano negli alunni da un lato, e l'utilità dello stesso per analizzare, riflettere e progettare la didattica dall'altro, incidono sul tipo di uso prevalente che essi fanno del test nel contesto di classe: un uso volto principalmente allo sviluppo di competenze piuttosto che al cosiddetto "teaching to test" e orientato a migliorare e adattare la proposta didattica agli alunni, per aiutarli a potenziare le loro competenze.

1. Theoretical and epistemological framework

Can INVALSI tests be a useful tool to support teachers' reflexivity on how to improve the quality of mathematics teaching in primary schools?

Can educational research answer this question? If so, through which approaches? On the basis of these questions, a study was developed to collect

and analyse mathematics teaching practices concerning the administration of national tests in primary schools. This research involved the development of an original tool to study the teaching of mathematics with regard to the INVALSI test framework. The project is part of an interpretation of pedagogical-didactic research as a scientific field that aims to increase the wealth of knowledge available on teaching action in order to guide teachers' practical-operational decisions. From this point of view, teachers are considered professionals whose job is to stimulate, guide, and support students' learning towards the acquisition of skills or rather deep, meaningful and useful forms of learning. The main levers of professional intervention by teachers are the selection, organisation and presentation of concepts and logics inherent to the study disciplines through the implementation of teaching activities tailored to the pupils' and classes' circumstances. The research presented below responds to a vision whereby educational research is developed as a process that surveys classroom management practice by involving the protagonists. It is then analysed and reflected upon with a view to enriching the existing theory in a key way that is beneficial to improving the practice itself (Zanniello, 2023) by means of professional reflection and decision-making processes of teachers. Didactic research is thus committed to building evidence that opens interpretative paths while adhering to the experience of teachers' classroom management in order to construct useful knowledge for their professionalisation. Let us analytically discuss the salient aspects of this theoretical-epistemological standpoint.

1.1. Studying didactic mediation and the "de facto curriculum"

Teaching action in terms of how a teacher teaches mathematics to support student learning with regards the administration of INVALSI tests has been identified as the object of investigation. Teaching action is considered in terms of "didactic mediation", from a processual and interactive perspective. Reference is made to a conception of knowledge and learning of a basically Piagetian and Brunerian constructivist matrix (Damiano, 2013), with Vygotskian references (Lenoir, 2017), according to which the indispensable condition for learning to take place is to allow the subject to consciously intervene in the situations he or she experiences. This dynamic is developed at school through the mediation of the teacher, who assumes the role of learning facilitator when, by activating appropriate forms of communicative exchange with pupils, he/she offers them the opportunity to interact with the teaching situation – in a mental and/or practical sense

– and to reflect on what they are experiencing, thus triggering a process of structuring and restructuring their knowledge. Teachers' actions can also be considered "medial" insofar as they are operationally translated into the production and management of "mediators" or representations of reality in terms of the various disciplinary knowledge to be taught. It has been observed that this "metaphorisation" activity (Damiano, 2013) can take place through the implementation of a combination of languages: active, based on direct and operational experience; analogical, based on forms of simulation, also of a ludic nature; iconic, based on used and produced images, fixed and moving, analogical and digital etc.; symbolic, through combinations of sign-formalisms with the power to evoke objects and phenomena in abstract and synthetic terms, to foster the construction of knowledge without having to resort to direct experience. In the context of this study, an approach was tested to try and "bring to light" some salient aspects of the phenomenology of the didactic mediation processes used by teachers for the administration of INVALSI tests, and thus to intercept aspects of the *de facto curriculum* (Perrenoud, 1993) generated in relation to the "Rilevazioni nazionali degli apprendimenti scolastici" (Italian National Surveys of School Learning). These are processes that are little known even by the teachers themselves, at least in an explicit and formalised perspective, the analysis of which could prompt reflection on how to improve teaching so as to promote increasingly effective forms of learning.

1.2. Studying teachers' processes of interpretation and action

The survey and analysis of the phenomenology of the teaching processes generated by the administration of National Tests was also conducted with reference to a conception of teaching as an activity of a professional nature, in line with interpretations that have been widespread in literature since the 1980s. It is believed that teachers can be considered professionals insofar as their work essentially consists of performing non-routine intellectual acts in pursuit of objectives in complex situations, in which they operate with a significant degree of autonomy and responsibility (Damiano, 2004), from a personal and educational background made up of multiple knowledge. (Paquay *et al.*, 2006; Tardif and Lessard, 2004). In their work, professional teachers draw on a multiplicity of interpretative frameworks, both of a theoretical-general nature and derived from experience; from these resources, they analyse the problems to be addressed in the classroom and possible solution strategies (Altet, 2010), in a kind of dialogue with the sit-

uation that passes through action – "reflection-in-action" – (Schön, 1983). This process involves the recognition, revision and development of one's theoretical-conceptual and operational models (Vergnaud, 2010), or habits of action (Barbier, 2017). According to this perspective, the theoretical-didactic criteriology on mathematics teaching implied in INVALSI tests can act as a framework for teachers to attribute meaning to classroom teaching processes "if" and depending on "how" it is used in relation to their professional experience and the concrete situations they habitually face. The research project thus developed around this focus of investigation in order to observe certain junctures in the practical processes of attributing meaning to the National Test framework.

1.3. "How" to study teaching practices: the choice of a "deliberative" approach

Within the theoretical-epistemological framework that connotes this study, the practical-operational knowledge of teachers was considered as a form of knowledge other than the scientific-formalised knowledge on classroom management, but which should be credited as a source for its construction (Viganò, 2019; Damiano, 2006). The literature invites us to pursue this path, suggesting that the study of teachers' practical-experiential knowledge is useful for the detection of previously unexplored problems, for the conceptual and operational definition of variables, for supporting the formulation of increasingly precise cognitive hypotheses (Zanniello, 2020), and for increasing the chances of practical spin-offs of research data (Hattie, 2016). This is a perspective that presents challenges to researchers, as it requires them to accept the confrontation with an "impressive complexity of the object" (Calvani, 2019), which lends itself to being described and analysed at various levels of generality, within different interpretative frameworks, with the use of terms often characterised by a large degree of semantic and pragmatic ambiguity. The "structural" difficulty of didactic research lies in the management and development of theoretical categories whose meanings can be shared between researchers and protagonists in the field (Cardarello, 2019). These concepts can have both a "pragmatic" value – needed to conceptualise teaching processes from the teachers' point of view – and a "pragmatised" value (Pastré, 2007) – which can be linked to the theorisations produced by scientific research on teaching in order to develop them. In this scenario, the study was developed according to a methodological pathway with a "deliberative" connotation (Savoie-Zajc, 2004), starting from predefined interpretative constructs-lenses. These were

validated and operationalised within the research team according to the data collection process, with the goal of systematically reconstructing the phenomenon in question with relatively broad quantitative coverage.

The research pathway is described below, with reference to the actors involved, the various phases, the characteristics of the instrument developed and adopted, the data collected and the results obtained.

2. Presentation of the overall research project design

As discussed in the previous section, the experience of the INVALSI tests has led to the emergence in schools of a series of macro-phenomena linked to the disciplines, to disciplinary teaching and more generally to didactic aspects and the culture of evaluating teachers.

In particular, the INVALSI assessment program has raised and still raises a series of research questions and problems concerning: the reading and interpretation of data, the analysis of teachers' training needs, the analysis of how the two variables listed above can influence teachers' attribution of meaning to the experience of evaluating the tests themselves.

In 2017, an interdisciplinary research project was launched within the "General and Disciplinary Didactics" Observatory of the SIRD (Italian Society of Didactic Research) which highlights the interpretation needs of the complex phenomena mentioned above, with the general purpose of identifying teachers' training needs at national level and proposing guidelines to improve teaching practices regarding the use of INVALSI mathematics tests.

The project was conducted by the research group "Teaching Mathematics and INVALSI Tests"[1] and involved mathematics teachers, experts in general didactics and experimental pedagogy.

The aim of the research was to analyse the link between INVALSI mathematics tests and teaching-learning processes at the Primary School level, through the voice of teachers.

[1] The researchers of the project are: coordinators: Ferdinando Arzarello, Ira Vannini (University of Bologna); members: Giorgio Asquini (La Sapienza University), Barbara Balconi (University of Milan), Giorgio Bolondi (Free University of Bozen-Bolzano), Eleonora Faggiano (University of Bari), Federica Ferretti (University of Ferrara), Violetta Lonati (University of Milan), Daniela Maccario (University of Turin), Annarita Monaco (Teacher, Rome), Ottavio Rizzo (University of Milan), Roberto Trinchero (University of Turin), Elisa Truffelli (University of Bologna), Valentina Vaccaro (INVALSI, Rome).

