

Public Performance Evaluation: from the International Debate to the Italian Legal Context

Irina Dokalskaya*

Abstract In recent years, public administration has experienced a period of strong pressure for change, innovation and modernization that has led in an increasing emphasis on measuring and reporting of results. Still public sector reforms continue to preoccupy governments all over the world since the controversies remain a problem with the integration of decisions concerning the use of the budget means allocated to the results obtained. Because of modernization reforms, the structural foundation of public organizations has changed in many countries. As part of NPM reforms, public organizations have adopted (or trying to) private sector-inspired management techniques, with the declared intention of improving both the quality and productivity of services. The primary cause of this situation has been the growing needs of citizens of services and policies in terms of quality, that public authorities have had and will have to satisfy with increasingly scarce resources. Considering the necessity of greater understanding of the economic and financial performance, this paper discusses the possible modalities, the experiences and the most appropriate processes needed to achieve the strategic and policy objectives for maintaining and improving the quality of services, ensuring the effectiveness of administrative function.

Keywords Public administration; Performance evaluation; Decision-making policies; Management of expenditures; Legal procedure

JEL Classification G28; H00; H11; H83; K4

1. Introduction

In order to improve the effectiveness and the efficiency of the use of budget funds, there is a shift towards the planning and execution of budgetary expenditure for programs and missions, which is the budgeting process based on the principle of public policies financing (results-oriented type). In this type of a financing system,

Irina Dokalskaya* (✉)

* Ph.D., Project manager at Leonid Srl, Innovative start up - research company for the innovations and improvements processes

e-mail: dokalska@yandex.ru

the decision-making core is represented by the analysis of a “cost-benefit” parameter, which assesses the conformity of the results achieved with respect to the initial objectives. So far, the integration aspect of decisions on the use of budget means allocated to the results achieved remains controversial (Andrews and Hill, 2003, Gilmour and Lewis, 2006). Therefore, public funding should be linked to the transition to the performance management system, based on the financing of specific programs aimed at the achievement of concrete results. This trend, which reflects the rules established by international practice, is determined by the reduction in public spending based on the estimate of the balance sheet items. Funding based on the estimation it is often associated with inflated figures justifying the amount of the costs for the execution of certain functions of the state. In addition, there is the need for additional costs to control the use of budgetary funds.

This makes the system of funding based on the estimated ineffective. The budget for the programs, on the other hand, aims to solve the problem of growth in public spending and to evaluate the effectiveness of these programs, in order to combine the numerous program budgetary tools used in the budget planning process. The development of this type of budget is an attempt to combine all the means to achieve public objectives and establish the budget in the context of the program, which inevitably leads to a rethinking of the functions and duties of the executive authorities. In this regard, one of the most important steps in budget reform becomes a clear consolidation of authority and responsibility among the public policy executors and co-executors in the process of policy planning and implementation. This paper is organized as follows: in the next section (section 2), the concept of performance and the need to measure it, as well as the major research directions that have been taken on the subject will be discussed in order to introduce the components of the main approaches to the performance evaluation, and what justifies in this case the choice of the valuation method (section 3). The final passage will be to go transversally through principal steps and efforts that have been made at the national legislative level to extend the debate on the subject for the management of the public performance evaluation process (section 4). Conclusion concludes.

2. Theoretical debate and perspectives

Evaluation in the public sphere has a precise meaning, which is primarily linked to the analysis of the effectiveness of policies or programs (Vecchi, 2007). Evaluation refers to the formulation of judgments based on the causal relationships that exist between an action (or a set of actions) and the relevant results through observations, measured with quantitative and qualitative techniques. The results, in this case, are not so much about the production of products, but the ability to change the situation of subjects who are in a condition considered unsatisfactory. Therefore, the basic evidence is a demonstration of the causal link between an intervention and the relevant changes, without which the results would not have been obtained.

The public system performance represents a theoretical construct design that cannot be measured directly, as the aims and objectives of public enterprise (services) are not

expressed exclusively in economic and/or financial terms, but are one-dimensional. Every attempt to measure it involves wide margins of arbitrariness in the choice of the dimensions to be evaluated and of the proxies to be used for the measurement, and must serve considerable approximations (Solipaca, De Belvis, Annexes, 2013). Performance measures, however, have traditionally been one-dimensional in public administrations, concentrated on the level of spending on the purchase of the factors or, at most, on the monetary value of the inputs consumed. This is due to the difficulty of developing a clear understanding of the quantitative and qualitative indicators of the effectiveness of public expenditure. Performance analysis cannot be limited to the quantitative evaluation of the results achieved, but must also focus on their quality and determinants, in a strategic approach (Kaplan, Norton, 1992; Lee, 1997; Cornett, 1998; Melkers, Willoughby, 1998; Kloot, Martin, 2000). Efficiency as an end in itself, meant as a reduction in resources, does not make sense in public services; it is, in fact, a “relative” concept: it measures the resources used with respect to the results achieved.

