

FUNDAÇÃO GETÚLIO VARGAS

ESCOLA DE ADMINISTRAÇÃO DE EMPRESAS DE SÃO PAULO

Programa Institucional de Bolsas de Iniciação Científica (PIBIC)

Accountability e mensuração de performance:

a relação entre os municípios e o Índice de Efetividade de Gestão Municipal (ieg-m) através
do estudo de caso dos serviços públicos de Educação

Accountability and performance measurement:

the relationship between municipalities and the Municipal Management Effectiveness Index
(ieg-m) through the case study of public education services

AUTOR: ALEXANDRE HANDFAS

ORIENTADOR: GUSTAVO ANDREY DE ALMEIDA LOPES FERNANDES

São Paulo – SP

2019

Accountability e mensuração de performance:

a relação entre os municípios e o Índice de Efetividade de Gestão Municipal (ieg-m) através do estudo de caso dos serviços públicos de Educação

Accountability and performance measurement:

the relationship between municipalities and the Municipal Management Effectiveness Index (ieg-m) through the case study of public education services

Abstract

We propose that the theoretical discussion of accountability is translated to the daily reality of public management through performance measurement systems. Given the fact that citizens want better public services, governments must be held accountable for the quality of the delivery of these services. Briefly, that is why governmental performance measurement deeply involves an accountability process. The goal of this work is to dive deeper in the assumption just described above.

We realized a structured review of public sector's performance measurement literature and developed a framework of the literature functioning dynamic. During all the review, our objective was to check if this literature discussed the accountability implications of performance measurement in public organizations.

Also, we investigated ieg-m, an index elaborated by São Paulo's Court of Accounts that is annually applied for all state's municipalities. More specifically, we studied i-Educ, the public education indicator from ieg-m. We visited three municipal education departments where we conducted interviews with political agents and bureaucrats. Our main goal was to

understand the relationship between these organs and i-Educ and investigate if, by any chance, the indicator increased their management capacity.

The literature review showed us that there is a gap between the concepts of governmental performance measurement and accountability. The majority of the analysed papers do not deal with any kind of accountability discussion. Regarding the field study, we verified that, despite of the fact it cannot be used as an internal performance measurement tool by municipalities, i-Educ have the potential of increasing management capacity.

Keywords

accountability, public services, performance measurement, local government, democracy

1. Introduction

Since the 18th century, the western world have been undergoing an extraordinary process of complexification. This dynamic of change is particularly visible and important regarding state's duties and functions. If in the centuries prior to the liberal revolutions one of the only functions of governments was to protect its citizens and its borders from external invaders, little by little, the state started to assume a myriad of functions, mainly the delivery of public services, such as health and education.

Furthermore, the western world has also gone through an important process of democratization. People began to elect its representatives, responsible for running government (executive branch) and for making laws and checking the work of government (legislative branch). A responsivity relationship between elected public officials and citizens was built and solidified in a way that politicians began to take responsibility for informing citizens about a diverse set of elements concerning their actions while performing public duty.

The american political scientist Robert Dahl described the convergence between both trends, public services and democratization/government responsiveness. His concept of *Polyarchy* empowered him with the appropriate theoretical elements to explain these dynamic. Merging the jeffersonian tradition of an almost unlimited populist democracy with the madisonian idea of checks and balances, Dahl (1997) defined his theoretical model of *Polyarchy* as the regime that is both centered in the fulfillment of the will of the majority and in the respect and protection of dissenting minorities.

Beyond this central dimension, Dahl also emphasized that, in a *Polyarchy*, government should always guide its actions by citizens preferences, equally considered. Here is where both trends explained above converge: given the growing demand for better public services, the responsivity relationship between government and society should be services-based. In other words, because people want good education and health services, e.g., public officials should be held accountable by their delivery of these public services.

Therefore, it is possible to observe that one of the core dimensions of a *Polyarchy*, and of a democratic regime, is accountability, here understood as the strategies undertaken within the public sector to deal with the diverse set of expectations, generated inside and outside the organization, regarding its performance in the execution of its duties (Romzek & Dubnick, 1987). As further explained and detailed, government accountability can take place towards citizens or towards state technical agencies composed by non-elected members, also known as oversight institutions¹. In both cases, this high-quality services demanding society drove accountability beyond its original stricted-legal focus to a performance-based framework.

This shift from a type of responsivity centered on assessing the legality of public administration acts to one more preoccupied with public sector's performance on delivering public services was accompanied by an ever-increasing tendency of importing business techniques to the private sector. These techniques began to be used not only by public agencies in order to enhance its management and finalistic outcomes but also by oversight institutions, which started to control government's results through performance gauging tools originated within the private sector.

It is possible to expect that, in practice, there should be a strong connection between public sector's performance measurement (PM) and accountability. Kloot (1999) expressed this relationship satisfactorily: "Not using managerial control, performance indicators and effectiveness indicators reduces accountability to the local community. Communities need to be assured that their resources are being used efficiently and effectively in providing the right service at the least (KLOOT, 1999; p. 580)"². We go as far as to state that PM is the practical translation from the theoretical discussion of accountability to public administration's factory floor. That is, we expect that the theoretical discussion about government responsiveness and the importance of holding public officials to account is being translated to public sector's everyday practice through performance assessment models that were mainly imported from the business world. In a public sector environment defined by service quality, accountability

¹ It is important not to forget about parliamentary oversight.