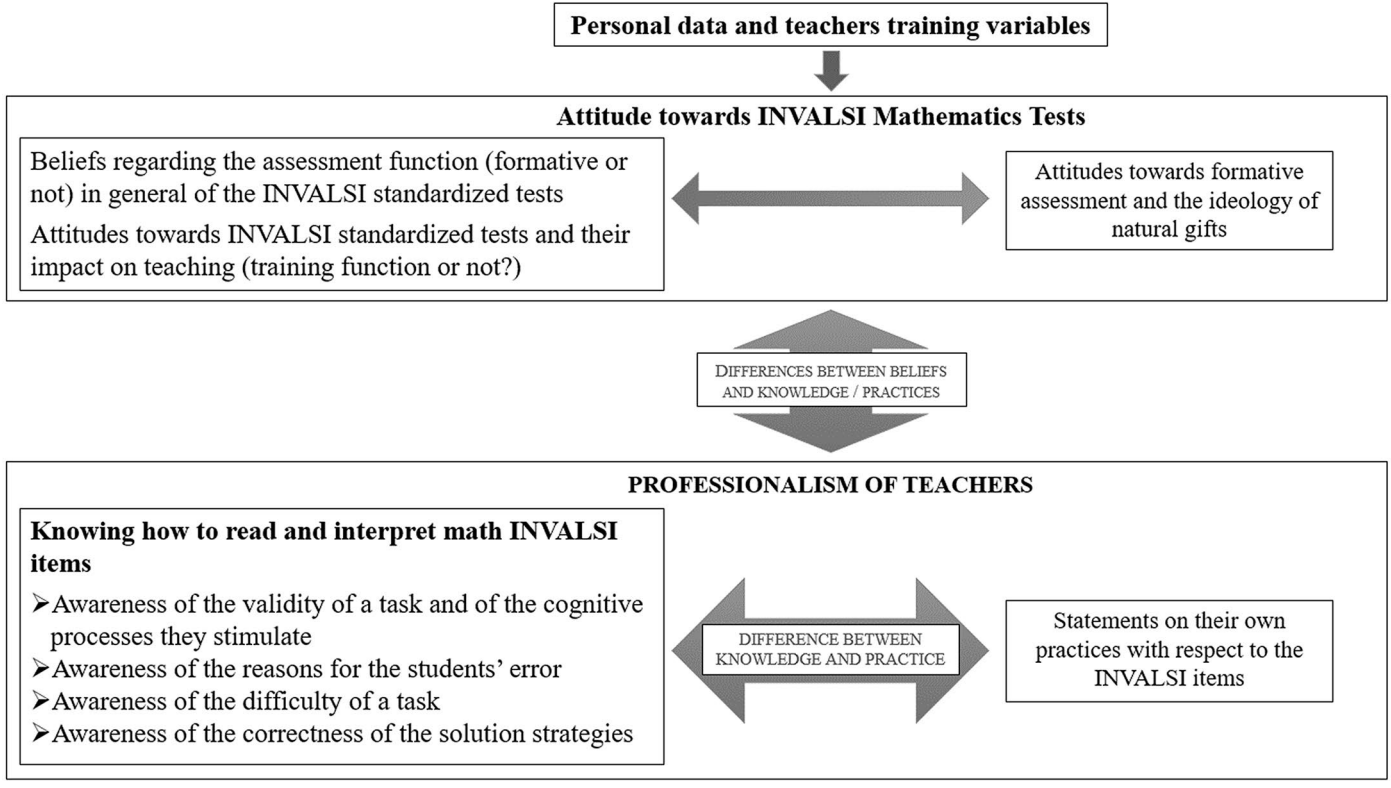


Fig. 1 – Variables and the reseach hypothesis

From the outset, the goal of the interdisciplinary group of researchers was to initiate an exploratory study to investigate the meanings teachers attribute to the INVALSI items, meanings that appear to play a crucial role in promoting or inhibiting an educational impact of the tests themselves. With this aim, the researchers shared the goal of building a research tool, i.e., a questionnaire, for collecting the perceptions of primary school mathematics teachers in relation to the reading and interpretation of the INVALSI questions and the data collected through the tests' administration.

In particular, teachers' beliefs regarding the knowledge and skills detected by the standardised INVALSI tests were investigated; the proximity/distance between the functions and contents of the INVALSI items and the teaching practices of the teachers was explored, the opinions of the teachers were collected in considering the INVALSI test as a possibility to be used in a formative assessment perspective.

To specify the different search variables and the research hypothesis of the relationships between the variables, the framework in Figure 1 was built.

2.2. The investigation tool

The framework of the research variables just described led to the configuration of the questionnaire, in three sections:

– Section1: Interpretation of INVALSI Test (examining how teachers interpret the INVALSI items and their results);
– Section 2: Style of teaching (exploring teachers' beliefs, attitudes, and their involvement in teaching practices);
– Section 3: Personal data and context information.

In the first section, seven INVALSI grade 5 or 6 items are presented in their original formulation. For each of them, questions were proposed to detect:

– the pedagogical knowledge of the mathematical content – the so-called Pedagogical Content Knowledge of Shulman (1986) – by teachers (misunderstandings, recurring errors, level of difficulty);
– the proximity/distance of the seven items from teaching practices and National Indications (2012), a reference document for the didactic planning of school teachers in the first cycle of education;
– the effectiveness of the INVALSI items with respect to the assessment of certain skills.

The second section presents four sets of questions concerning:

1) teachers' opinions on the INVALSI assessment programme;
2) didactic usefulness of the INVALSI items;

3) didactic practices related to the INVALSI items;
4) attitude towards formative assessment or the ideology of natural gifts (Ciani and Vannini, 2017).

The data collected in the third section refer to professional training as well as to the teachers' personal data. For example, they were asked to indicate: how many years they have been working as teachers; how long they have been teaching in the current school; how long they have been in their current role; what activities they carry out in their school; etc.

The tool just described was initially administered to 105 teachers (Field trial). This initial phase made it possible to test the questions in the questionnaire. Subsequently, based on the analysis of the data collected with the field trial, the questionnaire was partially modified.

The new version of the questionnaire was administered to 427 teachers (Main Study), of which 421 cases are valid. None of the previous responding teachers received the questionnaire again. The data collected in the two campaigns were encoded and analysed using a statistical software for data analysis (SPSS).

It was decided to consider the answers from both campaigns valid, capitalizing on the temporal proximity of the administration, with respect to the questions that were not modified after the field trial. Thus 526 valid cases were considered.

3. First results

We start by presenting the findings obtained through the items in the third section (personal data and context information) to delineate the characteristics of the sample of responding teachers.

– 95% of the respondents are female.
– 68% of the respondents received an invitation to fill in the questionnaire from their School and in particular from their Headmaster.
– 71% of the respondents teach in Piedmont or Emilia-Romagna (which together make up 15% of the Italian population).
– 90% of the respondents are tenured teachers.
– 21% of the respondents actively participate in school life (RAV, evaluation unit, INVALSI).

The average age of the responding teachers is 50.40 years, with a standard deviation of 8.56.

Although the sample is large it cannot be considered representative, yet the data collected offer a range of interesting information that will subse-

quently be considered along with the results from the other two sections of the questionnaire.

Before proposing deeper and more specific analysis, it seems appropriate to provide an overview[2] deriving from the analysis of the simple frequencies, inherent on the one hand, the purposes that the teachers attribute to the INVALSI tests and, on the other hand, the statements regarding the useful practices that the teachers make of the tests in the classroom. This frame of reference is also functional for the definition of training guidelines for primary school teachers – general objective of the research – to support them in more conscious planning of daily activities for managing mathematics teaching-learning processes starting from the INVALSI tests.

An interesting fact to initially note concerns the average score of the answers collected in item 30 of the questionnaire which request teachers to express their degree of agreement (from 1 to 4, where 1 is represented by the statement "completely disagree" and 4 by the statement "completely agree") with respect to the objective of assessment of the INVALSI tests.

The following table (Tab. 1) shows the question and teachers' answers:

Tab. 1 – Average score for each item of the question: "INVALSI aims to assess"

Item	*Average score (range 1 to 4)*
The Italian educational system	2.87
The individual student's learning in mathematics	2.17
The professional preparation of the individual teacher	2.05
The didactic effectiveness of schools	2.72

The highest averages recorded (2.87 and 2.72) concern the teachers' agreement in identifying the assessment aims of INVALSI in the effectiveness of the Italian educational system and subsequently in the educational effectiveness of the schools, but the other two show how, to date, despite the experience of the INVALSI tests in primary school which boasts a history of more than 10 years, there is a widely diversified panorama in terms of the teachers' understanding and sharing of the primary purposes of the test itself.

Assuming the hypothesis according to which the opinions and perceptions of teachers regarding the INVALSI tests promote or inhibit the didactic impact of the tests, a poor didactic impact of the tests can be hypothesised in the classes of teachers who think that the INVALSI tests assess the teach-

[2] This overview is the result of the joint work of the entire research group and aims to summarise the evidence drawn from the data and published in other works to date.

er's professional preparation. In this regard, returning to the training intent connected to the possibility of defining guidelines for teachers for improving teaching practices relating to the use of INVALSI mathematics questions – general objective of the research project – the data just presented arises in support of a formative perspective that not only takes care to explain the aims of the INVALSI tests, but also monitors the teachers' ongoing understanding and sharing.

Furthermore, the results regarding item 32 of the questionnaire are reported, aimed at investigating the statements of the teachers' use of the INVALSI test items in their classroom.

The teachers were asked how often they propose activities related to INVALSI questions to their students, or how often they draw inspiration from INVALSI questions to propose other activities to students.

The proposed index ranges from a minimum of 1, which indicates a subject who never uses similar tools, to a maximum of 4, which indicates a systematic use of these tools to assess students.

The average scores of the items of question 32 are shown below (Tab. 2).

