Planning (and working on) school improvement. Preliminary evidence from the Quality and Merit Project in Italy

Authors:

Andrea Caputo
Doctor
National Institute for the Evaluation of the Educational System of Instruction and Training (INVALSI)

Sara Mori
Ph.D
National Institute of Documentation, Innovation and Education Research (INDIRE)

Valentina Rastelli
Doctor
National Institute for the Evaluation of the Educational System of Instruction and Training (INVALSI)

Abstract

This paper provides preliminary evidence from the Quality and Merit Poject (PQM PON), an Italian in-service training program addressed to lower secondary school teachers which supports improvement plans (PdMs) and offers didactic alternative solutions in math teaching.

This study aims to evaluate the effectiveness of PQM on student math achievement and to explore the association between characteristics/contents of PdMs and some illustrative variables at school level, that are geographical area, improvement level in student math achievement and socio-economic status (SES).

The sample is composed of 248 lower secondary schools of Southern Italy regions, which includes 13816 students participating in the project in 2009/2010 e 2010/2011 school years. Pre-post standardized tests are used to assess the improvement of student math achievement and text analysis of PdMs is carried out in order to detect some school differences in planning strategies.

Results show student improvement in math achievement ($p < 0.01$), also controlling for SES and geographical area. The PDMs associated to better school improvements are those in which the schools have been able to carry out a more careful analysis in terms of context and detection of improvement goals and have been able to prioritize the various elements already in the diagnostic part of the process.

Key-words: school improvement, improvement plan, student achievement

Country Involved In The Program Under Assessment

PQM PON[1] (Italian acronym for National Plan for Quality and Merit) is an Italian in-service training program which aims to provide lower secondary school teachers some innovative teaching materials in order to enhance student achievement in math. It is a joint endeavour of the Ministry of Education, the National Institute of Documentation, Innovation and Education Research (INDIRE) and the National Institute for the Evaluation of the Educational System of Instruction and Training (INVALSI). It is addressed to the teaches of lower secondary school in the Southern Italian regions having access to the European Union funds for low income EU regions (Campania, Sicily, Calabria and Apulia).

The program is not intended to be a traditional content-focused training program, but it provides teachers with polyvalent training offering diagnosis instruments, didactic planning skills, and didactic materials. The teachers participating in the project are part of a network of schools coordinated by a tutor, who gives them both formal and online training, all along the school year. The training has two main goals: 1) help teachers to set up a Piano di Miglioramento (Improvement Plan, from now on PdM), based on student results in standardized test prepared by INVALSI and administered at the beginning of school year; 2) provide teachers alternative solutions to teach the usual curricular contents by using elements such as didactic material, team-work, and lab activities.

The drafting of the PdM is the most important moment of the training, because it determines the number and the type of remedial activities on which the teachers will be then trained. By setting up the PdM teachers should thus identify the skills they would need to

[1] The project is financed by the EU funding - PON Istruzione 2007-2013 (A-2-FSE-2009-2). Special thanks go to INDIRE (National Institute of Documentation, Innovation and Education Research) and especially to Samuele Calzone and Nicola Malloggi for the support provided with the recollection of the data necessary for the analyses. The opinions expressed in the article are those of the authors and do not represent an official position of INVALSI or INDIRE.

acquire both in didactic planning and teaching. The structure of PdM is organized in three sections:

1) analysis of the educational context, that is a fundamental step to plan effective and specific improvement interventions. It should be addressed to two levels. At school level, it should describe teaching organization and parental involvement in both the project and student learning more in general. At PQM class level, it should highlight classroom climate and student motivation with specific regard to math.

2) diagnosis of student needs, that are detected by the INVALSI assessment of math achievement deficits. The diagnosis should be integrated also with information on class background and ordinary teaching. In detail, it should identify both weak and strong points related to student cognitive processes and learning subject areas.

3) detection of improvement goals for planning specific and detailed activities. The main improvement goals deal with: remedy/empowerment of student education, teacher professional development and parental involving in school activities.

The activities that teachers can implement fall mainly in three categories:

- remedial and extra education outside the regular school time (15 hours each) with small groups of students (didactic units based on the main subject areas);
- producing new didactic materials;
- opportunities for sharing innovative teaching materials with other colleagues in the school in a sort of professional community (teacher peer-to-peer laboratory sessions).

At the end of the school year, students are tested a second time and the results are used as check of the activities of the current year and as starting point for the drafting of the PdM for the following year.

Aims Of The Study And Theoretical Framework

Many national and local projects focus on the improvement of student achievement, based on the capacity of schools to transform themselves into supportive environments for teacher learning and change. In this regard, high-performing school systems have shown three core competences (Curtis and City 2009): a deep understanding of the core business of facilitating learning; a theory of action for improving instruction, through a concrete vision and an effective line-up of resources; the strategies to stimulate self-assessment in key areas of competence and to build capacity at different levels and stages of development.

In line with the dynamic model of the educational effectiveness, schools which are able to recognize their weaknesses and take actions to improve their policy on aspects associated with teaching and their school learning environment (SLE) can improve their effectiveness status (Creemers and Kyriakides 2010, 2012). Indeed, research has shown that effective school improvement requires school-level processes (Reezigt and Creemers 2005), and teachers are considered an essential lever of change.

At the school level, the research in the Effective School Improvement (ESI) Project (Reezigt 2001) identifies three key elements: improvement culture, processes, and outcomes. The cycle of improvement processes expects five factors/stages: assessment of improvement needs, diagnosis of improvement needs and setting of detailed goals, planning of improvement activities, implementation/evaluation and reflection.

In this sense, schools can play a substantial role in supporting also teacher learning by creating continuous learning opportunities, promoting inquiry and dialogue, encouraging collaboration and team learning, and establishing systems to capture and share learning, in order to promote change as a result of this learning (Opfer et al. 2011). Participative decision-making, teaming, teacher collaboration, an open and trustful climate, cultures which value shared responsibilities, values and tasks, and transformational leadership practices can foster teachers' professional learning in schools (Thoonen et al. 2012) .