² That is why performance measurement tools develop the surveillance capacity of civil society, qualifying vertical accountability.

tools should assess government's performance. Hence, PM is one of these tools that allows the responsivity relation to happen.

This work is organized in two main parts. The first part is a structured review of the literature regarding performance measurement in the public sector. Our main objective with this review is to check whether the academia presents an accountability approach to public sector's performance measurement, verifying the relationships between accountability and performance measurement. In other words, are the PM academics studying the responsivity implications regarding PM? To investigate this issue, we chose seven international high-impact journals mainly from the public administration field and, through a stepwise process, we have selected 37 papers published between 2012 and 2017 that had public sector PM as its main theme. Moreover, we constructed a theoretical framework that represented this literature, that is, that represented its functioning dynamic.

The second product of this work is a field study regarding Brazilian subnational court of accounts, that from now on will be referred as Tribunal de Contas (TC). State of São Paulo's Tribunal de Contas (TCESP) has created an unprecedented index, called Municipal Management Effectiveness Index (iegm), that measures municipal government management performance regarding six dimensions: education, health, planning, public finances, civil defense and e-gov. Every year, each municipality from the State of São Paulo complete a series of forms regarding local authority's management of the six dimensions cited above. These forms are sent to TCESP, where they are analysed by its technical staff. Periodically, TCESP publishes the index result for each municipality; the global grade and a dimension-specified grade. All the grades vary from A (highly effective), to B+ (very effective), B (effective), C+ (adjustment phase) and C (low level of adequacy).

In this field study we have decided to focus only on the educational area, given its greater structure and higher budget compared to other areas. We have visited three education departments (secretarias municipais de educação) of municipalities located at the metropolitan area of São Paulo. In these visits we performed semi-structured interviews with middle-ranking bureaucrats and municipal education secretaries. In summary, we tried to

assess what was the relationship between the department and the i-Educ (the name of the educational dimension of ieg-m). Given the great effort that municipal departments of education must have to gather all the data needed to fill in the forms, we wanted to evaluate if the departments were benefiting from the index by making any kind of use of the ieg-m to improve its internal management and finalistic outcomes or if it was just another bureaucratic set of actions. We also wanted to assess the type of relationships established between i-Educ and pre-existing internal performance measurement tools from the local departments. Besides, as it will be further observed, the insights gained from this exploratory research addresses the question of accountability and PM because IEG-M is an index that exists within a broader movement of moving external control (the one that is made by TCESP) beyond the traditional legal evaluation to a performance and results assessment. That is why the two parts of the work are strongly intertwined.

This research intends to contribute to the broader study field of public sector accountability, specially regarding the audit-auditee relationship, and government performance measurement by providing both theoretical and practical inputs. The paper is organized as follows: in the next section the theoretical background that has based the research is presented and next the research methods are explained in greater detail. Following the research strategy, the results of the research are evidenced and discussed. Finally, in the last section, we present our concluding remarks.

2. Theory

2.1 The so called classic and modern controls:

The advent of the modern state was responsible for an important change in several of the world's political regimes, especially in Europe. Over time and with the enlargement of the state's scope of action, the quantity and the diversity of state's attributions and responsibilities increased substantially. In the heart of the study of this complex movement, O'Donnell (1998) added a very important element to be considered in this enlarged state understanding: the question of horizontal and vertical accountability. While the horizontal accountability concerns the existence of agencies and state organ networks (oversight institutions) that supervise and sanction state measures, the vertical type of accountability is the one constituted by citizens and the organized civil society that acts in relation to the state and its agents, whether elected or not (O'Donnell, 1998).

A distinction between classic and modern controls must be made. According to Groisman and Lerner (2006) what characterizes classic controls is the comparison between actions and norms, with the objective of verifying if the former respect the latter. Therefore, classic controls are sustained by the assumption that the mere compliance with rules and regulations will lead to a rational and efficient action by the public employee. The typical classic control is the account control.

The modern controls concept, a result-oriented type of control, was forged through the public administration development process and the establishment of the idea that the central vector that should drive public management is the effective delivery of high-quality public goods and services rather than just rule-based administrative acts. This type works as an *ex post* evaluation that assess the performance of the public sector by the results it presents (OECD, 1996). While the classical controls are based in the juxtaposition between acts and norms, in the modern controls what prevail is the opposition between the established goals and the obtained results. Mora Quirós (2006) affirms that this kind of control both supports the decision-making process, so it can fulfill its goals and increase its performance, and accountability.

2.2 Bringing the accountability discussion to the *factory floor* of public management: the performance measurement systems

The discussion of accountability and modern controls was translated to the *factory floor* of public management through the use of performance measurement systems in a great variety of public sector organizations, a process that was greatly marked by experience and concept sharing between public and private sectors. Measuring performance of public services is an important strategy to materialize a very theoretical concept such as accountability into the day-to-day reality (the *factory floor*) of the public service.

