

Accompanying organizational innovation by research: The case for performance evaluation of Italian school principals¹²

Marco Depolo³ & Michela Vignoli⁴

Abstract

In 2015, the Italian parliament approved the law no. 107, named “La Buona Scuola” (“The Good School”), in order to boost the quality of organizational processes in schools. Among these, one of the most innovative was introducing a performance evaluation procedure for the nearly 7000 school principals of Italian public schools, from primary to college. From 2000 on, the legal status of schools principals in the public system have been set to the managerial level. However, no formal performance evaluation had been really performed before.

In 2016, the INVALSI (the governmental agency for the evaluation of the national education system - <http://www.invalsi.it/invalsi/index.php>) was instructed to formulate a project aimed to: i) translate into concrete organizational procedures the goals dictated by the law; ii) train the evaluation teams needed; iii) monitor by an appropriate research design the outcomes of the new performance evaluation.

The paper describes and discusses the training programme for assessors, the research design and some preliminary results.

Keywords: Italian school principals; performance; assessment; INVALSI

1 The research and training activities described have been funded at INVALSI by the EU, through the PON Project “per la Scuola” (“ProDis – Strumenti e metodi per la promozione della professionalità del dirigente scolastico”), and through the PON Project “Misurazione diacronico-longitudinale dei livelli di apprendimento degli studenti”.

2 The views expressed are those of the authors and do not imply any expression of opinion on the part of the INVALSI.

3 Alma Mater Studiorum – Università di Bologna, Italy. Email: marco.depolo@unibo.it - ORCID: <https://orcid.org/0000-0002-4708-2715>.

4 Università di Trento, Italy. Email: michela.vignoli@unitn.it - ORCID: <https://orcid.org/0000-0002-2583-5384>.

Acompanhando a inovação organizacional pela investigação: o caso da avaliação de desempenho de diretores de escola italianos

Resumo

Em 2015, o parlamento italiano aprovou a lei nº 107, denominada “La Buona Scuola” (“A Boa Escola”), com o objetivo de impulsionar a qualidade dos processos organizacionais nas escolas. Entre eles, um dos mais inovadores foi a introdução de um procedimento de avaliação de desempenho para os quase 7 mil diretores de escolas públicas italianas, do ensino primário ao universitário.

Já desde 2000 que o estatuto legal dos diretores de escolas no sistema público ficou definido para o nível de gestão. No entanto, nenhuma avaliação formal de desempenho foi antes realmente realizada.

Em 2016, a INVALSI (a agência governamental para a avaliação do sistema nacional de educação - <http://www.invalsi.it/invalsi/index.php>) foi instruída no sentido de formular um projeto com o objetivo de: i) traduzir em procedimentos organizacionais concretos; os objetivos ditados pela lei; ii) treinar as equipes de avaliação necessárias; iii) monitorizar, por meio de um desenho de investigação apropriado, os resultados da nova avaliação de desempenho.

Este artigo descreve e discute o programa de formação para avaliadores, o desenho da investigação e alguns resultados preliminares.

Palavras-chave: Diretores de Escolas Italianas; desempenho; avaliação; INVALSI

INTRODUCTION

In the year 2015, the Italian parliament approved the law no. 107, also known as “La Buona Scuola” (“The Good School”), in order to boost the quality of both pedagogical and organizational processes in schools. Within this framework, “school” refers to the institutions directly held by the State, and includes primary (6 to 11-year-old pupils), to lower secondary (11 to 14), and upper secondary school (14 to 19).

This paper will concentrate only in the organizational side of the educational process developed by the school, taking into account more in detail the evaluation of the performance of school principals (hereafter, SP), which has been introduced by the aforementioned law.

In Italy, slightly more than 8000 public (i.e. directly managed by the State) school institutions were active in 2017/18. Each school is under the responsibil-

ity of an SP. At the same time, the number of SPs in service was around 6500. It implies that one school out of five was managed by an SP assuming a double (or, in rare instances, triple) position and responsibility, because of the scarce number of SP hired and in service.