In line with what suggested by the dynamic model (Creemers and Kyriakides 2008), PQM supports a whole school approach and school self-evaluation mechanisms for decision making about improvement of policies and actions. Indeed, the philosophy of the PdM is based on the assumption that schools which are able to identify their weaknesses and develop a policy on aspects associated with teaching and the school learning environment are also able to improve the functioning of classroom-level factors and their effectiveness status. The PQM project also gives opportunity for teachers to engage in continuous and sustained learning about their practice in the settings in which they actually work and to confront similar problems with colleagues and other schools. This is an essential principle of a theory of action which provides a through-line to the instructional core, what are the vital activities that need to happen to improve teaching and learning (City et al. 2009). In this sense, PQM supports change knowledge (Fullan 2005) as it shows some key-elements of theory of action, such as focus on motivation, capacity building with a focus on results, learning in context, changing context, a bias for reflective action, tri-level engagement persistence and flexibility in staying the course.

This paper aims at exploring the main features of PQM school improvement plans in relation to student achievement, given the theoretical relevance of them for an effective school practice. Thus, our research question concerns two specific aims:

1) Evaluate the improvement in student math achievement from 2009/2010 to 2010/2011 school year in order to provide a preliminary assessment of the effectiveness of PQM project;

2) Explore the association between characteristics/contents of PdMs and some illustrative variables of schools: geographical area, improvement level in student math achievement, Socio-Economic Status (SES).

Methods And Data Sources

Participants

In this paper we focus on the schools of the four regions of Southern Italy (Calabria, Campania, Apulia and Sicily) that started the PQM project in school year 2009/2010 (with sixth grade classes) and continued it in 2010/2011. Unfortunately, the reliability of the measures related to the entry test in 2009 was very low and only provided us information on the classes involved in the program (and not on the students), so we excluded it. We can use pre and post treatment measures for the second year of implementation. Thus we exclude both schools that participate in the program only in 2009-2010 or in 2010-2011. In more detail, we use pre and post results of the standardized test by INVALSI only for the students (n=13816) participating in PQM activities in both school years; they belong to 504 classes coming from 248 schools.

Data Sources

Data at the school level are provided by the Italian Ministry of Education through INVALSI. Data at the student level are collected directly by INVALSI, through standardized tests in mathematics at sixth (at the end of 2009-2010 school year) and seventh grade (at the end of 2010-2011 school year), the former being the pre-treatment and the latter the post-treatment outcome. The test measures knowledge of the mathematics contents and logical and cognitive processes used in the mathematical reasoning. The PdMs and data of the activities by schools and classes are provided by INDIRE. For each student, student questionnaire was also administered and provides us data of the student individual and socioeconomic characteristics.

Analysis Procedures

Given the nature of the research questions, we address the issue by adopting a mixed-methodology approach, a research paradigm that utilizes and assigns an equivalent status to both qualitative and quantitative components (Tashakkori and Teddlie 1998).

To assess student improvement in math achievement we calculate math test score simply as percentage of corrected answers out the total number of questions and that hence varies between 0 and 100. At this purpose, we use T-test to compare pre-post results (based on school average math score from PQM classes) in the two school years, controlling for regions and socio-economic status (SES). We calculate also the correlation (r coefficient) between school data on PQM intervention (number of didactic activities, school and class size, percentage of PQM students and classes out the total number of the school) and average math scores in order to better understand participation levels and treatment intensity.

Since this paper provides only a preliminary assessment of the effectiveness of PQM project, we will repeat the analyses on twin classes (selected in PQM schools) not participating in PQM program in order to compare them with PQM classes, also by using anchored scores of pre and post math tests that are not yet available.

We analyze PdMs written by schools with text analysis softwares (Lexico3 and T-Lab) focusing on each section (analysis of the context, diagnosis of student needs, detection of goals and activities). Besides, we explore the relationship between textual data of school PdM and some illustrative variables at school level (in our case, region, student improvement in math and SES). Given that illustrative variables need to be categorical, we split the distributions of both math improvement[2] and SES scores into five divisions at the 10^{th}, 25^{th}, 75^{th}, 90^{th} percentiles so to determine different levels for each variable (very high, high, medium, low, very low).

In more detail, we calculate some lexicometric indexes of PdMs in order to gather quantitative and qualitative information from the formal aspects of the texts, such as:

- Corpus dimension (N) in terms of total number of occurrences or word-tokens[3]

[2] In order to determine math improvement we calculate the difference between the school average math scores of 2010/2011 and 2009/2010 school years.

[3] Word-tokens are only occurring sequences of letters (graphic forms) taken from the alphabet and isolated by means of separators (blanks and punctuation-marks). Instead, each of the different graphic forms repeated in a text is a word-type.

- Vocabulary dimension (V) in terms of total number of different graphic forms or word-types
- Indexes of lexical richness, such as the Average Word Frequency (the occurrence of each word-type in the whole corpus) and the Type-Token Ratio (the number of type-words out of the total number of token-words)
- Indexes of lexical specificity and density, derived from the number of Hapaxes (word-types that occur only one time in the whole text) divided by the corpus (Lexical Variety) or the vocabulary (Hapax Percentage) dimension.

Computer-aided thematic analysis is also carried out to deepen the specific contents dealt with, this is to detect the main thematic repertoires (cluster analysis) and latent dimensions (multiple correspondences analysis) of PdMs texts. Indeed, thematic analysis allows to explore a representation of textual corpus contents through few and significant thematic clusters, related to different semantic nuclei (Lancia 2004). Each cluster consists of a set of elementary contexts (i.e. sentences) characterized by the same patterns of key-words and can be described through the lexical units (words or lemmas) and the most characteristic variables of the context units from which it is composed. Chi-square test allows to test the significance of a word recurrence within each cluster.