Tab. 2 – Average score for each item of the question: "In my classroom practices"

Item	*Average score (range 1 to 4)*
I train the students to work on many INVALSI test questions of the previous years	2.86
I discuss the theoretical aspects of taking the INVALSI Tests with the students (structure of the tests, timing)	3.12
I reflect with the students in order to find "quick" and "crafty" strategies to solve the questions of the INVALSI tests	3.18
I draw inspiration from the INVALSI test for activities related to argumentation in mathematics	3.22
I draw inspiration from the INVALSI test for activities related to problem solving	3.28
I draw inspiration from the INVALSI test for activities related to the justification of one's answers	3.03

The set of questions presented identifies two potential uses of the INVALSI tests. One corresponds to a more training, performance-oriented type: the items referable to this factor concern the discussion in class on technical aspects of the questions, outlining faster solution strategies, and training students to answer correctly (1, 2 and 3). The other concerns with a use which focuses on encouraging the development of students' skills: the items referable to this factor are those relating to implement problem solving activities and argumentation of a mathematical nature (4, 5 and 6) (Truffelli and Vannini, 2021). Although the survey averages reflect a widespread

practice of the proactive use of INVALSI tests to promote students' skills, training practices linked to the tests are still quite widespread, with the highest frequency detected in selecting "often" in their use to "find quick and crafty strategies with students to answer the questions" (190), while the higher frequencies with the "sometimes" option are based on discussing the theoretical aspects of taking the tests (241) and on training students by repeating the questions of previous years' tests.

Also in this case there is a considerable diversification with respect to the most widespread uses of the test questions among teachers; however, even starting from this data, a possible path for furthering the research can be identified. Also in support of the conception of training guidelines for teachers, a more organic exploration of what are, in concrete terms, the practices defined as "training" by teachers can be evaluated through a qualitative coring.

Furthermore, two other thematic analysis paths applied by the group of researchers should be noted, the first focused on the data emerging from the first section of the questionnaire, i.e., those concerning mathematics teaching, while the second mainly concerns the teachers' attitude towards formative assessments or the ideology of natural gifts (Ciani and Vannini, 2017).

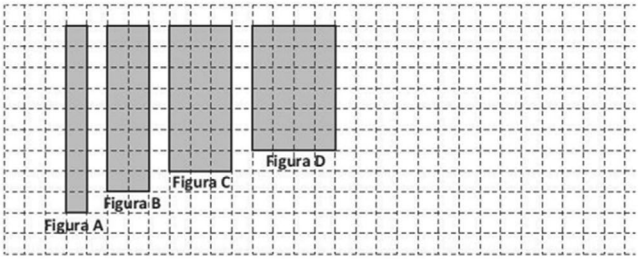
The first thematic analysis focused on the first four questions of the questionnaire which asked teachers to express their perceived degree of difficulty of the item for the students and to figure out possible reasons for the mistakes made.

The analyses conducted showed that teachers may have difficulty recognising the reasons for the errors made by students during the INVALSI tests (Vaccaro, Faggiano and Ferretti, 2021).

For each INVALSI item reproposed in the original formulation in the questionnaire, the teachers were offered four multiple choice options concerning possible correct and incorrect reasons for the error potentially made by students.

As an illustrative example of what is described above, the formulation of a question of the questionnaire that we indicate as Q3 (Fig. 1) is shown below.

D3. Look at this sequence of figures



a. Draw on the squares, next to Figure D, the next figure in the sequence.

b. Which of the following sentences is true?

A. The area of the figures is always the same
B. The area of the figures doubles at each step
C. The perimeter of the figures is always the same
D. The perimeter of the figures increases at each step

Fig. 2 − Question Q3 of the questionnaire

Following the INVALSI item, the possible reasons for the error made by the students were subsequently proposed in the questionnaire as follows:

– "Although 85.8% of students correctly answered part "a", only 35.7% of students correctly chose "C".
– In part "b", choose one of the reasons why:
  • "Students do not pay attention while reading the text".
  • "Students do not know area and perimeter formulas well".
  • "Students are led astray by the picture".
  • "Students believe that the area increases while the perimeter increases".
  • "Students do not pay attention while reading the text": we consider this as a boilerplate answer that we expect to be used by a teacher lacking knowledge of the didactic and epistemological issues.
  • "Students do not know area and perimeter formulas well": we can assume that most fifth-grade students have a working knowledge of computing areas and perimeters of such rectangles, but on the other hand the item does not ask for any explicit numerical result. Hence, we hypothesise that teachers who choose this option reduce the idea of "perimeter" and "area" to the computation of their values using the appropriate formulae, instead of considering the more general geometrical concept involved in the question.

- "Students are led astray by the picture": this option is very similar to the first, but might be chosen by teachers who recognise that the item is about geometry.
- "Students believe that the area increases while the perimeter increases": this is the answer we expect from a teacher aware of the didactical and epistemological issues at play.

Table 3 shows how the participants answered.

Tab. 3 – Distribution of the teachers' choices in answering Q3

Options	*%*
Students do not pay attention while reading the text	34.2
Students do not know area and perimeter formulae well	2.1
Students are led astray by the picture	32.7
Students believe that the area increases while the perimeter increases	21.5
Other	9.5

It can be seen that only 21.5% of the teachers recognised that the reason behind the students' error is connected to the misconception that areas and perimeters should behave in the same way.

The same reflection requested on this question was reproposed in relation to other items related to different mathematical contents. Similarly to what was proposed with the perimeter and area question, we asked teachers to figure out the possible mistakes made by the students.

It is interesting to report that in the four questions, the option that was generally chosen most was the students' "lack of attention in reading the text" (Tab. 3).

The weight that teachers attribute to the students' lack of attention in reading the text seems to be a symptom of a lack of awareness of the difficulties actually encountered by the students. There is the urgency of training need in terms of specific knowledge of didactic mathematics teaching and practices that make text comprehension an appropriate further subject in teaching mathematics.

Furthermore, thanks to the creation of a summary index of the teachers' responses to the four items listed, it is possible to show that 35% of the responding teachers did not identify correct interpretations with respect to the reasons for the errors and only five teachers, 1% of the total, correctly interpreted the reasons for the students' errors for all four questions.

Unfortunately, a strikingly high percentage of teachers (33.8%) identified only one of the correct reasons for the error, highlighting a concerning lack of awareness among this group (Vaccaro, Faggiano and Ferretti, 2021).

Tab. 4 – Frequencies of the "Awareness of the reasons for the error" index

Item	*Frequency*	*Valid percentage*	*Cumulative percentage*
Does not correctly answer any of the questions about the interpretation of the error	184	35.0	35.0
Correctly answers one of the questions about the interpretation of the error	178	33.8	68.8
Correctly answers two of the questions about the interpretation of the error	115	21.9	90.7
Correctly answers three of the questions about the interpretation of the error	44	8.4	99.0
Correctly answers four of the questions about the interpretation of the error	5	1.0	100.0
Total	526	100.0	

Also in this sense, it is possible to identify a further training path to follow with the teachers.

Moreover, this result seems to be in agreement with other considerations proposed regarding the first results emerging from the field trial relating to how primary school mathematics teachers read and interpret the data of the INVALSI tests (Arzarello and Ferretti, 2021). The latter had suggested the existence of a meta didactic conflict regarding the discourses on teaching processes such as student assessment, skills and errors, and not concepts of mathematics of thought per se, as in the case of usual epistemic or didactic conflicts reported in the literature (Brousseau, 1986). The cause of this conflict, referring to what Anna Sfard defines as «incommensurable discourse» (2008), seems to lie in a misunderstanding between teachers and students that occurs when the same words are used but have different meanings, being unaware of the reciprocal differences in their use. If not overcome, the resulting conflict can have serious consequences for the success of teaching/learning processes in the classroom. Arzarello and Ferretti propose considering this conflict as a reference also for discussing what emerged from the data presented in the first section of the questionnaire, identifying three different components: how teachers perceive the difficulties of students in the INVALSI questions; how teachers interpret students' responses and errors; how much teachers consider the INVALSI questions useful and how they use them in teaching practices (Arzarello and Ferretti, 2021).

Also in this case, a future research project line can be defined aimed at overcoming the identified conflict through a clearer understanding of its nature and an explanation of its components in order to obtain a real improvement of practices relating to the use of the INVALSI tests in schools.

The second thematic analysis path, relating to the teachers' attitude towards the formative assessment or the ideology of natural gifts, focused on the analysis of the data emerging from the attitude scale present in section two of the questionnaire, which measures the teacher's propensity to implement training assessment strategies, diagnosing errors in order to implement recovery strategies.

The scale refers to an attitude of the teacher oriented to analytically understanding the learning path carried out by the student, in order to control and self-regulate the didactic action in view of the students' specific cognitive needs. It is a conception of evaluation that aims at equity in learning outcomes and which strongly contrasts (in theoretical terms) the ideology of natural gifts, i.e., the belief that teaching is powerless in recovering the shortcomings in "certain" students (Truffelli and Vannini, 2021).

The results obtained will be presented and discussed in more detail in the following section, as they also serve to introduce the further analyses which are the subject of this chapter.

4. Analytic summary of evidence already obtained from this project about attitudes and use on INVALSI test in a formative perspective

In light of the overall picture outlined so far, in this chapter we have focused on blocks of variables relating to attitudes and beliefs on the formative function of assessment, on the INVALSI tests and their validity and on the declarations of use of the tests themselves in class, offering an analysis and interpretation of the connections that emerged on a correlational basis between the blocks of variables indicated. The general purpose of this work is mainly aimed at analysing the factors that influence the use practices of the INVALSI test in the classroom by primary school mathematics teachers. We asked ourselves if it is possible to trace a link between teachers' beliefs and attitudes towards assessment and towards the INVALSI tests and how the tests are used in the classroom, and lastly whether those beliefs and attitudes are related to specific practices of using the INVALSI tests.

Regarding these practices, we also asked what use teachers make of these tests. Do they use these tools to promote didactics that develop and strengthen students' skills? Or do they focus their use on results and performance to train their students on the mere passing of these tests?