The concept of the performance and the need to measure it were already present in Taylor’s work (1911) and their usefulness for the government of local administrations were identified by Ridley and Simon since 1943 (Poister, Streib, 1999). However, only with the managerial approach to public administrations, New Public Management (Osborne, Gaebler, 1992; Snell, Hayes, 1993; Hood, 1995; Garsombke, Schrad, 1999; Lynn, 2006) underlined the necessity of pursuing performance results and measurement, through the introduction of Performance Management Systems (Pollitt, Bouckaert, 2000; Hood, 2005; Van Dooren, Bouckaert, Halligan, 2010; Mettler, Rohner, 2009; Nielsen, 2014; Gerrish, 2015), which became the nucleus of intellectual thought in the field of public finance. According to the arguments of NPM supporters, the major results-based assumptions are able to improve the efficiency of public services and effectiveness, as they solicit managers to clarify strategic objectives and to develop results-oriented measures to achieve their goals.

Stimulated by this debate, the performance, since the 2000s, has become one of the most important issues at the international level. One of the research directions was represented by the studies aimed to determine whether performance measurement systems lead to better outcomes in the public context, with varied conclusions. In some cases, the factors contributing to the use of performance information were analyzed (e.g. Ammons, Rivenbark, 2008; Kroll, 2015); in the others, the main research focus has become the way performance measurement systems should be designed in order to achieve the intended results (e.g. Padovi et al., 2010). From an analysis based on the theoretical contributions made at the international level, the research themes were mainly carried out in three directions (Galli et al., 2010). *A first direction* is turned to the past and the compression of the characteristics that have distinguished the reform processes that over time have tried to implement systems for measuring and evaluating performance in different countries. *A second direction* is aimed at the analysis of the impacts produced by the introduction of performance measurement and assessment systems within public administrations. Finally, *a third direction* is addressed to conceptualizing the relationship between measurement and performance evaluation

systems and external subjects, all with a broad reference to network theories, the environment within which performance assumes a new connotation. The evaluation of performance in the public administration has developed and refined in the last decades due to strong requests from institutional subjects both internal and external to the public administration to increase productivity, that is to offer the same services by reducing and stopping the increase in public spending (Ruffini, 2010). Taking into account the peculiarity of public services and consequently the methods and difficulties encountered to measure the results, the identification of the links between the resources used and the expected results has become a priority for public administrations, especially in the context of limited financial resources available. Wherein, according to some authors, the conviction that the level at which performance measurement and evaluation systems need to be analyzed is not the micro one (that is, related to single activities or processes) has increasingly matured, nor the *meso* one, analysis of systems of specific measurement of some public administrations. ***At the center of attention, therefore,*** are the systems of public administrations and the networks in which these systems are inserted in order to define, implement and evaluate public policies, activities, and services rendered to users (Valotti et al., 2010).

The increasing number of empirical contributions which has been undertaken ultimately, concerns, in the first place, the attempts to better understanding the benefits, impacts, and challenges of the conceptual and value interpretations of performance management and measurement in the public sector. The undertaken works have employed different methodologies, like case studies examples, experience-based observations, and surveys. However, the analysis of the various international experiences does not seem to have produced conclusive evidence on the achievement of the assigned objectives.

As it emerges from the contrasting elements that arise from the research carried out in the area of performance evaluation in public administrations, ***the results for the most part would indicate the lack*** of a substantive testimony of the performance management benefits applied at the public sector. Even if in some cases, the focus on performance management seems to be shifting from simple collecting and performing data to actually using them for decision-making process (C. Bianchi, W. Rivenbark, 2013), still there is no clear evidence that performance information has been very successful in meeting the goal of improving performance or that it is being used extensively. In many cases, due to a fundamental misunderstanding of such issues as performance indicators, these last made little or no difference to performance (Parmenter, 2015).

Therefore, just as an important aspect becomes a lack of qualified personnel and/or information systems that are poorly standardized and integrated and considered unreliable by the users themselves (Vasilescu, 2011). Some of the recently conducted researches claim that public sector organizations have a negative image on productivity-related aspects of performance (e.g. Hvidman, 2015), although ***empirical evidence on public and private performance*** is yet weak and ***does not allow any firm conclusion*** that public organizations should perform worse (or better) than private business. The

question is whether negative stereotypes of the public sector nevertheless are reflected in how citizens perceive the performance of public sector provision (Andrews, Boyne, Walker, 2011).