Then, Correspondence Analysis enables to explore the relationship between clusters in bi-dimensional spaces, so to detect the latent factors which organize the main semantic oppositions in the textual corpus. In geometrical terms, each factor sets up a spatial dimension - that can be represented as an axis line - whose center (or barycentre) is the value ‘0’, and that develops in a bipolar way towards the negative (-) and positive (+) end, so that the objects put on opposite poles are the most different, almost like the ‘left’ wing and the ‘right’ wing on the political axes.

The relationship between the detected factors and illustrative variables is evaluated through Test Value, a statistical measure with a threshold value (2), corresponding to the statistical significance more commonly used (p. 0.05) and a sign (-/+) which helps in the understanding of the poles of factors detected through the Correspondence Analysis.

Results And Discussion

Student Improvement Analysis

Concerning the first research question, preliminary analyses limited to PQM classes have already provided some results. Pre-post analysis reveals an increase in PQM student math scores (p. < 0.01). On average students get 4 points percentage in correct answers from 2009/2010 to 2010/2011 school year. This difference remains significant also considering each region. In particular, Apulia has the highest improvement (almost 7 percentage points), Calabria shows the minimum one instead (close to 0 percentage points) (Table 1).

Table 1 – Pre-Post Measures Of Math Test Score (School Average Score)

		N	Mean	Std. Dev.	Standard error of the Mean
	Post	248	55.723054*	16.2291902	1.03055460
	Pre	248	51.385562*	9.9005092	0.6286829
REGION:					
Calabria	Post	32	51.136510*	14.7620674	2.6095894
	Pre	32	51.389127*	10.7608056	1.9022596
Campania	Post	82	54.626057*	16.0519628	1.7726427
	Pre	82	52.048199*	9.7763957	1.0796222
Apulia	Post	69	59.695688*	15.8234396	1.9049182
	Pre	69	52.730834*	7.9771509	0.9603361
Sicily	Post	65	55.147846*	17.0118938	2.1100657
	Pre	65	49.119805*	11.2206315	1.3917480

**Pre-post difference (2011-2010) is statistically significant (p<0.01).*

Participation and treatment intensity (number of didactic units, school and class size, percentage of PQM students and classes out the total of the school) has no relation with achievement. We find a correlation between class average SES and each math scores (for 2009/2010 and 2010/2011) (p. < 0.05) but not with the gap scores.

Table 2 - Correlations Between SES And Math Test Scores (2011 School Average; 2010 School Average; 2011-2010 Difference)

		SES
2011 Math Score	Pearson correlation	.141
	Sig.	.027
	N	248
2010 Math Score	Pearson correlation	.159
	Sig.	.036
	N	248
2011-2010 Math Score difference	Pearson correlation	.020
	Sig.	.750
	N	248

In this regard, results confirm the association between SES and student math achievement, when considering a single school year, in line with national (INVALSI 2009; 2010) and international data (OECD 2010), since socioeconomic background is widely recognized as an important contributor to student and school achievement (Coleman 1996; Sirin 2005). Instead, SES doesn't seem to affect student improvement in math achievement, when considering the gap score (between 2010/2011 and 2009/2010 school years), although an "incremental effect" of SES on student improvement may probably need a longer time range, and not just one school year.

A limitation of this study concerns the lack of anchorage measures of math tests and of a control group. As before mentioned, we intend thus to check the robustness of our findings and better estimate the size of improvement by using anchorage measures and not-PQM classes as controls.

Lexicometric Analysis Of PdMs

Our general corpus is composed of 248 texts and includes a total of 494538 word-tokens (N) and 51442 word-types (V).

Looking at each PdM text section (Table 3), the Type-Token Ratio is less than 20% and the Hapax percentage is less than 50% , hence it is possible to state the consistence of a statistical approach (Bolasco 1999). The comparison of the different corpora shows that , overall, *Analysis of the context* is longer and also characterized by higher lexical richness and variety, differently from *Diagnosis of student needs* that uses a repetitive (although detailed) vocabulary and from *Detection of improvement goals* whose lexicon is sufficiently rich but too generic.

Table 3 - Lexicometric Indexes Of Pdm Sections

PdM section	**Word-Tokens (N)**	**Word-Types (V)**	**Hapaxes (V_1)**	**Average Word Frequency (N/V)**	**Hapax Percentage (V_1/V)*100**	**Lexical Variety (V_1/N)*100**	**Type/Token Ratio (V/N)*100**
Analysis of the context	173413	11814	5342	14.68	45.22	3.08	6.81
Diagnosis of student needs	159560	6245	2488	25.55	39.84	1.56	3.91
Detection of improvement goals	161565	8776	2488	18.41	28.35	1.54	5.43

Using illustrative variables as text partition keys, we can compare PdMs' lexicometric indexes among different geographic areas (Regions), SES and achievement improvement levels[4] (Table 4). In sum, Calabria is the Region whose PdMs are generally characterized by greater richness and detail, whilst Campania PdMs tend to be slightly more stereotyped. SES is not significantly associated to any measure. Then, the PdMs of schools with very high student achievement improvement provide a more accurate analysis of the context and a greater originality and specificity of improvement goals, whilst the PdMs of schools with very low student achievement improvement show a more precise and articulated diagnosis of student needs.

[4] Cautious interpretation is needed regarding medium levels of SES and achievement improvement because this category includes 50% of schools out of the total (from 25th to 75th percentile). This means that lexicometric measures could be more unstable because of the greater corpus dimension.