The career of SP has two main entry requirements: having at least 5 years of service as a teacher, and passing a regional-based examination. From the year 2000 on, Italian SPs became public managers for law purposes, reporting to a general Director on a regional basis (“Regional Director”). SPs are responsible for all aspects of school functioning, including pupils’ learning results, coordination of teachers’ activities and relationship with local community and institutions.

Among the new features introduced by the aforementioned law 107/2015, one is of central interest for this paper: the new procedure for the performance appraisal of the SPs of Italian public schools. Previously, a formal performance assessment had never been realized. As officially stated, the whole assessment process “aims to professional enhancement and improvement of SPs, to progressively increase the quality of school service” (from the article 3 of the Directive implementing the law 107/2015).

The Ministry of Education was in charge to decide and implement the operational criteria for such performance assessment. It has been stated that the performance of every SP during the previous school year has to be assessed against three main criteria, namely: i) unified management, promotion of participation, managerial competences aimed to get results (accounting for 60% of the final evaluation rating); ii) enhancing and promoting human resources, professional effort and merit (30%); iii) recognition within professional and local community (10%).

1. THE PERFORMANCE ASSESSMENT OF SPS: THE CONTEXT

Organizational theorists have often considered schools as a rather special kind of organization, to be described and studied paying attention to their specificity, as Weick (1976) suggested in his famous paper that used the conceptual frame of “loosely coupled system”, to explain individual and collective behaviours in schools. Following Weick’s approach, a loosely coupled system shows some typical features, namely: i) situations where different means can get the same result; ii) a lack of coordination; iii) absence of regulations; iv) very slow feedback times, within highly connected networks. That is why schools (intended as organizations) would be quite agile and able to respond to local and contingent stimuli, but hard to regulate by tight organizational rules and linkages.

This condition of a loosely coupled system should not be considered a weakness, as if “tight connected” organizational systems would be more reliable and efficient than the loosely coupled ones. Loosely coupled systems – just because of their flexibility – tend to perform better in times of rapid environmental changes, to produce local adaptations more easily, and to allow for more self-determination by actors. However, one more time because of their flexibility, they also need special care in order to reach and maintain a reasonable amount of organizational congruence among their components.

Concerning the role of SP in schools, besides the focus put by researchers on the so-called instructional leadership (Davis, Darling-Hammond, LaPointe, & Meyerson, 2005; Hallinger, Adams, Harris, & Suzette Jones, 2018; Purinton, 2013), since schools still are *also* organizations, many conceptual tools of organizational sciences are applicable. Following this line, one of the most relevant topic for Italian SPs is a possible responsibility-authority gap. As the first example of such a gap, Italian SPs do not have any middle management in their staff, helping them in the everyday organizational life: this is a quite rare situation for a manager being directly responsible for approximately one hundred collaborators (mean 2017/18 of a number of teachers per school, from official Ministry database). The second example of such responsibility-authority gap is that an SP has not any voice in selecting teachers serving in his/her school. SPs have also a limited amount of real influence on teachers’ daily professional behaviour.

In other words, SPs are held responsible for the results of the school, while they have to manage schools more through moral authority than through real power. However, research showed that SP does have an influence on many crucial elements of school life. As Paletta, Alivernini and Manganelli (2017) showed, when the influences of the context where the school operates are controlled for, principal’s leadership actually influences the process variables related to teachers and educational climate: job satisfaction among the teachers, self-efficacy of teachers, quality of educational climate, as well the academic success of students are related to SP’s leadership behaviours.

A complete description and analysis of the role of SP in Italian schools goes beyond the scope of this paper. However, the aforementioned research data and information seem sufficient to show that SP’s performance is a focal point for school effectiveness. Before entering into the main topic of this paper, the reader should consider some specificities of the Italian situation, considering that all data that will be shown concern only public schools, since the private ones are not regulated by the same discipline.

From the year 2000 on, the government has set the legal status of SPs to a managerial level, at the same moment when schools have been declared “autonomous”,

stressing the right and duty to self-organize their effort towards the attainment of educational goals. Consequently, SP's role was redesigned, shifting from a "primus inter pares" role, towards a managerial one.