The Consciousness of the *empirically-based type on the effects* that causes information about performance on policy makers is still limited. One of these researches, however, has developed some hypotheses concerning the relationship between performance information and attitudes. The results of an experiment that included the survey among Danish municipal councilors, demonstrates that information on performance matters, in the sense that it affects the attitude of central decision-makers (Nielsen, Baekgaard, 2013). The results also indicate that the guilt-avoidance perspective applied by the authors to hypothesis development could be the guide of the politicians' responsibilities to performance information as opposed to performance-based fiscal recommendations, according to which high performance should be rewarded and low performance should be punished. In particular, it has been found that the experimental treatment of information showing high and low performance had a positive impact on the attitude towards spending, while information on the average performance negatively affected the same attitudes. Furthermore, information showing high performance has made politicians less favorable to the pursuit of organizational reform. An important conclusion of the survey is that a greater appreciation of high performance among policy makers could alter the incentives for public managers to focus not only on avoiding blame but also on the pursuit of success.

3. Methodological approaches to the performance evaluation

Choice of the evaluation methods

The whole variety of evaluation methods found in the literature can essentially be traced back to two main types: evaluation by factors (or fixed parameters) and evaluation by objectives. In the first case, those who evaluate performance do so use a series of parameters, which should identify those components of performance (proficiency, knowledge, ability, organisational behaviour, skills) considered critical for work success. The judgement on each of these components then flows into a synthetic judgement on the whole service provided by the collaborator (Borgonovi, 1998). In the second case, on the other hand, the evaluation is based on the results actually achieved in relation to the objectives initially set; this creates a close integration between the personnel evaluation system and the management planning and control system, since it is the latter that provides the data relating to the results themselves and the planned objectives. This is the method used in Management By Objectives (MBO), which attributes economic recognition to the achievement of objectives. If the objective and factor evaluation methods are used together, then the term "mixed evaluation" is used.

When choosing which evaluation method to adopt, the relative pros and cons should be taken into consideration. With regard to fixed parameters, their definition is characterized by high uncertainty: first of all, there is the problem of identifying "factors that are really significant and observable, so as to allow a thorough

examination of performance”; secondly, it is necessary to clearly define the content, after which the various evaluators can give a homogeneous interpretation (Borgonovi, 1998). On the other hand, however, performance evaluation based on results obtained has the advantage of being more objective and concrete, but there can also be different complexities and problems in defining objectives: “What is the appropriate number of objectives?”, “Are the objectives to be expressed only in numerical or also descriptive terms?”, “How do we weight the different objectives?” ...Moreover, the objectives fail to cover all areas of responsibility which a position includes, so the problem remains to find a way to ensure a comprehensive performance evaluation. This latter problem can be solved by resorting to a mixed system, which examines performance from the dual perspective of the results and behaviours shown (Borgonovi, 1998).

The evaluation, regardless of whether it is related to public policies or not, is a process that is divided into six main stages (Hatry 1999):

- Development of a conceptual model related to the programme subject to evaluation and identification of the key points of the evaluation
- Identification of the evaluation questions and definition of the measurable outcome variables
- Definition of the evaluation design
- Data collection
- Data processing
- Dissemination of the resulting evidence

Identifying the expected results of an evaluation process means defining which information needs the evaluation must respond to. Considering cognitive needs, the evaluation can take five different forms (Martini, Cais, 2000), which represent the classification widespread in the literature and which seem to also be accepted by the European Union¹:

- 1) *Policy and programme design*, i.e. evaluation as a tool to allocate resources to the most deserving uses
- 2) *Management control*, i.e. evaluation as a tool for monitoring performance within organisations
- 3) *Accountability*, i.e. evaluation as a vehicle to account for achievements in a certain field of public action
- 4) *Implementation*, i.e. evaluation as a tool for critical analysis of the implementation processes of a policy
- 5) *Learning*, i.e. evaluation as an estimate of the effects produced by a policy.

Thus, each evaluation must be “tailor-made”, and although the evaluation design can vary in complexity, the evaluation work is always influenced by (Rossi, Lipsey, Freeman, 2004):

6) The purposes of the evaluation.

¹ All public policies financed by the Structural Funds are subject to an evaluation procedure, whose need is reiterated in the Community Regulations, both in Regulation 1260/99, relating to the 2000-2006 programming, and in European Council Regulation 1083/2006, as it concerns the 2007-2013 programming, laying down general provisions on the European Regional Development Fund, the European Social Development Fund and the Cohesion Fund.