A first study (Truffelli and Vannini, 2021) conducted as part of this survey investigated these aspects, exploring them through second-level analyses. We will briefly present the results to then introduce the further analyses which are

the subject of this contribution, given the close connection between the former and the latter. Two sets of questions, one relating to the perceived usefulness of the INVALSI questions by teachers and the other relating to the type of use of the questions in classroom teaching, were subjected to factor analyses of the main components. We defined the two components that emerged regarding the perception of use as a) usefulness of the INVALSI tests to analyse, reflect on and plan teaching; b) usefulness of "training" for the INVALSI tests. These components reveal two distinct orientations underlying the use of the tests, one aimed at planning actions such as the definition and redefinition of ad hoc didactic objectives for the class, the other aimed at teaching-to-test to prepare students to pass the system assessment tests. The components resulting from the same analyses conducted on the set concerning the type of use of the INVALSI tests in class are: a) use of the INVALSI tests to favour the development of students' skills; b) use of the INVALSI tests for training. Different uses are therefore made of the items taken from these tests: in some cases oriented to the process of building competence, favouring an argumentative exchange with the teacher and among students and the development of critical thinking; in other cases it is more focused on performance, technical aspects and faster resolution strategies.

To give a further answer to the first questions mentioned above, Truffelli and Vannini (2021) developed an exploratory regression model. The model investigated beliefs and attitudes that can favour teachers' propensity for didactic choices consistent with a formative perspective. The exploratory regression was conducted on a summary index relating to the statements about the use of the INVALSI questions which we called "Systematic use of INVALSI tests to promote skills". The analysis in question identified three elements that significantly regress on the dependent variable: the factor that reveals an attitude about the usefulness of the INVALSI tests for analysing, reflecting and planning didactics (beta = .401); the "Engaged teachers" index (beta = .236) which identifies those who have particular roles related to assessment within their own school and have been specifically trained in mathematics assessment and teaching; the scores on the scale that measures the propensity towards a diagnostic-formative use of the assessment (beta = .188) (Ciani and Vannini, 2017). As the authors conclude, «We can affirm that the propensity to act in a formative perspective is greater in teachers who study and are involved in relation to the themes of didactics and assessment, who have the ability to recognise and use the formative function of assessment and above all interpret the INVALSI tests as a tool and occasion for error analysis and enhancement» (Truffelli and Vannini, 2021, p. 391).

Starting from the hypothesis that attitudes and beliefs can influence the practices that teachers implement in their professional action, we focused our

attention on further aspects, more specifically an analysis and interpretation of the links that emerged starting from statistical correlations between three elements: a) teachers' beliefs about the type of knowledge and skills detected by the INVALSI test; b) its validity; c) statements regarding the teaching practices connected with the use of the INVALSI test in the classroom.

5. Use of INVALSI questions at the microsystem level

The analyses presented in this section, carried out on the database taken from the administration of the questionnaire to over 500 primary school mathematics teachers, allowed us to develop some reflections regarding the use made of this nationally administered learning survey by teachers in the micro-contexts to which they belong. We focused in particular on the correlations between two indices constructed in a first phase of the research starting from specific questions of the questionnaire and one of the factors that emerged in the factor analyses.

The first index measures the extent to which teachers believe INVALSI items are suitable for assessing students' learning. In fact, the questionnaire included seven INVALSI mathematics items that were released after their use for the national testing and the following question was asked for each of them: "On a scale from 1 to 4, how well do you think this question is suitable for assessing your students' learning in the fifth year of Primary School?". The scores were indexed in a single variable.

The second index measures how often teachers use INVALSI questions to develop students' skills. We selected and included in the index three specific items that were particularly relevant to the topic (see Tab. 5) from the set of questions: "Please indicate how frequently you use the following practices in class in reference to each of the following statements. In my classroom practices...: (omitted)".

Tab. 5 – Items included in the index

	Items included in the index
1	I draw inspiration from the INVALSI tests for activities related to argumentation in mathematics
2	I draw inspiration from the INVALSI tests for activities related to problem solving
3	I draw inspiration from the INVALSI tests for activities related to justifying one's answers

Tab. 6 – Correlations under analysis

		INVALSI test usefulness for analysing, reflecting and planning	*Use of the INVALSI test to enhance math skills*	*INVALSI test validity*
INVALSI test usefulness for analysing, reflecting and planning	Pearson Correlation	1**	,515**	,331**
	Sig. (2-tailed)		,000	,000
	N	463	463	463
Use of the INVALSI test to enhance mathematical skills	Pearson Correlation	,515**	1**	,310**
	Sig. (2-tailed)	,000		,000
	N	463	463	463
INVALSI test validity	Pearson Correlation	,331**	,310**	1**
	Sig. (2-tailed)	,000	,000	
	N	463	463	463

** Correlation is significant at the 0.01 level

Finally the third factor concerning the teachers' attitudes towards the usefulness of the INVALSI tests to analyse, reflect and plan didactics was found to correlate consistently and significantly with both indices, in particular with the second.

The highest correlation reported in the last column puts in evidence a positive and statistically significant link[3] between recognition of the validity of the INVALSI test to measure and assess mathematical skills and the usefulness (perceived by teachers) of the items for analysing, reflecting and planning their own didactics (.331**) (Tab. 6).

This correlation shows that those who consider the measures obtained from this test as non-random, reliable and referable to a clear conceptual framework, are inclined to consider the test itself as a useful tool capable of providing a solid foundation of data. They attribute a diagnostic assessment function to the test (Vertecchi, 2003; Domenici, 2003), able to highlight the skill levels of each student. Starting from these data, these teachers can reflect on their own didactics, on the needs that emerged in their class and on the opportunity to plan work paths that give specific answers with respect to the picture that emerged from the analysis of the test results.

A second interesting correlation is the one that highlights a positive and significant link between the recognition of the validity of the INVALSI test and the statement of use of the test (in teaching practicies) to enhance their mathematical skills (r = 0.310**) (Tab. 6). Our analyses led us to consider that those teachers, who consider the INVALSI test as a useful tool for assessment, not only recognise the diagnostic potential to exploit upstream of the didactic action, but also appreciate the reinforcing potential, which can strengthen learning downstream or during explanations, consolidating the message through operational practices. According to our interpretation, recognising and giving value to the theoretical construct that underlies the INVALSI test represents a non-obvious element: this element represents a necessary condition for the attribution of validity by teachers to each of the items released and available for didactic use, including the microsystem level evaluation function in the classroom context.

A third high correlation indicates a marked covariation between the use of INVALSI questions to analyse, reflect and plan didactics, and the habit of using INVALSI items to work with students to enhance their mathematical skills (.515**) (Tab. 6). Some teachers consider the INVALSI questions as a

[3] The correlations were calculated with Pearson's r and the strength is reported in parentheses, with the level of significance with asterisks in the contribution for each, according to conventional standards recognised in the scientific community.

valid basis for structuring and restructuring teacher interventions in a targeted and individualized way based on the needs of the students. We interpreted that belief as a prerequisite for teaching practices don't oriented as much to train students to pass the INVALSI test, but rather to help students in developing those specific mathematical skills involved in their resolution.

6. Reflections based on the data

What do the correlations that we have analysed in this contribution indicate as a whole? Undoubtedly, they are revealed as being strongly interconnected. This first observation makes us understand something of interesting. The three summary variables constructed, 1) the recognition of the validity of the INVALSI test to assess students'skills, 2) the use of INVALSI questions in the classroom context to develop students' skills, 3) the recognition of the usefulness of the INVALSI tests for analysing, reflecting and planning didactics, are linked by a common ground. We can trace a convergence in them towards certain convictions and attitudes relating to the teaching-learning process: it is intended above all as the result of a work that requires specific interpretations of the context supported by diagnostic assessments. The teacher who emerges from this vision is therefore the same person responsible for the process. Furthermore, the emerging vision of the teaching-learning process shows strong connections with specific professional skills: 1) teachers' reflective and analytical capabilities 2) teachers' ability of planning the didactic activities according to a medium and long-range perspective 3) teachers' skills in carry out what planned. In this sense, the teacher assumes the responsibility of a programmatic vision of the work in the classroom and manages its direction. Moreover, the vision of the teaching-learning process, that is outlined starting from the correlations discussed, testifies to a conception of learning that is far from an innate vision, which considers the act of learning as an almost exclusive result of natural gifts the students already have. Rather, it appears to be conceived as the result of a positive, intentional and guided interaction with the notion that skills can be acquired, trained and developed, an interaction that enhances the use of valid tools to work on strengthening students, also in a targeted way.

These correlations show that their beliefs have influence teachers' didactic choices and actions. Recognising the validity of the INVALSI tests, in relation to the skills they demand of students on the one hand, and their usefulness for analysing, reflecting and planning didactics on the other, affect the type of prevalent use they make of the test in the classroom context:

a use mainly aimed at developing skills and less oriented towards teaching-to-test. The teachers with high scores on the indices and factors used for these second-level analyses, interpret themselves as professionals who do not act merely by their simple vocation and believe that learning can and should be constructed. They develop reflections to improve and adapt their didactic proposal to students and help them enhance their mathematical skills. In summary, among the respondents, we can trace the figure of a teacher who works on two fronts: the first direction is oriented to build a stimulating learning environment; the second direction aims to strengthen students' skills for encouraging them to approach their learning path in a critical and reflective way. It is a type of teacher who recognises the validity of the INVALSI test constructs and is therefore inclined to identify them as useful tools for analysing their teaching action, reflect on it and redesign it in a manner which is functional to the learning objectives, and to use the items derived from the INVALSI tests in classroom assessment practises more frequently.

7. Conclusions and prospects

In conclusion, the study that was carried out intercepted a current and interesting topic and provided significant results. It allowed us to shed light on the potential ways of improving mathematics teaching thanks to the monitoring offered by the INVALSI tests.