Within this framework, SPs should have been treated in the same way than other public managers in other areas of Civil Service. Among the many elements to be considered, performance assessment is an important one. Although performance evaluation is mandatory by law for all public managers, the unique characteristics of educational organizations suggested not to automatically translate for SPs the same procedure in use for "standard" administrative managers of the public sector. This choice seems adequate, following good practices in organizational theory and in professional consulting, but in fact, any formal performance evaluation had not been performed for many years before the aforementioned law no. 107 ("The Good School"), which in 2015 dictated the guidelines for the performance assessment of SPs.

One year later, in 2016, the National Institute for the Educational Evaluation of Instruction and Training (hereafter: INVALSI), a governmental agency dedicated to study and research on the whole education system from kindergarten to college, was requested to formulate a project concerning SPs' performance assessment, aimed to:

- i) translate into concrete organizational procedures the guidelines dictated by the law;
- ii) train the teams of assessors needed;
- iii) monitor, by appropriate research design, the whole process.

2. THE PERFORMANCE EVALUATION OF SPS: CONTENT, PROCESS, AND OUTCOMES

As seen before, Italian SPs are public managers. It implies that the assessment of their performance is quite complex, at least for three main reasons.

First, it has been necessary to clearly state who should have been responsible for such an assessment. In organizations, the immediate supervisor is usually in charge of it: in our case, the definition of "immediate supervisor" is not completely applicable. The Italian schools system is organized on a regional basis: 18 Directors of "Regional School Office" are responsible in each of the 18 administrative regions that form the Italian Republic. In principle, in each region, an SP reports directly to the Regional Director, so that the Ministry made the Regional Director responsible for the performance assessment of all the SPs working within the region.

Secondly, following the law and professional good practices, managers have to be assessed not against their organizational behaviours, but against the objectives

they have been assigned to. Managers are expected to behave in order to reach goals, not only in conformity with pre-fixed behavioural standards like punctuality, collaboration with colleagues, and similar. Taking into account the goals assigned to SPs, one should note that a good amount of variability has been found in the nature and format of the goals actually delivered by each regional director to their SPs. And, following the law, every SP has to be evaluated by appreciating “his/her contribution to the attainment of the goals related to the improvement of the service offered by the school [that s/he is managing]”.

Thirdly, this is the first massive performance appraisal campaign implemented within the Italian school system. The national school community did not experience any previous programme, any learned habits, any customary practices of implementation or negotiation between actors about what a performance assessment is, and how to deal with it: the national educational community was confronted with a completely new situation, like in a true textbook case of organizational learning.

Once defined the content of the performance assessment (i.e., measuring the contribution of the SP to the attainment of the assigned goals), the definition of the assessment process was the second challenge. As seen before, the regional Director was responsible of the assessment: however, it should be noted that the ratio of SPs per Director varies between 1:44 (in the smallest region) and 1:922 (in the biggest one). It is quite evident that the regional Director, although formally responsible for the whole process of assessment, would not be able to manage personally every step of this task.

The reader should also consider that regional offices have in average a small amount of staff adequately skilled for this job so that some new teams of assessors had to be established. Drawing on the existing role of “inspector” (regionally located expert collaborators of the Director, mainly used to investigate, help and report to the Regional Director in case of critical events within the schools), a number of a team of assessors have been created. Each team is formed by two SPs (preferably not serving in the same area where they have to act as evaluators), coordinated by one “inspector”.

The tasks of the teams are: i) studying the personal file of the SP they have to evaluate; ii) meeting (preferably by an audio-visual tool like Skype) the SP, to give him a voice in the process; iii) formulating a personalized “feedback for improvement”, for the personal use of the SP; iv) translating the evaluation into a final statement of appraisal, and proposing it to the Regional Director.

As seen before, only the regional Director is legally responsible for the assessment, so that the team *proposes* the final evaluation, which is adopted and/or amended by the Director, who, at the end of the process, signs the final official act that communicates to every SP the evaluation received for his/her performance.