- 7) The conceptual and organizational structure of the programme/project/service.
- 8) The resources available for the evaluation itself.

There are therefore no rigid guidelines to guide the design of an evaluation plan, which is always a creative and collaborative effort, - the methodological approach should first be chosen in line with the objective of the evaluation; each evaluation design, however, is structured around the following three points:

- the nature of the relationship between the evaluator and the stakeholders
- the evaluation questions
- the methods of evaluation

Rossi, Freeman and Lipsey propose a systematic approach to the evaluation of public policies that is essentially based on some questions that the evaluator cannot avoid:

- What are the purposes of the evaluation in question?
- What is the structure and context in which the analysed programme is found?
- How should the relationship with the stakeholders be set up?
- What are the evaluation questions that need to be answered?
- What are the applicable methods to give solid answers to the questions asked?

These five evaluation questions constitute the same stages of the process of evaluating a public policy, which are in a hierarchical relationship, constituted by the evaluation of the social problem to which the programme is addressed:

- Evaluation of the social needs to which the programme intends to respond
- Evaluation of the theory underlying the programme
- Evaluation of the process/implementation
- Evaluation of the impact/effects
- Evaluation of efficiency

With *the evaluation study of social needs*, we try to answer the question “What is the problem?”, around which arise a series of questions and epistemological issues with the aim of defining the problem in its dimensions, trends and particular characteristics, making use of both quantitative techniques and data and also *qualitative* methods of investigation that allow the problem to be described, and take into account the different perceptions of it by the different stakeholders.

The theory of the programme is broken down by the authors into two main components:

- impact theory, which refers to the change expected after the programme according to a chain of causal events and links;
- process theory, which describes the steps and activities to be performed in the organisation and implementation of the events.

The theory of a policy can be both “explicit” and “implicit”, and it is up to the evaluator to define the average through different information sources, which make it possible to gather a policy’s objectives, its functions, components and activities of the programme, as well as the sequential logic that binds functions, activities and products. If the theory underlying a policy is considered well defined and justified, then it is possible to continue with the evaluation. Otherwise, it would not make sense to continue, because every result obtained would be impossible to interpret in relation to the policy.

The purpose of *the evaluation of the implementation process* is to understand “what the

programme really is”, and whether the services/interventions actually reach the target population or not. It can be carried out in retrospect (programme process evaluation), necessary for impact evaluation, or in an ongoing manner (continuous programme monitoring), which is necessary mainly for programme managers, as it provides regular feedback on the implementation of interventions. Evaluation of a policy’s implementation process also takes on a certain importance in terms of *accountability*, since it helps all possible stakeholders realise “what the organisation is doing”. Some aspects of the analysis are the level of participation/involvement of the *target* population, the level of “*bias*” in the provision of a service, and the level of dropouts. By comparing the results of this analysis with the theory of change underlying the policy, it is possible to know whether the policy is implemented as planned or not (*implementation failure*).

The identification and measurement of the causal link between the process variable (the product or output) and this dependent variable (effect or outcome variable) constitute the challenge of *evaluating the effects of the policy*. The outcome variable can be considered from different points of view:

- outcome level
- outcome change
- programme effect (impact)

The authors stress the importance of the distinction between monitoring and measuring outcomes and impact evaluation. Monitoring (*programme monitoring*) can be defined as the systematic collection of aspects of a programme’s performance that indicate whether it is functioning as originally intended or in accordance with certain previously established *standards*. The subject of the monitoring is the programme’s *performance*, which can concern the products, *the outcomes*, and the implementation process. Monitoring, even if based on observation of indicators related to outcomes, cannot be defined as “evaluation of the effects”. Impact evaluation can only be carried out after making the evaluations placed at the lower stages of the “hierarchy”. We talk of a “*black box*” evaluation when the evaluator conducts an impact analysis without knowing anything about the processes that generated these effects. This is basically a comparative analysis: the group addressed compared to the control group; or factual counterfactual value, that is, the estimate of the value that the variable result would have had if the group addressed had not been subjected to the process.

Efficiency evaluation, as well as impact evaluation, is more appropriate for mature and stable programmes. It can be carried out *before the event* (for allocative choices in the presence of scarce resources - allocative efficiency) or *after the fact* (as an extension of the impact evaluation). The authors identify two sub-categories:

- cost-benefit analysis (the effects of the programme are expressed in monetary terms)
- cost-effectiveness analysis (the effects of the programme are not transformed into monetary terms).