Table 4 - Lexicometric Indexes Of Pdms Partitions By Illustrative Variables

Illustrative variables	Word-Tokens (N)	Word-Types (V)	Hapaxes (V_1)	Average Word Frequency (N/V)	Hapax Percentage (V_1/V)*100	Lexical Variety (V_1/N)*100	Type/Token Ratio (V/N)*100
Analysis of the context							
Geographic area (Region)							
Calabria	20436	3663	1857	5.58	50.70	9.09	17.92
Campania	61699	7064	3385	8.73	47.92	5.49	11.45
Apulia	54238	6508	3200	8.33	49.17	5.90	12.00
Sicily	37040	5192	2583	7.13	49.75	6.97	14.02
Socio-Economic Status (SES)							
Very High	20872	3920	2070	5.32	52.81	9.92	18.78
High	22269	3983	2086	5.59	52.37	9.37	17.89
Medium	85308	8318	3912	10.26	47.03	4.59	9.75
Low	26448	4439	2348	5.96	52.89	8.88	16.78
Very Low	18516	3635	1940	5.09	53.37	10.48	19.63
Achievement improvement							
Very High	14659	3248	1829	4.51	56.31	12.48	22.16
High	24008	4095	2119	5.86	51.75	8.83	17.06
Medium	89692	8665	4087	10.35	47.17	4.56	9.66
Low	27961	4476	2292	6.25	51.21	8.20	16.01
Very Low	17093	3566	1941	4.79	54.43	11.36	20.86
Diagnosis of student needs							
Geographic area (Region)							
Calabria	15314	1951	704	7.85	36.08	4.60	12.74
Campania	54716	3603	1432	15.19	39.74	2.62	6.58
Apulia	60582	3710	1446	16.33	38.98	2.39	6.12
Sicily	28948	2633	1005	10.99	38.17	3.47	9.10
Socio-Economic Status (SES)							
Very High	16362	2046	883	8.00	43.16	5.40	12.50
High	20314	2496	1139	8.14	45.63	5.61	12.29
Medium	83986	4264	1604	19.70	37.62	1.91	5.08
Low	25134	2482	967	10.13	38.96	3.85	9.88
Very Low	13764	1959	821	7.03	41.91	5.96	14.23
Achievement improvement							
Very High	15857	1699	713	9.33	41.97	4.50	10.71
High	24885	2362	885	10.54	37.47	3.56	9.49
Medium	77727	4553	1875	17.07	41.18	2.41	5.86
Low	26954	2628	1058	10.26	40.26	3.93	9.75
Very Low	15137	1948	839	7.77	43.07	5.54	12.87
Detection of improvement goals							
Geographic area (Region)							
Calabria	25176	3401	1451	7.40	42.66	5.76	13.51

Campania	62965	5547	2153	11.35	38.81	3.42	8.81
Apulia	39454	4385	1907	9.00	43.49	4.83	11.11
Sicily	33970	3785	1408	8.97	37.20	4.14	11.14
Socio-Economic Status (SES)							
Very High	16042	2987	1409	5.37	47.17	8.78	18.62
High	26113	3968	1886	6.58	47.53	7.22	15.20
Medium	81868	6306	2509	12.98	39.79	3.06	7.70
Low	23528	3170	1324	7.42	41.77	5.63	13.47
Very Low	14014	2557	1186	5.48	46.38	8.46	18.25
Achievement improvement							
Very High	10046	2119	1051	4.74	49.60	10.46	21.09
High	23321	3534	1609	6.60	45.53	6.90	15.15
Medium	80522	6185	2385	13.02	38.56	2.96	7.68
Low	29913	3873	1626	7.72	41.98	5.44	12.95
Very Low	17763	3234	1559	5.49	48.21	8.78	18.21

Thematic Analysis Of PdMs

Analysis of the context

The analysis detects four thematic clusters of which we report both some of the most characteristic lemmas (Table 5) and examples of elementary context units, indicating their percentage out of the total (Table 6).

Table 5 - Analysis Of The Context. The Most Characteristic Lemmas In Each Cluster (Chi-Square)

Cluster 1		Cluster 2		Cluster 3		Cluster 4	
Lemma	**Chi-square**	**Lemma**	**Chi-square**	**Lemma**	**Chi-square**	**Lemma**	**Chi-square**
Difficulty	162.90	Laboratory	138.13	Territory	387.58	Population	126.44
Study	60.20	Teacher	98.20	Social	194.50	Family	82.31
Mathematics	52.12	Training	71.69	Cultural	166.08	Immigrant	78.62
Ability	48.94	Multimedial	67.68	Context	111.33	Poverty	24.50
Remedy	44.58	Tool	64.98	Offer	107.21	Social Class	16.33
Lacking	41.99	Assessment	63.31	Users	101.14	Unemployment	15.55
Knowledge	41.01	Project	62.56	Parish	67.70	Service Industry	15.12
Learning	40.73	Information Technology	39.32	Resource	65.06	Economic	13.58
Improvement	36.79	Experimentation	18.31	Educational	62.02	Community	11.43
Competence	20.30	Workshop	18.24	Agency	59.03	Ethnic Group	10.25

Table 6 - Analysis Of The Context. Examples Of Elementary Context Units Of Each Cluster (Percentage)

Cluster 1 (31.84%)
Participating in the project may be an opportunity to consolidate and develop logical mathematical knowledge and skills for some students, to remedy basic abilities for other ones instead'.
Cluster 2 (18.03%)
PQM project tutor in the school department has provided teacher training on ICT, with specific regard to the use of didactic softwares. The school offers several laboratories: information technology, scientific, linguistic and musical'.
Cluster 3 (18.44%)
'The school is set in a densely populated suburb, recently characterized by an indiscriminate council housing and lack of suitable community services. This territory offers few meeting places for youth. The only educational agencies are family, school and parish'.
Cluster 4 (31.68%)
'The working population is heterogeneous: professionals, office workers, shopkeepers, artisans, casual workers and unemployed; it follows that the socio-economic conditions of children attending the school are different and the level of education of their families is lower middle. There are aslo poor pupils who live in residential children's homes, and some without families'.

Cluster 1: Student ability and performance.

It deals with student characteristics regarding study method and basic abilities. In more detail, it shows a specific focus on general learning difficulties of class and the need of improving student performance.