The aforementioned feedback deserves a short comment. Professional literature suggests (cf. Aguinis, 2009; Aguinis, Joo, & Gottfredson, 2011) that the perfor-

mance feedback has an influence on individual and team performance, as well as on worker engagement, motivation, and job satisfaction. For a professional, getting (and using) competent feedback about one's own performance could be considered at the same time duty and a right since competent feedback is a powerful tool to increase the performance quality and improve personal skills. Following this approach, such "feedback for improvement" can be considered as the empirical proof of the fact that the whole process of evaluation was not put in practice only to rate SPs, but also to help them to improve professional skills and competences. Moreover, assessors having to deliver a personalized feedback suggestion are "forced" to become more deeply acquainted of the SP's unique situation (the school where s/he works; the phase of his/her career; the strengths and weaknesses of his/her managerial behaviour).

3. THE TRAINING PROGRAMME FOR THE MEMBERS OF THE ASSESSMENT TEAMS

The first main task for INVALSI has been the training of a very large number of members of the assessment teams. Although the choice of the members remained under the responsibility of the Regional Directors, the training has been designed and implemented on a national basis. In total, more than 300 assessment teams (corresponding to around 800 evaluators, some of them being part of more than one team) were trained in a series of 2-day intensive training seminars, replicated 17 times in different locations.

The goal was to provide a national-based shared model of the procedure, through a participated training, made of some informational inputs, and of a larger amount of small group activities, mainly aimed to become familiar with the procedure, and to anticipate and discuss possible problems and solutions. Special attention has been paid to allow participants to practically experience the various steps, simulating a complete evaluation procedure, starting from the analysis of documents, through the direct interaction with the SP, choosing the professional feedback for improvement, and proposing a simulated final judgment to the Regional Director.

All the procedure has been accompanied by a pre- and post-seminar questionnaire, aimed to collect data about: i) satisfaction about the course; ii) motivation to apply learned skills; iii) expectations about one's own performance as assessor; iv) expectations about the quality of the whole assessment system that will take place in the near future.

4. THE PERFORMANCE EVALUATION OF SPS: THE IMPLEMENTATION PHASE

After the end of the training programme, regional Directors started the evaluation procedure in their region. Using a dedicated digital platform, every SP was requested to upload his/her own portfolio, documenting activities performed and results obtained, together with a self-evaluation against the same criteria used by the assessors.

During approximately three months, all the evaluation teams performed the assessment sessions required, taking into account official documents of the school, SP's personal portfolio, direct interaction with the SP, in every region. Then, they delivered the proposed assessment to the regional Director, including for every SP the personalized "professional feedback for improvement".

Every SP finally received the final assessment and the feedback, directly from his/her regional Director.

During this period, factual data (e.g. the number and timing of the performed assessments) were collected. At the same time, INVALSI conducted a broad research activity, aimed to monitor the whole process, through the point of view of both assessors and SPs assessed.

5. THE PERFORMANCE EVALUATION OF SPS: THE RESEARCH AND MONITORING PROGRAMME

In order to monitor the implementation of the performance appraisal system, we investigated two areas.

The first area is a subsidiary one and concerns the efficacy of the training delivered to assessors, in order to get an estimate of the transfer of training obtained. As noted below, this was the widest training programme in this domain in the Italian school: getting data (although based only on self-rating by assessors) was crucial.

The second area is the central one and concerns the way the SPs perceived the assessment procedure. With this scope, we adopted the Appraisal Effectiveness model firstly developed by Cardy and Dobbins (1994) and subsequently integrated by Levy and Williams (2004). This model postulates that appraisal effectiveness consists of three main components: a) rater errors and biases; b) rating accuracy; c) appraisal reactions. The research and monitoring programme specifically focused on appraisal reactions held by SPs, which included dimensions such as: satisfaction

for the whole assessment system; satisfaction for the assessment session; perceived utility; perceived accuracy; procedural justice; distributive justice; interactional justice; motivation to use feedback; fairness and competence exhibited by the assessors.