Measurement of the performance: the indicators' problem

The multidimensionality of the performance, as it was noted before, requires a measurement and evaluation process, powered by qualitative and quantitative information (Bouckaert, 1995; Hatry, 1999; Jackson, Palmer, 1992). It requires the development of a system of indicators, as well as a process of monitoring, reporting and use of functional information in decision-making. The link between performance management and performance measurement is fundamental here, as performance management incorporates and uses in the decision-making process the information produced by the performance measurement (Van Dooren, Bouckaert, Halligan, 2010). The existence of a performance measurement system is a necessary condition for directing and focusing the attention of the decision makers on the results and their determinants, in order to make employees aware of what is expected of them and to raise awareness of the stakeholders for the purposes, the objectives, and the results of the organization and the manner in which they were achieved (Nisio, 2012).

On the subject of indicators, as Palumbo points out, two different conceptions face each other (2001). According to the first concept, which is closer to the lexicon of social research, the indicators are specifications and articulations of a given concept in measurable or otherwise detectable dimensions. As a rule, one must observe something that is possible to know (and perhaps measure) in place of something else that cannot be accessed directly. So, in many concrete situations, variables of the most interest are almost never immediately translated into direct data collection operations. The degree of functioning of a program, the level of achievement of its objectives, can be known through indicators, which are the result of the decomposition of a more general concept into one (usually more than one) more precise, susceptible to lead to detection and measurement procedures. The indicator is, therefore “a property that the researcher conceptualizes both as such, and as a surrogate of another that interests him more”, where “the necessary and sufficient condition for the use of indicators is, therefore, the impossibility (or the serious difficulty) to give a satisfactory operational definition of a property that affects, regardless of its level of generality, theoretical relevance, or other quality” (Palumbo, 2001 see Marradi, 1994).

The second concept, which is closer to the statistical exception, as elementary processing of basic data, does not start from the concept to then define the indicators, but from the existing data to ask how these can be elaborated, in a more or less complex way, in order to provide a known contribution to the problems. The two meanings actually correspond to two different strategies for the construction of indicators: the first one favors the process of conceptual clarification and the subsequent operations of identifying indicators that are, firstly, valid and only secondly readily detectable or, at the limit, obtainable from data already available; the second strategy starts from the available data (or easily detectable) to define only in the second measure which their elaborations can be considered valid indicators of the concept to be measured. The two strategies, according to the same author, can also be traced to a deductive and an inductive approach (see Palumbo, 1995, 2001), where, in both cases, different paths

can be defined between the following steps:

- data – available or detectable
- information – data structured according to given criteria
- indicators – structured information and possibly elaborated according to given criteria
- concepts – to which the indicators are referred. (Palumbo, 2001)

In fact, an intermediate way is often followed in the evaluation process, inasmuch the specification of the concepts develops in tandem, normally with the analysis of the available information and the indicators help to specify the concept itself more than it determine the choice. Both in the case of the defined “sociological” and “statistical”, to be adequately used in the evaluation process, the indicators must possess the following characteristics (Palumbo, 2001 see Zajczyk, 1997):

- *validity*, understood as the ability to effectively represent the concept indicated in an acceptable way
- *reliability*, when the outcome of the operations set up during the operation is positive, i.e. it effectively captures the status of the properties that the researcher had set out to observe
- *sensitivity*, or power of resolution, understood as the ability to discriminate between the different forms that a phenomenon can take
- *adequacy*, or degree in which the indicator meets the evaluator’s cognitive needs
- *comparability*, which may concern both the spatial and temporal dimensions, i.e. the possibility of comparing the same phenomenon in different territorial areas, the same phenomenon in the same area but at different times
- *timeliness*, or the inclusion of the indicator within a decision-making or evaluative process of support that makes it necessary to have the indicator itself in adequate time to modify a program or an intervention thanks to the opinion formulated on the basis of the indicator.

The use of indicators, however, is not very productive if it is not linked to the theory of the program (as mentioned in the previous pages). The problem arises, in particular, when using indicators in a statistical sense, that is when one appeals to elementary processing of available data, already recorded for other purposes, intended to bend to the assessor’s cognitive needs. On the one hand, the specific context in which the survey will take place must be entered (Palumbo, see Cannavò, 1999) and, on the other, the theoretical frame of reference that provides sense and meaning to the concept. Therefore, the indicators are constructed, and in the case of the indicators used in the evaluation, the theoretical references are inextricably linked to those of a practical nature: an intervention is carried out because the production of certain consequences is expected and, on the other hand, on the basis of the consequences will be judged not only the quality/value of the intervention achieved but also the inspiring theory.