Cluster 2: School resources.

It highlights school educational and organizational resources, also in terms of didactic tools (availability of school equipment, laboratories, ICT, etc.). This cluster is associated with a tendency to invest on student assessment, teaching innovation and experimentation, and to create a teacher professional community inside the school.

Cluster 3: External socio-educational agencies.

It focuses on the relationship between school and external socio-educational agencies (parish, youth associations, social services). Since youth education is seen as a shared responsibility with territorial partners outside the school, the context is given a central role in supporting school efficacy.

Cluster 4: Family background and social context.

This cluster mainly relies on the description of student family background and origin. It also refers to wider social, cultural and economic context and some critical issues (immigration, youth problems, unemployment, poverty) which are likely to affect student education and development.

Correspondence Analysis enables to explore the relationship between the four thematic clusters detected in a bi-dimensional space (Figure 1). Thus it allows to analyze the latent factors which organize the main semantic oppositions in the textual corpus, from the different position of clusters on the first two factorial axes - as indicated by Test Values (Table 7) - which explain about 86% of total inertia.

Figure 1 - Analysis Of The Context. Factorial Space

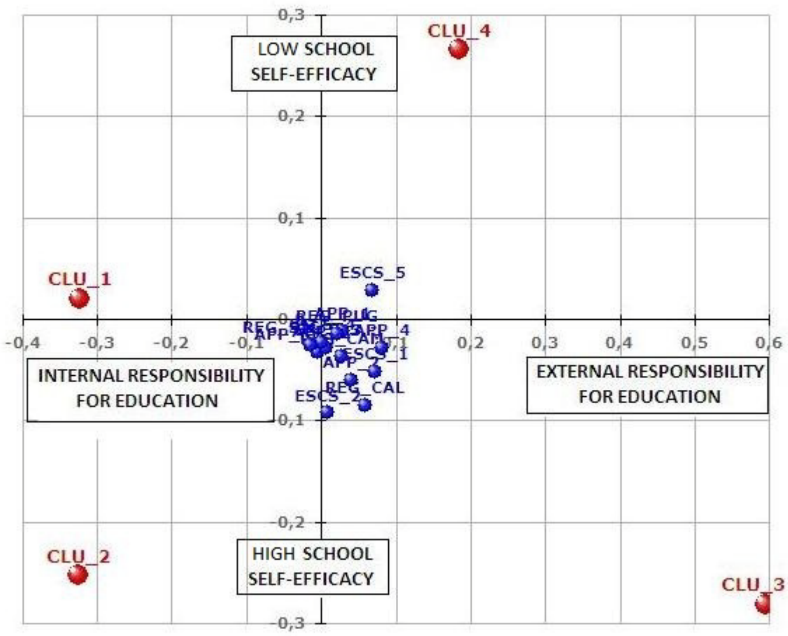

Table 7 - Analysis Of The Context. Relation Between Clusters And Factors (Test Value)

	Factor 1	Factor 2
CLU 1	-50.33	3.11
CLU 2	-33.98	-26.53
CLU 3	60.86	-28.70
CLU 4	29.12	41.55

In more detail, the first latent factor (horizontal axis) seems to refer to *school responsibility for student education* (61.11% of the variance). On the negative pole, the analysis of the context focuses on the role of the school in improving both student achievement (Cluster 1) and teaching methods (Cluster 2). Instead, on the positive one, there is a scarce internal commitment because schools tend to delegate their institutional function to other external socio-educational agencies (Cluster 3) or to families and wider social context (Cluster 4).

The second latent factor (vertical axis) expresses *school self-efficacy* (23.44% of the variance): the negative pole is associated to the school perception of internal (Cluster 2) and external (Cluster 3) resources that can support its educational aims and efficacy; on the contrary, the positive pole is mainly represented by the perception of powerlessness in front of unchangeable characteristics of student environment (Cluster 4).

Analyses also show some associations between latent factors and illustrative variables: schools with low achievement improvement level mainly attribute responsibility for student education to the social context (Test Value=2.18); whilst schools with high SES perceive themselves effective in education improving (Test Value=-2.21).

Diagnosis of student needs

The analysis detects five thematic clusters of which we report both some of the most characteristic lemmas (Table 8) and examples of elementary context units, indicating their percentage out of the total (Table 9).

Table 8 - Diagnosis Of Student Needs. The Most Characteristic Lemmas In Each Cluster (Chi-Square)

Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5	
Lemma	**Chi-square**	**Lemma**	**Chi-square**	**Lemma**	**Chi-square**	**Lemma**	**Chi-square**	**Lemma**	**Chi-square**
School	651.12	Mathematics	221.68	Activity	633.64	Answer	201.76	Master	404.66
Score	365.76	Information	189.39	Remedy	442.52	Correct	97.49	Representation	233.85
Region	324.28	Use	178.68	Improvement	358.84	Question	72.54	Algorithm	169.74
Result	256.56	Tool	125.52	Planning	313.54	Level	63.35	Procedure	166.09
Achievement	206.83	Scientific	121.00	Proposal	61.59	Item	45.20	Thinking	164.87
South	201.17	Description	119.45	Qualification	49.74	Percentage	44.78	Symbolic	120.19
Test	189.43	Solve	116.49	Didactics	41.13	Gap	30.11	Argument	119.51
National	180.37	Detect	116.49	Subject	32.73	Weakness	27.43	Problem	117.59
Average	157.46	Explain	116.26	Curriculum	\17.65	Output	16.45	Arithmetic	93.65
Comparison	84.61	Quantitative	114.50	Performance	6.68	Error	5.83	Process	56.39

Table 9 - Diagnosis Of Student Needs. Examples Of Elementary Context Units Of Each Cluster (Percentage)

Cluster 1 (9.93%)

'In order to formulate the diagnosis for the class, we analyze results from the test administered in May 2010. The class scored 51.9% of correct answers, in line with our school, PQM schools and national average, and higher than the regional and Southern Italy one'.