The relationship between teaching practices and tests, as emerging from the data, still needs to be explored and deepened from many points of view. It however seems to be a crucial link to confirm the importance of recognizing and enhancing the professionalism of teachers. Their key role consists in giving an interpretation of the INVALSI tests results and in drawing inspiration from them to implement processes capable of raising the quality level of teaching. The highlighted elements contribute to outlining a rich picture, which may, however, be further confirmed and clarified by the collection of wider range of data, both in terms of the representativeness of the sample and in terms of depth of the analysis. For example:

– the preparatory activities in view of the administration of tests could be also explored;
– the processes of design and implementation of educational interventions suggested in some way by the tests could be analyzed;
– dynamics relating to the return of the information collected and its exploitation may be studied in order to improve the quality of teaching.

The results set out above also encourage to explore the logical-pragmatic patterns in order to describe and validate "good practices" of teaching mathematics as favoured by the INVALSI tests. Such dynamics can be considered as "habits" of professional activity. They do not always give rise to processes of mentalization and speech and may, have a "pre-semantic" nature. A possible development of this study could therefore go in the direction of deepening their knowledge and understanding with a set of ad hoc tools, hopefully capable of also reaching the practices actually implemented. The aim is to focus on elements of educational quality that may be exported as part of training and accompanying initiatives to enhance the INVALSI tests as potential tools for improving teaching and learning processes. Furthermore, these are 'ways' of interpreting class work that are constructed during the activity itself also on the basis of the relationships that the teacher maintains with students as the literature suggests. This may suggest an effort to rebuild the context and ways of participating in the communities of professional practices that have generated certain approaches to teaching from tests. On the other hand, precisely the emergence from the data of distinctive dynamics potentially capable of bringing "added value" to the use of INVALSI tests in the classroom recommends not to neglect different professional interpretations, characterised by greater misalignment between the use of tests and classroom teaching practices. It would be the subject of analysis in a pro-active and improving key, possibly also "in" and "with" the teacher communities. This, in the wider dialogue framework of the relationships between INVALSI tests and didactic-evaluation processes, to be explored, in perspective, also in the development of innovative methods and strategies for conducting the class.

On a methodological level, the study contributes to confirming the importance of promoting collaborative dialogue between researchers and teachers. This strategic synergy could create the conditions so that "experiential knowledge" and formalized knowledge about teaching can find basis for discussion. Practical and theoretical problems on how to promote the best learning by each student, indeed, can find a ground for mutual clarification and for progressive synthesis. In this way it is possible to give life to forms of knowledge that are increasingly responsive to the professional action of teachers, promoting operational choices that are based on cognitive foundations.

References

Altet M. (2006), "Le competenze dell'insegnante-professionista: saperi, schemi d'azione, adattamenti ed analisi", in M. Altet, É. Charlier, L. Paquay, Perrenoud P. (a cura di), *Formare gli insegnanti professionisti. Quali strategie? Quali competenze?*, Armando, Roma.

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

Barbier J.M. (2017), *Vocabulaire d'analyse des activités*, PUF, Paris.

Brousseau G. (1986), "Fondements et méthodes de la didactique des mathématiques", *Recherches en Didactique des Mathématiques*, 7 (2), pp. 33-115.

Calvani A. (2019), "Prefazione", in G. Bonaiuti, *Le strategie didattiche*, Carocci, Roma.

Cardarello R. (2019), "Dimensioni metodologiche nella Ricerca-Formazione", in G. Asquini (a cura di), *La ricerca-formazione. Temi, esperienze, prospettive*, FrancoAngeli, Milano.

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

Damiano E. (2004), *L'insegnante. Identificazione di una professione*, La Scuola, Brescia.

Damiano E. (2013), *La mediazione didattica. Per una teoria dell'insegnamento*, FrancoAngeli, Milano.

Domenici G. (2003), *Manuale della valutazione scolastica*, Laterza, Roma-Bari.

Hattie J. (2016), *Apprendimento visibile, insegnamento efficace. Metodi e strategie di successo dalla ricerca evidence-based*, Erickson, Trento.

Lenoir Y. (2017), *Les médiations au coeur des pratiques d'enseignement-apprentissage: une approche dialectique. Des fondements à leur actualisation en classe: éléments pour une théorie de l'intervention éducative*, Éditions Cursus Universitaire St-Lambert, QC.

Paquay L., Altet M., Charlier E., Perrenoud P. (2006), *Formare gli insegnanti professionisti. Quali strategie? Quali competenze?*, Armando, Roma.

Pastré P. (2007), "Quelques réflexions sur l'organisation de l'activité enseignante", *Recherche et formation*, 56, pp. 81-93.

Perrenoud P. (1993), "Curriculum: le formel, le réel, le chaché", in J. Hussaye (*ed.*), *La pédagogie: une encyclopédie pour aujourd'hui*, ESF, Paris.

Perrenoud P., Altet M., Lessard C., Paquay L. (2008), *Conflits de savoirs en formation des enseignants: entre savoirs issus de la recherche et savoirs issus de l'expérience*, De Boeck, Bruxelles.

Savoie-Zajc L. (2004), "La recherche qualitative/interpretative en education", in T. Karsenti, L. Savoie-Zajc (*dir.*), *La recherche en éducation: étapes et approches*, CRP, Sherbrooke.

Schön D.A. (1983), *The Reflective Practitioner. How Professionals Think in Action*, Basic Books, New York.
Sfard A. (2008), *Thinking as communicating: human development, the growth of discourses, and mathematizing*, Cambridge University Press, Cambridge.
Shulman L.S. (1986), "Those who understand: Knowledge growth in teaching", *Educational Researcher*, 15 (2), pp. 4-14.
Tardif M., Lessard C. (2004), *Le travail enseignant au quotidian*, PUL, Laval.
Truffelli E., Vannini I. (2021), "Convinzioni e atteggiamenti degli insegnanti di scuola primaria italiani e orientamento all'uso formativo delle prove INVALSI di matematica", in P. Lucisano (a cura di), *Ricerca e didattica per promuovere intelligenza comprensione e partecipazione*, Pensa Multimedia, Lecce.
Vaccaro V., Faggiano E., Ferretti F. (2021), "Consapevolezza degli insegnanti delle ragioni degli errori degli studenti", in P. Lucisano (a cura di), *Ricerca e didattica per promuovere intelligenza comprensione e partecipazione*, Pensa Multimedia, Lecce.
Vergnaud G. (2011), "Au fond de l'action, la conceptualization", in J.M. Barbier, *Savoir Théoriques et savoir d'action*, PUF, Paris.
Vertecchi B. (2003), *Manuale della valutazione*, FrancoAngeli, Milano.
Viganò R. (2019), "Parlare di metodi di ricerca pedagogica non è questione di tecniche", *Pedagogia Oggi*, 18 (1), pp. 27-35.
Wenger E. (1998), *Communities of Practices, Leaning, Meaning and Identity*, Cambridge University Press, Cambridge.
Zanniello G. (2023), "Educational practices, evidence-based experimentation and R-A whith project", *Italian Journal of Educational Research*, 16 (30), pp. 11-22.
Zanniello G. (2020), "Relazione introduttiva", in *Le Società per la società: ricerca, scenari, emergenze. Atti del Convegno Internazionale SIRD*, Roma 26-27 settembre 2019, pp. 32-43.

4. Linking student achievement to teacher assessment practices: issues, challenges, and implications for educational research and practice in Italy

by Gianluca Argentin, Serafina Pastore

The use of assessment for instructional guidance represents one of the most powerful levers for school improvement. With a strong emphasis on the alignment of teaching, learning, and assessment, the reconsideration of assessment methods and strategies has become relevant, as well as the need for a new assessment culture more responsive to educational policy requirements, school system needs, and teacher practice instances. In this vein, for example, formative assessment has been recognized as integral to curriculum design and to the learning journey of all students. More specifically, teacher assessment practice and teacher assessment literacy have been recently considered as crucial for teacher professionalism, as well as for the improvement of student learning. However, despite the emphasis on teacher assessment literacy, recent research confirms that teachers are not prepared to effectively integrate assessment into their daily teaching practice and that low levels of assessment literacy tend to negatively affect student learning achievement.

Entering this lively debate, the present chapter identifies which assessment practices are more frequent in the Italian school system and tries to understand if some assessment practices are more likely to influence student learning as measured in the INVALSI tests. Using Maths and reading test scores from the INVALSI database (grade 5), as well as data from the INVALSI Teacher Questionnaire (2017/2018), OLS regression models and student fixed effects models have been performed. Analyses show that teachers' own design of standardized tests and the use of online assessment tools (a rare practice) are associated to positive student performance. We interpret this association as the consequence of students' socialization to the practice of testing.

Con una forte enfasi sull'allineamento dei processi di insegnamento, apprendimento e valutazione, la riconsiderazione dei metodi e delle strategie

di valutazione è stata considerata rilevante così come la necessità di una nuova cultura della valutazione, più rispondente ai requisiti delle politiche educative, alle esigenze del sistema scolastico e alle istanze pratiche degli insegnanti. In tal senso, la valutazione per l'apprendimento è stata, ad esempio, riconosciuta come parte integrante della progettazione curriculare e del percorso di apprendimento degli studenti. Nel corso degli anni, la pratica valutativa e la competenza valutativa degli insegnanti sono state identificate come componenti cruciali della loro professionalità e funzionali al miglioramento dei risultati di apprendimento degli studenti. A dispetto dell'enfasi sulla competenza valutativa dei docenti, alcune ricerche confermano come gli insegnanti non siano preparati per integrare in modo efficace la valutazione nelle loro attività didattiche quotidiane e come bassi livelli di competenza valutativa influenzino negativamente i livelli di apprendimento degli studenti.