4. Legal background

The need to improve performance is not exclusive to the private sector but also applies to public administrations. Unlike the private sector, where efficiency is due to the need

to obtain real profits compared to operating costs and where the market bears witness, there is no such evidence in the public administration, and the NPM reforms represent an attempt to establish the public sector as an equivalent to market discipline.

In Italy, the growing interest in evaluating performance is not only due to the problem of the system's economic sustainability but also to the effects of the profound political and organizational changes observed over the last 20 years (Solipaca et al, 2013). Already in the 1980s, when people began paying attention to reducing spending and improving the competitiveness of the public sector, some awarding institutions were introduced for the first time in public employment contracts, but the recognition of merit remained more than anything else a formal declaration, valid "only on paper". In fact, the interventions were calculated on the basis of the contractual level of classification and the rate of presence, without any attention to commitment, results, skills, etc. (Mattalucci, 2011).

In the 1990s, through some legislative measures, the theories of modernization of the Public Administration were tested, defined by some authors as a modernization process managed by law (Rossi, 2013). The key concept promoted by many reforms of the Italian public administration at the time concerned the full use of systems for evaluating staff, and in particular management, the establishment of bodies responsible for measuring and evaluating effectiveness and efficiency of the public administrations' activity, the refinement of punishment/reward systems, as well as the efficiency of public spending and better organization of public administrations. In the literature there is a distinction of several reform cycles in terms of public management (e.g. Ongaro, 2009), which summarize the administrative activities and legislative procedures performed in the 1990s with the aim of improving the quality and efficiency of programming public policies, up to the need to reorganize and strengthen the system with the aim of "ensuring high quality and economic standards of the service through the valorization of results and organizational and individual performance" (Legislative Decree No. 150/2009 Art. 2) , and therefore *putting the concept of performance at the center*.

The new reform (so-called Brunetta reform) is inserted, as noted by some authors (e.g. Vasilescu, 2011), in a context of growing distrust in the ability of public administrations to provide adequate responses to the social and economic problems of the country, explicitly acknowledging that the pre-existing systems had not achieved the expected results. An official document notes that the verification of objectives is currently limited in the majority of cases to the sole implementation of spending programs and performance reports do not allow an assessment of the degree of achievement of the programmatic objectives, but are limited to a mere description of the administrative activity for the expenditure of allocated funds (Council Presidency, 2009). Although inspired by previous attempts to reform the public administration, the document develops classical and central themes introducing novelty and some radical turns. In this context, performance is intended as the contribution (result and modality of achievement of the result) that a subject (system, organization, organizational unit, team, single) contributes through its action to the achievement of goals and objectives

and, in ultimately, to the satisfaction of needs for which the organization was formed. Moreover, the legislative decree 150/2009 emphasizes that the administration's performance has to be intended as the performance of the organization based on the performance of its staff.

The "Sistema della Performance" introduced by Legislative Decree 150/2009, can, therefore, be basically conceived as the set of tools through which the Public Administrations and the Territorial Bodies, plan, manage, measure, control, evaluate and communicate the own performance, from an organizational and individual perspective (Corteselli, 2016).

In spite of the fact that it was a strong application, supported by specific legislation, the implemented performance management systems had significant limitations and encountered slowdowns. The rapid start-up of the new systems did not respond effectively to expectations, as they were connected to the period (2009-2014) of the financial and economic crisis and frequent changes in government (Rebora, Ruffini, Turri, 2015). As noted by the authors themselves, in practice the developments of the 2009 reform experienced the same difficulties that had hindered previous attempts, imposing very heavy management obligations on the administrations without producing significant benefits (Rebora et al, 2015).

Among the most important topics of the recent turn, that the reform of public administration has taken² was that of open data and transparency, as tools to make what the public administration does and what results it produces more understandable. Unlike international experiences, however, in Italy there seems to prevail, as some scholars have pointed out (Corteselli, 2016), a "particular type of transparency focused on the right-duty to information, but above all on the objective of increasing widespread control of public administration".

Transparency often intended mainly as a communication of things that "should" means more than an administrative action value. In turn, the suppression and manipulation of negative data leads to the perception that there is a lack of transparency in public performance reporting, leaving the obstacle for the effective implementation of performance measurement. In the Italian public administration, a legal and bureaucratic culture evidently prevails (Borgonovi, 2005), which has traditionally left out analysis of the managerial component and in particular of performance; so transparency means primarily widespread controls against illegality and then as a stimulus to improve the results of public action.