Cluster 2 (23.10%)

'Use learned math to analyze information through quantitative data in scientific, technological, economic and social fields (describe a phenomenon in quantitative terms, interpret it with statistical tools or functions, use math models)'.

Cluster 3 (15.04%)

'I propose to plan some remedial didactic units concerning *Arithmetic* and *Geometry* and some empowering units on *Relations and Functions* already used last year'.

Cluster 4 (34.42%)

'There are low percentages of correct answers to the questions n. 14 and n. 19, which concern the subject area *Data handling*. Another negative element is the high rate of missing answers to the question n. 25 dealing with *Arithmetic'*.

Cluster 5 (17.51%)

'The class seems to have greater difficulties in two subject areas, *Arithmetic* and *Data Handling*, and in the following processes: Problem solving by math tools; Know and master various forms of representation; Progressively acquire typical forms of math thinking'.

Cluster 1: School achievement.

This cluster is associated to the use of results from INVALSI standardized test, mainly focusing on school ranking at regional and national level in the wider context. General information is derived on math score at school or class level, without further reference to math subject areas or in-depth analysis on students.

Cluster 2: Student ability in using math.

This cluster is completely focused on one of the cognitive processes measured by INVALSI standardized test, which deals with the student ability in using math for analyzing quantitative information and interpreting reality. In this sense, it seems to emphasize a competence-based approach for student assessment rather than a knowledge-based one.

Cluster 3: Curriculum-based information.

This cluster highlights the importance of curriculum-based information to diagnose student needs. In this regard, teacher experience and continuity in education are some key elements to plan effective goals, consistently with the usual teaching practice.

Cluster 4: Detailed analysis of math test.

This cluster refers to the detailed analysis of student results deriving from INVALSI standardized test. The external assessment is accurately used as the main source to detect student deficits, from a multi-focused and analytical view.

Cluster 5: Cognitive processes and subject areas.

In this cluster student needs are detected from test results concerning both cognitive processes and subject areas measured in math test. In this sense, the use of external assessment allows to identify priorities and specific learning skills to address in planning next activities.

The analysis detects two factorial axes (Figure 2) which overall explain 80.53% of the total inertia and are differently associated to the five thematic clusters (Table 10).

Figure 2 – Diagnosis Of Student Needs. Factorial Space

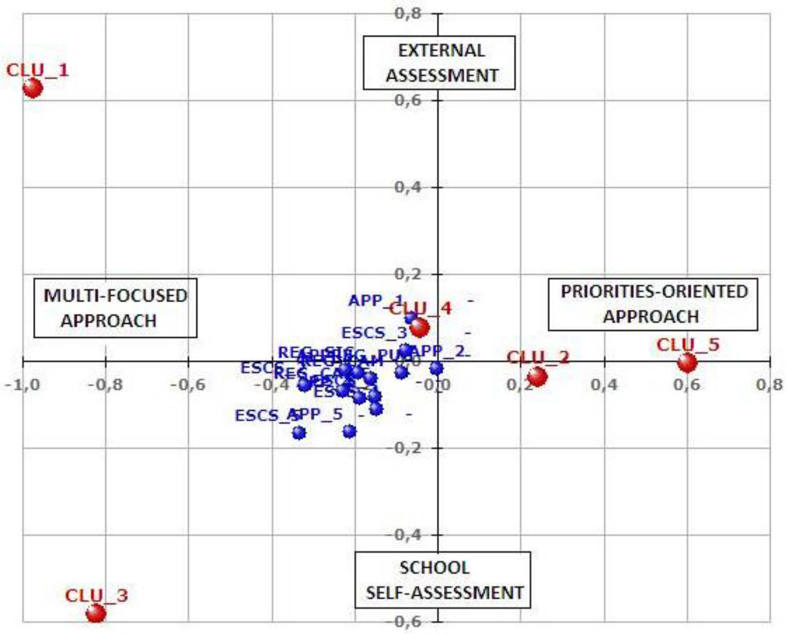

Table 10 – Diagnosis Of Student Needs. Relation Between Clusters And Factors (Test Value)

	Factor 1	Factor 2
CLU 1	-72.46	46.53
CLU 2	36.31	-5.50
CLU 3	-72.95	-51.96
CLU 4	-7.38	13.63
CLU 5	75.72	-0.88

The first latent factor refers to the specific *utilization of standardized test* (61.97% of the variance) for the diagnosis of student needs. On the negative pole, the diagnosis relies on a multi-focused approach that is mainly curriculum-based (Cluster 3) and fail to use test results correctly, because the analysis of the test is more school than student-centered (Cluster 1) or is overdetailed but without identifying key-elements to improve (Cluster 4). Instead, on the positive pole, more attention is paid to detect specific student skills (Cluster 2) or learning processes (Cluster 5) as priorities to enhance.

The second latent factor focuses on the degree of *integration between internal and* *external assessment* (18.55% of the variance). It mainly opposes a self-assessment approach - negative pole - giving greater importance to the curriculum and teaching processes inside the school (Cluster 3) to an external assessment - positive pole - which provides objective data and allows to compare school achievement with other schools in the wider context (Cluster 1).

The results on the association between factors and illustrative variables show that the schools with low (Test Value= -2.71) or very low (Test Value= -3) SES tend to highlight the centrality of self-assessment processes and teaching practice inside the school in order to detect student needs, rather than using external assessment. It is also true for the schools with low (Test Value= -2.23) or very low (Test Value= -2.96) student achievement improvement that, in addition, mainly rely on a multi-focused diagnosis and fail to use test results correctly (Test Value=-4.69 for low level; Test Value=-3.81 for very low level).

Detection of improvement goals

The analysis detects five thematic clusters of which we report both some of the most characteristic lemmas (Table 11) and examples of elementary context units, indicating their percentage out of the total (Table 12).