Nel solco di questo dibattito, il presente contributo identifica quali pratiche di valutazione sono più frequentemente utilizzate dagli insegnanti italiani e mira a comprendere se le pratiche di valutazione tendano a influenzare l'apprendimento degli studenti e quindi anche le loro performance nelle prove INVALSI. Usando i risultati delle prove di Italiano e Matematica del database INVALSI (V primaria) e i dati ricavati dal Questionario Insegnante per gli anni scolastici 2017-2018, si sviluppano modelli di regressione OLS e modelli a effetti fissi di studente. Si osserva che l'auto-produzione di test standardizzati e il ricorso a strumenti di valutazione online da parte degli insegnanti (pratica molto rara) si associano a migliore performance degli studenti. Si interpretano queste associazioni come conseguenza di una loro socializzazione al testing.

1. Introduction

During the last two decades, the widespread diffusion of the neoliberal ideology (Lewis and Lingard 2015; Schleicher, 2018; Smith, 2016) in schooling policy led to the radical review of different crucial aspects of education (from parental involvement to teacher and student conceptions and practices). The launch of reforms in teacher education (e.g., in the US, UK, Australia, Singapore, Hong Kong or, within the EU context, in Germany and Finland) raised the standards for education quality (Hargreaves, 2020; Fullan *et al.*, 2018; Zao, 2016). Teachers and teaching have been progressively recognized as key levers in educational reforms, as well as in the debate about educational quality (Hanushek and Rivkin 2006). More specifically, the emphasis on assessment (as a component of teaching practice) and on the

use of assessment information has been unprecedented. In this perspective, several adjustments have been suggested to generate both accurate accountability data and robust evidence of individual mastery that support effective instructional decision-making and better learning within national and local school systems (Darling-Hammond, 2017; Farrell and Marsh, 2016; Stiggins, 2017). Pursuing the alignment of teaching, learning, and assessment (Birenbaum *et al.*, 2015; Gardner *et al.*, 2010), the call for a reconsideration of assessment methods and strategies arises (e.g. in terms of the enlargement to a broad spectrum of activities from constructing paper-pencil tests to grading, to interpreting tests scores). At the same time, the need of a new assessment culture, more responsive to educational policy requirements, school needs, and teacher practice, has been also recognized (DeLuca and Bellara 2013; OECD, 2018).

The interest reserved for student learning assessment, as well as the growing emphasis on standardized tests and high-stakes assessment, led national school systems to diversely frame assessment in order to reply both to accountability requirements and educational improvement (e.g., in Australia, Canada, South Korea, Hong Kong) (Smith, 2016).

Firstly, following the theoretical perspective of assessment for learning (Black and Wiliam 1998; McMillan, 2017), classroom assessment practices, far from the measurement tradition (e.g., assessment as functional to produce accurate estimations of student learning to monitor and report on processes) have been recognized as relevant, because they provide robust information to document and support both teachers and students in fostering student learning (Andrade and Heritage 2018; Filsecker and Kerres 2012; Kingston and Nash, 2011). The main idea, in this vein, is that assessment effects student learning and, hence, that assessment practices and strategies have to be competently used by teachers.

Secondly, despite the traditional resistance to and persistent confusion around its purposes and relevance (Stiggins, 2017), large-scale assessments have been progressively recognized by school administrators and school personnel (i.e., principals and teachers) as a component of educational governance and instructional quality systems.

Thirdly, within the framework of evidence-based education, data and data-use have become crucial «to make many types of decisions from school improvement to classroom and instructional decision-making» (Mandinach and Gummer, 2016, p. 1). However, some studies (Boardman and Woodruf, 2004; Brookhart, 2011) already demonstrated that teachers are reluctant to change their assessment practices (and conceptions), especially when new practices are framed within the rationale of institutional reforms (Brown,

2004; Klieger, 2016; Remesal, 2007), or within new scenarios such as those that emerged during the Covid-19 pandemic. Despite the recognition of the importance of assessment, some studies (Hopfenbeck, 2015; Looney *et al.*, 2018) focused on the lack of modernisation and indicated that assessment has not really changed. Evidence from the ground shows that core assessment practices remain problematic, both at theoretical and methodological levels. More precisely, using educational research to change assessment and feedback practice, or promoting student learning and academic self-regulation through assessment, are often perceived by teachers as difficult actions (Baird *et al.*, 2014; Hopster-den Ottera *et al.*, 2017).

On the backdrop of this debate, the present chapter intends to investigate the relationship between teacher assessment practices and student learning performances using data from the Italian national assessment of students learning (i.e., INVALSI database), a data source providing information both at student and teacher levels.

Following the theoretical perspective on school improvement and teacher assessment literacy, this study examines teacher assessment practices and testing, to detect whether they generate impacts on student learning.

To this end, in the first section, the article explores issues around the concept of assessment practice and assessment literacy. The second section, instead, reports the results of the analyses performed on the INVALSI database. The last part discusses policy implications deriving from our results, both for teacher professional development and for educational research, especially in Italy.

2. Background and conceptual framework

At the cornerstone of standards-based educational systems stands the growing attention that assessment received by educational policies during the last decade. On the one hand, the research on educational assessment and standardized testing has dominated most of the 20th century; on the other hand, classroom assessment and formative assessment have attracted a new considerable interest. More specifically, a great deal of attention has been turned to research documenting the potential impact of assessment on student learning (Hattie, 2009; Shavelson *et al.*, 2008). Teacher classroom assessment practice – which involves all aspects of designing, administering, and utilizing assessment to support or report on student learning – has been widely scrutinized. The benefits of assessment-based teaching to student achievement, meta-cognitive abilities, motivation, positive self-perception,

and enhanced teacher instruction led, indeed, to a progressive differentiation of teacher assessment practices in the classroom. Tests, quizzes, homework assignments, and questions at the end of the learning unit, as well as projects, rubrics and checklists are now widespread in schools and are used by teachers to enact assessment approaches, which promote student learning towards educational standards.

In this scenario, investigating teacher assessment approaches became relevant in terms of alignment between educational policies and educational practices, as well as in terms of educational quality and learning standards improvement. The emphasis on the impact that teacher assessment practice has on student learning, moreover, led to carefully reconsider knowledge, skills, and dispositions required by a teacher to ensure a sound assessment practice. Consequently, educational policies and educational research started to call for assessment literate teachers.

Teacher assessment literacy, generally, corresponds to a set of knowledge and skills required to ensure appropriate design, selection, interpretation, and use of assessment for instructional practice. Popham (2017; 2009) has defined assessment literacy as the teacher's understanding of principles and main concepts of a valid and reliable assessment. Thus, an assessment literate teacher is able to:

– identify teaching priorities;
– define and implement effective instructional strategies;
– identify strength and weakness in student learning;
– decide how to adjust and align instructional strategies.

Teacher assessment-literacy, in this vein, is based on a set of elements, such as the educational assessment rudiments, the ability to identify and to differentiate the assessment purposes, the ability to analyse data in order to improve teaching strategies. Therefore, assessment literate teachers are likely to better respond to student learning needs (DeLuca and Johnson, 2017; Smith *et al.*, 2013).

These conceptualizations of assessment literacy, initially focused on practical aspects of assessment, moved progressively towards more complex definitions. The recognition of teacher professionalism as a key component in educational systems' effectiveness has triggered a different interest for assessment literacy and for teacher professional development. From a socio-cultural perspective, current educational research suggests a more nuanced working definition of assessment literacy, linked to the social identity of teachers as assessors, as well as to their assessment dispositions, assumptions, and conceptions (Chan and Luo, 2020; Looney *et al.*, 2017). Moreover, recent attempts to re-conceptualize teacher assessment literacy have ad-

dressed the importance of data-use and decision-making (Cowie and Cooper, 2016; Mandinach and Jimerson, 2016; Schildkamp, 2019). However, this recognition, especially in some school systems (e.g. in the U.S., the Netherlands, Australia) provoked an increasingly overlapping of the assessment literacy concept with those of data literacy and statistical literacy. As a consequence, it is still difficult to clearly define the assessment-literacy concept in the compulsory education field.

Within such fluctuating research streams, moreover, scant is the evidence on the links between teacher assessment practice and student learning improvement. In addition, considering the Italian case, where national assessment is quite recent, even descriptive evidence about teacher behaviour is still extremely scant.

Given this framework, the present chapter, grounding on the debate on teacher assessment literacy and assessment approaches, tries to shed light on Italian teacher assessment practices and their consequences on pupils' achievement.

3. The research study

3.1. Research aims and questions

This chapter tries to describe Italian teacher assessment practices and to estimate whether they have an impact on pupil achievement. We aim to identify which assessment practices are more frequent in the Italian school system and to estimate to what extent they influence student learning, thanks to the fact that it is possible to connect student scores in the INVALSI national assessment and the self-reported assessment practices of their teachers. Considering that Italian teachers generally received no specific training on student assessment (especially in pre-service education paths), we also investigate whether there is an association between their participation in professional development in the domain of assessment and student learning achievements, as measured by INVALSI national large-scale assessment.

3.2. Data and methods

In this study, we focus on learning levels achieved by students at the end of primary school (5th grade) and on the assessment practices of their teachers. We use the national main samples of students assessed in school

year 2017/18 and the answers reported by their Italian and Maths teachers in the related *Teacher Questionnaire*. Statistical uncertainty of estimates is mainly driven by the (relatively small) size of teacher samples. Nonetheless, as shown in Table 1, we rely on large students samples, a condition allowing us to also detect tiny associations.