Other aspects of the reform evoke the need to provide for the digitization of the performance measurement and evaluation process to allow coordination at national level³; they expect the processes of independent evaluation of the level of efficiency and quality of the services and activities of the public administrations and of the impacts produced, also through the use of standards of reference and comparison.

Conclusion

² Entry into force on 2 July 2016 of Presidential Decree No. 105 of 9 May 2016.

³ Law 124 of 7 August 2015.

The need to improve performance is not exclusive to the private sector but also applies to public administrations. Unlike the private sector, where efficiency is due to the need to obtain real profits compared to operating costs and where the market bears witness, there is no such evidence in the public administration, and the NPM reforms represent an attempt to establish the public sector as an equivalent to market discipline. An important feature in the evaluation is that it can be considered as a comparative activity, based on the methodology of research in the human sciences, which in the case of the public administration aims to establish a comparison between institutions, processes, and behaviors with a reference to different contexts, and implies satisfaction of two conditions: the availability of at least two concrete cases; and the adoption of a descriptive approach aimed at illustrating the differences between different contexts and providing interpretative keys (Pickvance, 2001). However, scientific research on the benchmarking of evaluation systems has been more focused on a national rather than a regional level. The numerous analyzes conducted at the regional level have been mainly oriented towards establishing comparisons between regions within the same country, rather than between regions of different countries. As an interpretative key to this phenomenon, it is identified by scholars in the *difficulties connected to the search for homogeneous conditions* in the comparison between two different institutional contexts (e.g. Bianchi, Rivenbark, 2013).

As a key aspect of evaluation analysis of the public policies effects is considered a possibility to make the specificities and institutional characteristics that subtend its implementation more transparent (Potì, Cerulli, 2010). The concept of transparency, which is the basis (after the measurement and assessment of organizational and individual performance) of public administration reform, is often cited as an indispensable element inasmuch it allows citizens to know, not just the situation of individuals which operate within the framework of every entity of the public administration, but also, in general, the way the administration operates. In practice, however, it happens that only a small percentage of the stakeholders are able to learn the information, which leads to thinking of their involvement in the definition of the priorities of the administration and in the decision-making processes apparently not considered.

It is obvious that the qualification of the activities of the use of budget allocations as “ineffective” requires special tests proves of the existence of a practical alternative to the possibility of activities, which could lead to a better use of the resources. Consequently, the core issue, in this case, becomes the need to set a legal mechanism appropriate for the reception and presentation of such tests and therefore, to assess the effectiveness. It is a fact that reforms, in general, are projected for the long term, when the launch of a standard does not immediately bring benefits, but instead is the starting point of a path that must be followed and monitored in order to understand the gaps and to develop the appropriate supplementary corrective measures (Rossi, 2013).

Building a performance-based administration *does not seem to result from the adoption of horizontal regulations*, but mostly from the triggering of change and learning processes within administrations. The tools to support performance are

characterized by the implementation process and their proper functioning also depends on the actors involved and their interaction (Butera, Dente, 2009, Vecchi, 2010). The orientation of the administrative action towards the introduction of performance systems is *not so much a direct consequence of the laws, but of the action* of development agents and leaders who have a strong interest in the use and diffusion of such systems to support the decision-making processes. It is, therefore, a matter of a lot that is much more than a purely normative process and also of a technical order because it configures a challenge between alternative solutions and administrative models, a competition that invests decision and power centers (Rebora et.al. 2015). This represents a difficulty, not only because it is challenging for all the subjects involved, but because it produces significant consequences through sense-making processes, which may have unexpected and contradictory results compared to the expectations of the promoters of the reforms.

However, based on the experiences of research works carried out on the subject matter taking into consideration practice of other countries, the influence of the regulatory systems is *an important reference element* for the comparative analysis on performance management.