Besides the scope of reliable information about the perception of the procedure and its outcomes, we were interested in locating SPs' answers into a situated work context. It is well known that the same procedure may be considered more or less positive, useful, and acceptable, also depending on the characteristics of the respondent's work situation. Specifically, with the aim to analyse such a work situation, we adopted the Job Demands-Resources Model (JD-R) (Bakker & Demerouti, 2017). This model assumes that job performance, as well as its personal consequences, is determined by two processes. The first is the *health-impairment* process, which postulates that job demands are able to increase strain (such as exhaustion) in employees and consequently decrease job performance. The second is the *motivational* process, which states that job and personal resources are able to increase motivation (such as work engagement) and consequently increment the levels of job performance.

This model appeared to be useful for two aims. First, in order to locate perceptions about the assessment procedure within a specific work context (as perceived by respondents). Secondly, this model helped to analyse the relationship between lower and higher job performances (as rated by assessors), and the working conditions (as perceived by respondents). This second part of the questionnaire proposed to the SPs included scales of effort/reward imbalance, professional exhaustion, workload, received support, adequacy of skills, skill discretion, decision latitude, and organizational identification.

6. POPULATION AND PROCEDURE OF THE RESEARCH AND MONITORING PROGRAMME

Two populations have been studied.

The first one was composed by the 800 assessors, who received three questionnaires: i) Q1, at the beginning of the 2-day training seminar; ii) Q2, at the end of the same seminar; iii) Q3, after the end of the evaluation work. Q1 and Q2 investigated through a pre-post design the perceived efficacy of the training session, while Q3 (administered via a dedicated digital platform) collected assessors' perceptions about the task performed, checking also changes in confidence previously expressed by Q2, at the end of the training session.

The second population was formed by all the 6500 SPs. They too have been called to participate in this research through the same dedicated digital platform.

They answered two questionnaires: i) Q1, before the beginning of the assessment procedure; ii) Q2, after they received the final formalized assessment by their Regional Director. Referring to the structure described in paragraph 5, Q1 comprised the scales exploring the expectations about the assessment procedure, while Q2 proposed the same scales after receiving the final assessment, plus the scales describing SP's work context (following his/her perception).

It is worth of note that during the first implementation of this new performance appraisal system, a rude confrontation between government and SPs occurred. The unions asked for a greater recognition (both in terms of salary and of professional autonomy at school), and the government decided to resist. Within this framework, SPs decided to use their participation in the assessment procedure as a tool of political pressure. As an effect of such confrontation, SPs' major unions invited to boycott the assessment itself and, of course, the related research questionnaires. This hot political climate-induced the government to declare that this first campaign of performance assessment was an experimental one, also "freezing" any impact (monetary and other) of the evaluation performed.

The impact of such a political climate was relevant, both for assessors and for SPs. The whole group of assessors answered the questionnaires Q1 and Q2 (pre- and post- training session), administered in presence at the very beginning and at the end of the training sessions. Q3 questionnaire (after the end of the assessment procedure) was answered by 74% of the assessors. As for the response rate of SPs, despite this very unfavourable climate, 60% of the SPs who decided to give their contribution to the assessment process (completing their portfolio, and accepting the interaction with the assessment team) collaborated also by answering the questionnaires we proposed. In total, and summing up Q1 and Q2, more than 5500 questionnaires have been collected.

7. FINAL REMARKS

The main scope of this paper was to present briefly the interplay between design, implementation and research monitoring of a wide and innovative organizational intervention. However, although this paper does not want purposely to be a classic research report, centred on hypothesis testing and data analysis, some preliminary results that emerged from questionnaires may add some interesting information.

Concerning the population of assessors, data from Q1 and Q2 (pre- and post-training session) showed in general high satisfaction for the training itself, and good levels of confidence about future performances, both for the respondents themselves

as assessors, and for the new appraisal system. It suggests that participants appreciated the efficacy and the quality of the massive training intervention performed. Data from Q3 (answers given by assessors after the assessment procedure closed) are currently under examination, but a first analysis suggested that, in average, the real assessment has been harder than foreseen (compared with pre-assessment expectations at Q2). However, assessors in average self-rated their team performance as satisfying, able to ensure organizational equity, and to capture the core quality of SPs' performance they had to assess.

Concerning the SPs respondents, preliminary data (still under examination) are suggesting that satisfaction for the final evaluation received is mainly a function of fairness, competence, and genuine respect shown by assessors, as predicted by the model. Of course, more analyses are needed to explore this large database, linking SPs' perceptions also to personal and situational characteristics, following the frame of the Job Demand-Resources model adopted.