Tab. 1 – Sample sizes by subject and percentage of pupils whose teachers answered the questionnaire

	Students	*Teachers*	*% teachers answering*
Italian	17,323	950	66
Maths	16,968	952	69

We use the data to develop three sets of analyses:

a) descriptive analyses simply reporting to what extent Italian and Maths primary teachers declare regularly implementing the 10 assessment practices suggested in the questionnaire or declared participating in professional development on the topic, in the previous two years;
b) correlational analyses based on OLS regression models, where we investigate the association between having a teacher declaring the use of each assessment practice/to have attended professional development on the topics, on one side, and their pupils' scores in that subject, on the other side. We run separate models for the 10 assessment practices considered in the questionnaire plus the participation in professional development, due the impossibility to consider in a unique model all these elements together, since we would generate estimates based on an over-fitting regression model. In other words, we would generate a multidimensional space too complex considering the occurrence of cases displaying the high number of combinations of teachers' answers among the many practices. In these models, we consider together Italian and Maths scores for pupils, simply counting a student twice if he/she participated in both assessments (generating a long format dataset). Since students' scores are affected by the subjects and related assessment tools, the associations between teacher practices and students' scores are controlled for subject. In order to reduce the risk that these associations are spurious, our multivariate models control also, at the student level, for ESCS, gender, migration background, pre-primary school attendance, previous failures, geographical region, while at teacher level for gender, age, having a temporary contract, age of experience, organizational roles in the school and a general index of exposure to professional development. In the following tables, we will label parameters coming from these models as "net asso-

ciations", since we cannot exclude that the associations detected between teacher practice and student outcome may be due to omitted variables in the model, at context, school, class, teacher and student level. We do not separate the associations between Italian and Maths, in order to make the results comparable to those coming from the following set of models, where this distinction is not feasible;

c) additional correlational analyses, moving a step forward in the direction of a causal estimate, were implemented with students fixed-effects models (FE). These models take advantage from the fact that, generally, we have for each student two outcome measures (Italian and Maths) and, quite often, two reported teacher behaviours. Since these do not always overlap on the same student, meaning that their Italian and Maths teachers implement different assessment practices, we compare their results; in the subsample of students in this condition (variability at teacher level in assessment practice) we estimate an association between teacher practice and student outcomes, hence based on within students' differences. In this case, our models apply a student fixed-effect parameter controlling for all characteristics at student level, even those not considered in our dataset. The advantage of this second set of models is that they move us a step forward in the direction of exploiting a causal relation, because we are now controlling for a group of characteristics occurring at student, class, school, and context levels, even if not directly observed: indeed, the comparison takes place on the same student, looking at teacher variability. For this reason, in the FE models we still control also for the teacher characteristics listed above. Our association is now more reliable in terms of causal estimate, but the price we pay is that our sample size sharply decreases and that we lack external validity. More precisely, our sample shrinks only to students where both Italian and Maths teachers answered (865 couples) and, even worse, only to students displaying variability at teacher level regarding the independent variables, namely to the subsample of Italian and Maths teachers who answered differently regarding their assessment practices. Parameters coming from fixed-effects models will be labelled as "robust association" in the following tables, because there is still the possibility that we are omitting relevant variables at teacher level, hence they cannot be fully considered as casual effects.

4. Results

As shown in Figure 1, there is high variation among teachers in terms of practices used regularly to assess pupils: traditional written examinations are extremely common while the use of online tools is, conversely, rare. Oral examinations are also widely used; the main difference here is between the planned or unplanned strategy adopted by teachers.

Interestingly, there is not so much variation in terms of the adopted assessment practices between Italian and Maths: indeed, the bar length is essentially the same for all the alternatives suggested in the questionnaire.

We also detect that attention is paid in the school to assessment of students: at least among respondents, more than one teacher out of five participated in professional development on this topic, in the two years prior to the survey.

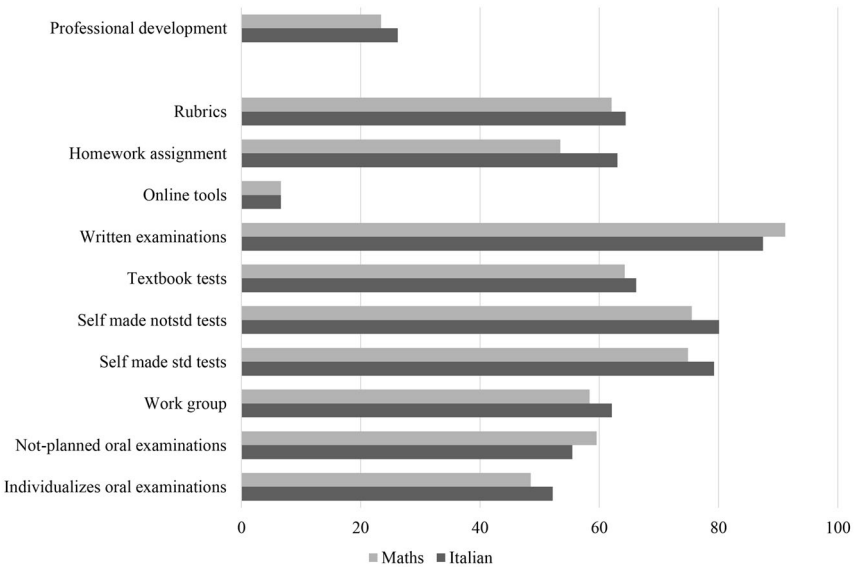


Fig. 1 – Teachers reporting to regularly implement assessment practices or to have participated to related professional development, by subject (%)

We wonder now whether teacher training and practices have an impact on pupil achievement. As stated above, there are good reasons to think so; at the same time, it is also possible that different practices do not make a big difference, due to the fact that skilled use of alternative assessments may not be so common among the Italian teacher population.

In the following tables, we report the associations between the assessment practices and the professional development, on one side, and, on the other, pupils' achievement. Considering the trade-off between external and internal validity of the two estimation methods adopted, we will now comment jointly results. It is exactly choice to compare results coming from OLS and FE models implying that it would not be useful to separate the OLS models results by subjects.

First of all, we can observe that all the reported parameters are small (considering that performance scores are reported on a scale with mean 200 and standard deviation 40). More importantly, we observe that only two net associations are confirmed as not null by the second regression model, the one estimating in a more reliable manner a potential casual link: teachers' use of online tools and of self-made standardized tests, like those administered by INVALSI.

Tab. 2 – Teacher assessment practices and their estimated impacts on pupil achievement (regression parameters and related standard error)

	Net association (OLS models)	*Significance*	*Robust association (FE models)*	*Significance*
Individualized oral examinations	-0.358		-0.5	
	(0.452)		(0.466)	
Unplanned oral examinations	-0.274		1.608	***
	(0.464)		(0.466)	
Work group	0.502		0.112	
	(0.447)		(0.455)	
Self-made std tests	1.232	**	1.251	**
	(0.524)		(0.518)	
Self-made non-std tests	1.247	**	0.554	
	(0.515)		(0.503)	
Textbook tests	-0.415		0.418	
	(0.478)		(0.465)	
Written examinations	1.841	***	0.529	
	(0.686)		(0.634)	
Online assessment tools	4.771	***	2.405	**
	(0.959)		(0.943)	
Homework assignment	-2.095	***	0.303	
	(0.466)		(0.470)	
Rubrics	1.53	***	-0.507	
	(0.468)		(0.481)	
Professional development	0.198		-0.123	
	(0.533)		(0.552)	

Note: standard errors are reported in brackets; statistical significance levels are reported according to the usual standards: *** $p < 0.01$; ** $p < 0.0$; * $p < 0.10$

A third effect detected in our analyses is that related to unplanned oral examinations, where we do not detect a net association.

5. Conclusions and future developments

Previous analyses suggest that testing pupils matters to increase their achievement, both when teachers develop self-made standardized tests and when they use online tools. These beneficial effects, despite being limited in size, are promising, considering that the vast majority of teacher assessment practices do not make a difference and that not even teacher professional development matters. A plausible interpretation of the positive effects of exposing children to standardized self-made assessments and related online tools is that these practices are a form of pupil "socialization" to standardized test, helping them to deal with a not so-common assessment experience, at least in Italian primary schools. What we are detecting seems not teaching to the test, being Italian standardized assessment low stakes, especially in primary school. Instead, we are observing that, when very young pupils are regularly exposed to this kind of assessment tool, they better express their skills, probably thanks to reduced anxiety. On the one hand, the widespread use of self-generated standardized assessments among teachers is encouraging; on the other hand, this evaluative tool being quite recent in Italy, we are probably not using it fully effectively. Teacher professional development should invest more in promoting the skills necessary to self-produce high quality standardized assessment; this would also lead to the improved use of data coming from the national assessement implemented by INVALSI, reinforcing the linkage between micro-practices and macro-policies. Connecting the topic of standardized assessment to online tools in teacher professional development seems necessary, considering the still rare use of ICT in the evaluation process and its effectiveness. It must be noted that teachers, especially in Italian primary schools, are mainly middle-aged women, only partially exposed to these new digital developments. On the research side, scholars should investigate more deeply the phenomenon we detected, focussing on pupil response style, behaviour, and emotions during the assessment. In addition, expanding the analyses to further grades should not lead to the same "socialization effect". However, it must be kept in mind that we rely on a sample of respondent teachers who may be more in favour of standardized assessment than those who did not participate in the INVALSI survey. Additional investigation to better understand teacher behaviour would be to explore how assessment practices combine (i.e.,

through cluster analyses). In order to develop robust analyses, a promising strategy seems to be getting larger samples of teachers, adding up INVALSI data for school years 2017/18 and 2018/19. Moreover, for future research it would be useful to go beyond the current version of the questionnaire scale used by INVALSI. More precisely, it would be necessary to identify more detailed items regarding assessment practices, in order to catch more finely promising teacher behaviours.