References

- Abdel-Maksoud A., et al. (2015). The Use of Performance Information in Strategic Decision Making in Public Organizations. *International Journal of Public Sector Management*, Vol.28, No.7, pp.528-549.
- Ammons, N., and Rivenbark, W. (2008). Factors Influencing the Use of Performance Data to Improve Municipal Services: Evidence from the North Carolina Benchmarking Project. *Public Administration Review*, Vol.68, No.2, pp.304-318.
- Andrews, M., and Hill, H. (2003). The Impact of Traditional Budgeting Systems on the Effectiveness of Performance-based Budgeting: A Different Viewpoint on Recent Findings. *International Journal of Public Administration*, Vol.26, No.2, pp.135-155.
- Borgonovi, E. (1998). *Formare la Dirigenza per Riformare la Pubblica Amministrazione. LAVORO E SICUREZZA SOCIALE*. Francoangeli.
- Cope, G.H. (1987). Local Government Budgeting and Productivity: Friends or Foes? *Public Productivity Review*, Vol.10 (3), pp.45-57.
- Cornett, C. (1998). California's performance budgeting pilot project: A view from the legislative. *Cal-Tax Digest*, February.
- Garsombke, H.P., and Schrad, J. (1999). Performance measurement systems: Results from a city and state survey. *Government Finance Review*, February.
- Gilmour, J.B., and Lewis, D. (2006). Does Performance Budgeting Work? An Examination of the Office of Management and Budget's PART Scores. *Public Administration Review*, 66(5), pp. 742-752.
- Hood, C. (1995). The "New Public Management" in the 1980s: Variations on a theme. *Accounting Organization and Society*, Vol.20 (2-3), pp.93-109.
- Hvidman, U., and Andersen, S. (2015). Perceptions of Public and Private Performance: Evidence from a Survey Experiment. *Public Administration Review*, Vol.76, No.1, pp.111-

120.

- Kaplan, Robert S.; Norton, David P. (September 1993). "Putting the Balanced Scorecard to Work". *Harvard Business Review*.
- Kroll, A. (2015). Exploring the Link Between Performance Information Use and Organizational Performance: A Contingency Approach. *Public Performance and Management Review*, Vol.39, No.1, pp.7-32.
- Kroll, A., and Moynihan, Donald P. (2015). Does Training Matter? Evidence from Performance Management Reforms. *Public Administration Review*, Vol.75, No.3, pp.411-420.
- Lee, R.D., Jr. (1997). A Quarter Century of State Budgeting Practices. *Public Administration Review*, 57(2), pp. 133-140.
- Melkers, J., and Willoughby, K. (1998). The State of the States: Performance-Based Budgeting Requirements in 47 out of 50. *Public Administration Review*, 58(1), pp. 66-73.
- Moullin, M. (2007) 'Performance measurement definitions. Linking performance measurement and organisational excellence', *International Journal of Health Care Quality Assurance*,20:3,pp. 181-183.
- Ngacho, C., and Das, D. (2015). A Performance Evaluation Framework of Construction Projects: Insights from Literature. *Project Organisation and Management*, Vol.7, No.2, pp.151-173.
- Nuti, S. (2008). La Valutazione per la Governance del Psr. *Salute e Territorio*, Vol.166, pp.53-64.
- Ongaro, E. (2009). Public management reform and modernization. Trajectories of administrative change in Italy, France, Greece, Portugal and Spain. Edward Elgar, 2009, pp.316.
- Osborne, D., Gaebler, T. (1992). *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector*. Plume. New York.
- Poister, T.H., and McGowan, R.P. (1984). The Use of Management Tools in Municipal Government: A National Survey. *Public Administration Review*, Vol.44 (3), pp.215-223.
- Poister, T.H., and Streib, G. (1999). Performance Measurement in Municipal Government: Assessing the State of the Practice. *Public Administration Review*, Vol.59 (4), pp.325-335.
- Poister, T.H., and Streib, G. (2005). Elements of Strategic Planning and Management in Municipal Governments: Status after Two Decades. *Public Administration Review*, Vol.65 (1), pp.45-56.
- Pollitt, C., Bouckaert, G. (2000). *Public Management Reform: a Comparative Perspective*. Notes from supporting the international conference on modernization and state reform.
- Poocharoen, O.-O., and P. Ingraham (2003). "Integration of Management Systems in Government Performance Project," *Paths to Performance in State and Local Government: A Final Assessment from the Maxwell School of Citizenship and Public Affairs* (Syracuse: Maxwell School).
- Poti, B., and Cerulli, G. (2010). La Valutazione Ex-post di Uno Strumento di Politica della Ricerca Industriale: Modello Analitico, Processo di Realizzazione, Eterogeneità degli Effetti. Ceris-CNR, Roma.
- Rutherford, A., and Meier, Kenneth, J. (2014). Managerial Goals in a Performance-Driven System: Theory and Empirical Tests in Higher Education. *Public Administration*, Vol.93, No.1, pp.17-33.

- Snell, R., and Hayes, K. (1993). Performance budgeting and the states. Presented to the Nebraska Legislative Appropriations Committee, London, NB.
- Valotti, G., Galli, D., Orlandi, A., Rota. (2010). La Misurazione e la Valutazione della Performance nella PA Centrale. SDA Bocconi School of Management.
- Valotti, G. et al. (2012). Il Livello di Diffusione del Performance Management tra I Dirigenti Pubblici delle amministrazioni Centrali. SDA Bocconi School of Management.