The performance assessment of the 6500 Italian SPs has shown the complexity of this task and the challenge set by its first implementation. The research activity conducted by INVALSI is showing that the indicators of the quality of the process proposed by the literature, both scientific and professional, have to be considered a necessary requirement of the system.

In particular, the quality of the whole assessment process, and of its outcomes, seems to rely on two main factors. The first one is the degree of participation allowed for the SPs assessed. The second one is the high level of commitment and performance that the process requires the teams of assessors to show. The main challenge for them seems to switch from an old-fashioned concept of "assessment as rating", towards a more flexible and modern one, considering "assessment as help for improvement".

It is worthy of note that this is a very crucial point in the international debate too, as shown quite recently by a brilliant paper (Adler et al., 2016) under the title: "Getting Rid of Performance Ratings: Genius or Folly? A Debate". Following this debate, it appears that performance appraisal systems would be no longer beneficial (not even in financial terms) for organizations, if confined to a simple rating classifying "good" and "bad" workers. On the contrary, the right question would be if the performance assessment in use is really able to give a concrete contribution to organizational performance. In other words, performance assessment adds value when it is able to suggest what kind of knowledge, skills, and behaviours should be adopted by organizational members as targets of professional improvement.

We hope that researchers may be able to accompany more and more the implementation of new performance assessment programmes, mainly in relatively new domains, like the school. Many elements suggest that such research would show

the additional contribution that organizational sciences might offer for a better understanding of the factors influencing the quality of processes and outcomes, even in so “special” organizations like schools.

REFERENCES

- Adler, S., Campion, M., Colquitt, A., Grubb, A., Murphy, K., Ollander-Krane, R., & Pulakos, E. (2016). Getting Rid of Performance Ratings: Genius or Folly? A Debate. *Industrial and Organizational Psychology, 9*(2), 219-252. doi: 10.1017/iop.2015.106
- Aguinis, H. (2009). *Performance management* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Aguinis, H., Joo, H., & Gottfredson, R. K. (2011). Why we hate performance management – and why we should love it. *Business Horizons, 54*(6), 503-507. doi: 10.1016/j.bushor.2011.06.001
- Bakker, A., & Demerouti, E. (2017). Job Demands-Resources Theory: Taking Stock and Looking Forward, *Journal of Occupational Health Psychology, 22*(3), 273-285. doi: 10.1037/ocp0000056
- Cardy, R. L., & Dobbins, G. H. (1994). *Performance appraisal: Alternative perspectives*. Cincinnati, OH: South-Western Publishing.
- Davis, S., Darling-Hammond, L., LaPointe, M., & Meyerson, D. (2005). *School leadership study: Developing successful principals (Review of Research)*. Stanford, CA: Stanford University, Stanford Educational Leadership Institute.
- Hallinger, P., Adams, D., Harris, A., & Suzette Jones, M. (2018). Review of conceptual models and methodologies in research on principal instructional leadership in Malaysia: A case of knowledge construction in a developing society. *Journal of Educational Administration, 56*(1), 104-126. doi: 10.1108/JEA-03-2017-0025
- Levy, P. E., & Williams, J. R. (2004). The Social Context of Performance Appraisal: A Review and Framework for the Future. *Journal of Management, 30*(6), 881-905. doi: 10.1016/j.jm.2004.06.005
- Paletta, A., Alivernini, F., & Manganelli, S. (2017). Leadership for learning: The relationships between school context, principal leadership and mediating variables. *International Journal of Educational Management, 31*(2), 98-117. doi: 10.1108/IJEM-11-2015-0152
- Purinton, T. (2013). Is instructional leadership possible? What leadership in other knowledge professions tells us about contemporary constructs of school leadership. *International Journal of Leadership in Education, 16*(3), 279-300. doi: 10.1080/13603124.2012.732244
- Weick, K. E. (1976). Educational Organizations as Loosely Coupled Systems. *Administrative Science Quarterly, 21*, 1-19. doi: 10.2307/2391875