References

Andrade H.L., Heritage M. (2018), *Using formative assessment to enhance learning, achievement, and academic self-regulation*, Routledge, New York.

Baird J.A., Hopfenbeck T.N., Newton P., Stobart G., Steen-Utheim A.T. (2014), *State of the field review assessment and learning (case number 13/4697)*, Knowledge Center for Education, Oslo.

Birenbaum M., DeLuca C., Earl L., Heritage M., Klenowski V., Looney A., Smith K., Timperley H., Volante L., Wyatt-Smith C. (2015), "International trends in the implementation of assessment: Implications for policy and practice", *Policy Futures in Education*, 13, pp. 117-140.

Black P., Wiliam D. (1998), *Inside the black box*, Kings College, London.

Boardman A.G., Woodruff A.L. (2004), "Teacher change and "high stakes" assessment: What happen to professional development?", *Teaching and Teacher Education*, 20, pp. 545-557.

Brookhart S.M. (2011), "Educational assessment knowledge and skills for teachers", *Educational Measurement: Issues and Practices*, 30, pp. 3-12.

Chan C.K.Y., Luo J. (2020), "An exploratory study on teacher assessment literacy: Do novice university teachers know how to assess students' written reflection?", *Teachers and Teaching*, 26, pp. 214-228.

Cowie B., Cooper B. (2016), "Exploring the challenge of developing student teacher data literacy", *Assessment in Education: Principles, Policy & Practice*, 24, pp. 147-163.

Darling-Hammond L. (2017), "Teacher education around the world: What can we learn from international practice?", *European Journal of Teacher Education*, 40, pp. 291-309.

DeLuca C. (2012), "Preparing teachers for the age of accountability: Toward a framework for assessment education", *Action in Teacher Education*, 34, pp. 576-591.

DeLuca C., Bellara A. (2013), "The current state of assessment education aligning policy, standards, and teacher education curriculum", *Journal of Teacher Education*, 64, pp. 356-372.

DeLuca C., Johnson S. (2017), "Developing assessment capable teachers in this age of accountability", *Assessment in Education: Principles, Policy & Practice*, 20, pp. 107-126.

Farrell C.C., Marsh J.A. (2016), "Contributing conditions: A qualitative comparative analysis of teachers' instructional responses to data", *Teaching and Teacher Education*, 60, pp. 398-412.
Filsecker M., Kerres M. (2012), "Repositioning formative assessment from an educational assessment perspective: A Response to Dunn & Mulvenon (2009), *Practical Assessment, Research & Evaluation*, 17, pp. 1-9.
Fullan M., Quinn J., McEachen J. (2018), *Deep learning: Engage the world to change the world*, Corwin, Thousand Oaks.
Gardner J., Harlen W., Hayward L., Stobart G., Montgomery M. (2010), *Developing teacher assessment*, McGraw Hill, London.
Hanushek E., Rivkin S. (2006), "Teacher quality", in *Handbook of the Economics of Education*, Elsevier, Amsterdam, vol. 2, pp 1051-1078.
Hardgreaves A. (2021), "Large-scale assessments and their effects: The case of midstakes tests in Ontario", *Journal of Educational Change*, 11, pp. 1-28.
Hattie J. (2009), *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*, Routledge, London.
Hopfenbeck T.N. (2015), "Formative assessment, grading and teacher judgement in times of change", *Assessment in Education: Principles, Policy & Practice*, 22, pp. 299-301.
Hopster-den Ottera D., Woolsb S., Eggena T.J.H.M.,Veldkamp B.P. (2017), "Formative use of test results: A user's perspective", *Studies in Educational Evaluation*, 52, pp.12-23.
Kingston N., Nash B. (2011), "Formative assessment: A meta-analysis and a call for research", *Educational Measurement: Issues and Practice*, 30, pp. 28-37.
Klieger A. (2016), "Principals and teachers: Different perceptions of large-scale assessment", *International Journal of Educational Research*, 75, pp. 43-145.
Lewis S., Lingard B. (2015), "The multiple effects of international large-scale assessment on education policy and research", *Discourse: Studies in the Cultural Politics of Education*, 36, pp. 621-637.
Looney A., Cumming J., van Der K., Haris K. (2018), "Reconceptualizing the role of teachers as assessors: Teacher assessment identity", *Assessment in education: Principle, Policy & Practice*, 25, pp. 442-467.
Mandinach E.B., Gummer E.S. (2012), *Navigating the landscape of data literacy: It is complex*, WestEd/Education Northwest, Washington, DC/Portland.
Mandinach E.B., Jimerson J.B. (2016), "Teachers learning how to use data: A synthesis of the issues and what is known", *Teaching and Teacher Education*, 55, pp. 452-457.
Marsh J.A., Bertrand M., Huguet A. (2015), "Using data to alter instructional practice: The mediating role of coaches and professional learning communities", *Teachers College Record*, 117 (4), pp. 1-41.
McMillan J.H. (2017), *Classroom Assessment. Principles and Practice That Enhance Student Learning and Motivation*, Pearson, New York, 7th ed.
OECD – Organization for Economic Cooperation and Development (2018), *Education at a glance 2018: OECD indicators*, OECD Publishing, Paris.

Popham W.J. (2018), *Assessment literacy for educators in a hurry*, ASCD, Alexandria.
Popham W.J. (2009), "Assessment literacy for teachers: Faddish or fundamental?", *Theory into Practice*, 48, pp. 4-11.
Popham W.J. (2017), *Classroom assessment. What teachers need to know*, Pearson Education, Boston, 8th ed.
Remesal A. (2007), "Educational reform and primary and secondary teachers' conceptions of assessment: the Spanish instance. Building upon Black and Wiliam (2015)", *The Curriculum Journal*, 18, pp. 27-38.
Schildkamp K. (2019), "Data-based decision-making for school improvement: Research insights and gaps", *Educational Research*, 61, pp. 257-273.
Schleicher A. (2018), *World class: How to build a 21st-century school system*, OECD, Paris.
Shavelson R.J., Young D.B., Ayala C.C., Brandon B.R., Furtak E.M., Ruiz-Primo M.A., Tomita M.K., Yin Y. (2008), "On the impact of curriculum-embedded formative assessment on learning: a collaboration between curriculum and assessment developers", *Applied Measurement in Education*, 21, pp. 295-314.
Smith W.C. (*ed.*) (2016), *The global testing culture. Shaping education policy, perceptions, and practice*, Symposium Book, Oxford.
Stiggins R.J. (2017), *The perfect assessment system*, ASCD, Alexandra.
Zhao Y. (2016), *Counting what counts: Reframing education outcomes*, Solution Tree Press, Bloomington.

The authors

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

Barbara Balconi, researcher at the "Riccardo Massa" Department for Human Sciences at the University of Milano-Bicocca, Academic discipline: Area M-PED/03 – methodologies of teaching and special education. Her main research interests are: teaching and learning strategies; teacher education; general didactics and disciplinary didactics. She is member of a national research center (CERTP) for Educational Research on Teachers as Professionals.

Giovanna Brancato, researcher in Istat since 1996. Since mid-2020, she works in the sector of education statistics where she is responsible of coordinating the planning and development of the new Statistical Register on Education and Training. Previously, she has been in charge of coordinating ISTAT quality function.

Donatella Grassi, researcher in ISTAT since 1996. She has carried out research in various fields. She is currently in charge of the Education and Training sector as responsible of the systematization of the information sources with the aim of harmonizing the processes of data acquisition and dissemination.

Daniela Maccario, Associate Professor at Department of Philosophy and Educational Sciences, University of Turin, Academic discipline: Area M-PED/03 – methodologies of teaching and special education. Her research focuses on Competency based education in formal and non-formal educational contests; teachers and social educators education and training; research and development of didactic models; analysis of didactic-educational practices.

Lorenzo Maraviglia, PhD in Sociology (title obtained at the University of Pisa). In the last twelve years I have directed the statistical office of the province of Lucca and I have participated in various research projects in the social and statistical fields promoted by ISTAT and various research bodies. Actually I work as a statistical researcher at INVALSI.

Serafina Pastore, PhD in 'Instructional Design and Educational Assessment' and Fulbright Research Fellow, is associate professor in the Research and Humanities Innovation Department at the University of Bari (Italy). Her research examines the complex intersections of assessment practice, teacher education and educational policy as operating within the context of school and university innovations.

Elisa Truffelli, PhD, is Associate Professor at the "Giovanni Maria Bertin" Education Studies Department, University of Bologna. Academic discipline: Experimental Pedagogy. Elisa Truffelli's research focuses on Evaluation and Assessment for Learning at different school levels, from Pre-school to Higher Education.

 ISBN 9788835167938

 ISBN 9788835167938

Education is a crucial issue for the economic, cultural and social development of all countries. International empirical evidence has actually highlighted the close correlation between school learning and the economic growth of a country, reaffirming the value of education as a central development lever. INVALSI data, therefore, in addition with other data sources – such as the ones of the Ministry of Education and Merit – may be precious in order to know, better understand, and deepen school processes. INVALSI tests, in particular, provide both educational institutions and policy makers with a wide range of useful and important information to respond to different school needs.

This volume collects some research papers presented at the sixth edition of the Seminar "INVALSI data: a tool for research and teaching" (held in Rome from 25th to 28th November 2021) aimed to explore this topic.

As the Statistical Service, we hope that these examples of the use of INVALSI data, either alone or integrated with other databases, will reinforce the belief that a proper use of data can make an important and crucial contribution to decision-making processes and be a determining factor in making worthy and strong decisions.

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