Table 11 - Detection Of Improvement Goals. The Most Characteristic Lemmas In Each Cluster. (Chi-Square)

Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5	
Lemma	**Chi-square**	**Lemma**	**Chi-square**	**Lemma**	**Chi-square**	**Lemma**	**Chi-square**	**Lemma**	**Chi-square**
Work	269.71	Figure	454.69	Strategy	169.39	Represent	187.18	Class	318.56
Practice	160.95	Geometry	165.72	Instrument	159.76	Interpret	178.25	Student	260.73
Tutor	119.14	Space	78.96	Identify	134.93	Master	158.40	Remedial	206.35
Laboratory	115.08	Didactic Unit	76.53	Scheme	122.23	Know	112.35	Empowering	172.39
Activity	84.78	Drawing	61.41	Link	113.39	Describe	104.97	Improvement	134.13
Time	84.58	Calculate	36.48	Usefulness	90.40	Meaning	66.97	Involving	92.05
Cooperative	81.25	Play	20.30	Resolving	80.19	Learn	46.20	Goal	43.90
Method	81.18	Handle	18.71	Measure	74.26	Codify	46.13	Responsibility	34.33
Meeting	58.11	Dynamic	10.62	Operational	57.64	Understand	31.49	Self-evaluation	30.40
Innovative	42.98	Writing	9.86	Explore	35.14	Cognitive	10.48	Skill	30.14

Table 12 - Detection Of Improvement Goals. Examples Of Elementary Context Units Of Each Cluster (Percentage)

Cluster 1 (8.55%)
Laboratory practice promotes exploration, modeling and operational deduction. Cooperative methodology (small working group) activity enhances argumentation about experiences and procedures used. Along the different steps of the work, after individual production it is possible to exchange views and analyze the process'.
Cluster 2 (10.09%)
'The training activity will be centered on Geometry, in particular on the didactic unity "Equi-decomposability of elementary figures: the Tangram" that will allow to deal with the concepts of equi-extension and equivalence of plane figures in order to explain simple literal formulas'.
Cluster 3 (22.73%)
'Problem solving by using math tools (i.e. detect and connect relevant information, compare solution strategies, identify solution schemes such as calculation sequences, explain the solution process)'
Cluster 4 (29.11%)
'Know and master specific contents of mathematics (mathematical objects, properties, structures). Know and master algorithms and procedures (in arithmetic and geometry). Know and master various forms of representation (verbal, written, symbolic, graphical) and know how to switch between them'.
Cluster 5 (29.52%)
'On the basis of this diagnosis, it is appropriate to plan an empowering training activity, addressed to ten students, and two remedial activities concerning Arithmetic and Relations and Functions, each addressed to another group of fifteen students'.

Cluster 1: Training methods.

This cluster mainly focuses on training methods and procedures used for goals implementation plan. In detail, it refers to a teaching approach that is based on laboratory and group working in order to enhance cooperative learning among students, consistently with PQM philosophy on school improvement.

Cluster 2: Didactic units.

This cluster deals with the selection of specific didactic units related to math contents that need to be improved. It thus refers to teaching materials and activities, as concrete and practical dimensions of the experimentation, in close relation with the curriculum.

Cluster 3: Problem-solving.

The focus is on a specific cognitive process of student learning that is a key-element of math teaching experimentation. In detail, it concerns problem-solving seen as the ability of using math tools to solve real life problems in everyday situations.

Cluster 4: Cognitive processes.

This cluster includes several student cognitive processes to which PQM intervention is addressed. These processes are associated to cross-curricular sub-competences that students are asked to know, use and reflect in relation to mathematics.

Cluster 5: Detailed description of activities.

In this cluster more attention is paid to the specific context at class and student level. The main goal of PdM is declared in terms of remedy or empowering education, and some differences are detected among students – based on previous assessment – in order to diversify PQM classroom activities.

The analysis detects two factorial axes (Figure 3) which overall explain 79.12% of the total inertia and are differently associated to the five thematic clusters (Table 13).

Figure 3 - Detection Of Improvement Goals. Factorial Space

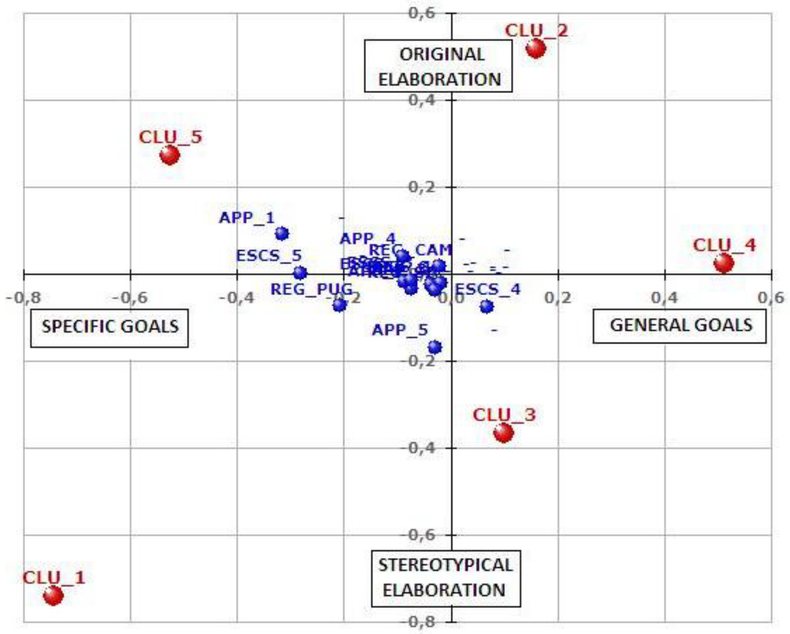

Table 13 - Detection Of Improvement Goals. Relation Between Clusters And Factors (Test Value)

	Factor 1	Factor 2
CLU 1	-47.42	-47.42
CLU 2	13.14	41.82
CLU 3	13.34	-49.17
CLU 4	81.56	3.63
CLU 5	-78.72	40.73

The first latent factor seems to rely on the degree of *specificity of planned goals* (49.87% of the variance). It mainly opposes a specific to a general goal orientation. On the negative pole, there is a greater tendency to explicit educational aims and detail activities (Cluster 5), also focusing on training methods and materials (Cluster 1). Instead, on the positive one, general suggestions are given about student cognitive processes that need to be improved, but without indication on how to do it (Cluster 4).

Then, the second latent factor highlights the degree of *originality in PdM elaboration* (29.25% of the variance). On the negative pole, the focus is on the wider PQM framework concerning a teaching approach that is oriented to collaborative learning (Cluster 1) and problem-solving (Cluster 3). PdMs seem thus to address general key-elements of PQM activities, that are reported in a very stereotypical way[5] without further elaboration. On the contrary, the positive pole is characterized by higher originality and autonomy in setting up PdM because various strategies are accounted in order to select didactic units (Cluster 2) and to diversify classrooms activities (Cluster 5).

The results on the association between factors and illustrative variables show that the schools with very high student achievement improvement tend to report more specific activities and better explicit improvement goals (Value Test=-5.34); whilst the schools with

[5] This stereotypical tendency is also noticeable from the use of exact sentences (elementary context units), derived from official materials and general information on PQM project.

very low improvement are characterized by a stereotypical tendency in PdM elaboration (Value Test=-3.63). Then, Apulia is the region which shows the greatest degree of specificity of planned goals (Value Test=-6.75).

Conclusions

Regardless the identified limitations (need of anchoring of the tests, need of repeating the analyses on control group, unavailability of more waves of data), some evidence exists on the fact that PQM students are improving, as indicated by pre-post analysis on math scores.

Moreover, the PDMs associated to better school improvements are those in which the schools have been able to carry out a more careful analysis in terms of context and reflection on aims and have been able to prioritize the various elements already in the diagnostic part of the process. Thus, it is possible that these schools have then been able to identify and implement more effective improvement activities than what happened in school that were not able to focus on specific needs and tried instead to address simultaneously a multitude of aspects, as suggested by both lexicometric and thematic analysis of PdMs.

In regard to improvement planning strategies at school level, the key-factors which seem to promote better student achievements concern the school capabilities of taking account of the educational context and, above all, detecting specific and detailed improvement goals, as shown by results on Apulia that is also the region with highest student improvement. On the contrary, the main obstacles to school improvement seem to refer to: the tendency to attribute responsibility for student education to external socio-educational agencies, the exclusive focus on school self-assessment to detect student needs, a diagnosis of student learning deficits not based on cognitive processes or specific subject areas to enhance, a poor autonomy and originality in setting up improvement activities. Then, socio-economic background does not seem to be specifically associated to student improvement, but can influence school planning strategies: high-SES schools are likely to perceive themselves more self-efficacy in determining student outcomes; whilst low-SES schools seem to rely on internal evaluation mechanisms to take decisions on how to improve school functioning, without perceiving benefits from external assessment.

Such very preliminary conclusions certainly require further research and reflection; yet, they feed into a stream of discussion that is currently growing in Italy and related to the ability of schools and school staff to use strategically information and autonomy of action for initiating school improvement processes.

References

City, Elizabeth, et al. *Instructional Rounds in Education: A Network Approach to Improving Teaching and Learning*. Cambridge, Massachusetts: Harvard Education Press, 2009.

Coleman, James Samuel, et al. *Equality of Educational Opportunity*. Washington: US Government Printing Office, 1966.

Creemers, Bert, and Leonidas Kyriakides. *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools.* London: Routledge, 2008.

Creemers, Bert, and Leonidas Kyriakides. "Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors", *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice* 21 no.4 (2010): 409-427.

Creemers, Bert, and Leonidas Kyriakides. *Improving Quality in Education: Dynamic Approaches to School Improvement*. London: Routledge, 2012.

Curtis, Rachel and Elizabeth City. *Strategy in Action: How School Systems Can Support Powerful Learning and Teaching*. Harvard Education Press, 2009.

Fullan, Michael. *Leadership and sustainability: System thinkers in action.* Thousand Oaks, CA: Corwin, Toronto, Canada, Ontario Principals Council, 2005.

INVALSI Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione, *Rilevazioni degli apprendimenti. Scuola primaria. Prime analisi.* INVALSI, Rome, 2009.

INVALSI Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione, *Le prove del Servizio Nazionale di Valutazione 2008-2009 - Analisi Tecnica*. INVALSI, Rome, 2010.

Lancia, Franco. *Strumenti per l'analisi dei testi. Introduzione all'uso di T-LAB*. Milano:

FrancoAngeli, 2004.

OECD Organization for Economic Co-operation and Development, *Education at a Glance 2010 Revised edn*. OECD, 2010.

Opfer, Darleen, et al. “The influence of school orientation to learning on teachers' professional learning change”, *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice* 22, no.2 (2011): 193-214.

Reezigt, Gerry, and Bert Creemers. “A comprehensive framework for effective school improvement”, *School Effectiveness and School Improvement: An International Journal of Research. Policy and Practice 16*, no.4 (2005): 407-424.

Reezigt, Gerry. *A framework for effective school improvement. Final report of the ESI project.* Groningen, The Netherlands: GION, Groningen Institute for Educational Research, University of Groningen, 2001.

Tashakkori, Abbas, and Charles Teddlie. *Mixed methodology: combining qualitative and quantitative approaches.* CA: Sage, Thousand Oaks, 1998.

Thoonen, Erik, et al. “Building school-wide capacity for improvement: the role of leadership, school organizational conditions, and teacher factors”, *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practic*e *1,* no.20 (2012): 1-20.

Sirin, Selcuk R. “Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research.” *Review of Educational Research* 75, no.3 (2005): 417–453.