Performance measurement is a process of quantifying action, measuring performance and the efficiency and efficacy of a given action (Neely, Gregory and Platts, 1995). This system may be structured and understood from three levels: (i) individual performance measures, (ii) the set of performance measures and (iii) the relation that is established between the performance measurement system and the environments in which they are situated. The literature converges in the direction that indicators and performance gauging models must always be constructed through the identification of the organization's mission and strategic objectives. Thus the performance measurement system must be a reflex of these two important elements, so that changes in these points result in changes and adaptations of the performance measures.

The performance measurement process can be divided in three steps: (i) system design, (ii) implementation and (iii) use. The first step consists in the identification of important objectives and in measures design. The second one concerns the put into practice process of the pre-elaborated systems and proceedings and the third step is constituted by the actions that guarantee the success of the implementation strategy and feedback, whose objective is to assure process improvement. There is a high probability that, through a natural evolution of its use, the measures of performance deviate from the organization's strategic objective. That is why feedback, responsible for guaranteeing the alignment between strategic organizational objectives and performance goals at all times, is so important (Bourne, *et al* 2000).

Still in this topic, there is a clear indication of the need of periodic assessment and internal and external monitoring that are able to deal with rapid changes of the internal and external environments of the organization (Bititci, 2000). These tools must operate together with an efficient revision system that gives way to the indications of the monitoring system. The author also highlights the need of an internal implementation system that can change organizational objectives and priorities so that there is always a symmetric relation between these elements and the performance indicators.

A literature review effort suggests two basic types of performance measures that are present in any organization: results - e.g: competitiveness and financial performance - and determinants of results, such as quality, flexibility, resource use and innovation (Neely, Gregory and Platts, 1995). These are the goals that the performance measurement will control and if by any chance the results demonstrate that the organization overflowed these parameters, action plans must be executed (feedback logic). It is thus possible to observe an enlargement of the performance measurement models frameworks, a process that started in the private sector and is more and more frequent in the public sector. This process is characterized by the new importance given to the characteristics that must be present in the organizations so that they are able to achieve these good results instead of a closed results approach. Still dealing with this process, it is worth mentioning the movement of creating performance measurement models that integrate different performance dimensions, including not only financial but non-financial aspects as well.

The Balanced Scorecard (BSC), by Kaplan and Norton (1992), is one of the most well-known examples of multidimensional models. The model is structured on the principle that a performance measurement system must feed the manager with informations that allow him to provide an in-depth investigation of the aspects of the organization related to (i) financial perspective, (ii) internal perspective, (iii) client perspective and (iv) innovation perspective.

The establishment and use of performance measurement systems in the public sector increased throughout the 20th and 21st centuries, specially after the 1980's fiscal crisis. This crisis, that had a major impact on almost every developed and developing country in the end of the 20th century, was one of the main motivations behind a series of reforms that aspired to

make government more efficient through an approximation between public and private sector practices (Randma - Liiv and Kickert, 2017), a process commonly known as New Public Management (Hood, 1995) that is still happening nowadays.

2.3 Performance Measurement Systems (PMS) in the context of local government

Given the fact that the practical part of this work concerns local governments, it is also important to dive into the academic literature regarding performance measurement at the local level.

Studying two small cities in New Jersey (USA), Hoontis & Kim (2012) identify the elements that precede the well succeeded implementation of governmental PMS. Firstly they observe that, according to previous academic expertise, the main factors that are necessary for the successful implementation of a PMS are: available resources; technical capability regarding the development of goals and measures; supportive leaders that demonstrate to civil servants the importance of using PMS; and the participation of internal and external actors. The authors add to this knowledge some empirical evidence gathered during their research: mayors or managers that have previous experience with performance measurement are more likely to implement a governmental PMS because this accumulated experience allows them to have a sense of the value and importance of this tool concerning performance development.

Hoontis & Kim (2012) also observed that the expectation of managers and mayors is that using PMS will improve management, reduce costs and increase customer satisfaction. Ironically, at the same time that this positive expectation is higher at smaller municipalities, given the closer administration-citizen relationship, the costs of adopting a PMS are also higher in small cities. Implementation costs, specially where municipal data systems were not compatible with the performance measurement systems that ought to be used (which implies a double provision of data), were one of the main elements of risk expectations, altogether with responsible and trustworthy use of performance information and the possibility of PMS being a source of bad performance accusations. The authors state that the development and distribution of a single PMS to similar-size cities would diminish the implementation costs. With regard to across-cities benchmarking, mayors and managers wish to have a simple PMS

that presents performance information of similar municipalities through indicators immediately comparable with the ones possessed by the city hall and its departments.

Still in this topic, Padovani, Yetano & Orelly (2010) investigated the factors that influence the implementation and use of such systems at the local context through the case study of four medium-size municipalities in Spain and Italy that have implemented PMS. First of all, they discovered that longer implementation periods are important because they allow adaptations to specific contexts and to singular managerial needs. Besides that, longer periods also increase the measurement system's legitimacy within the public machine. The authors have also noticed that it is important for the PMS' success that it continues improving and changing, like an alive and mutant system, even after the implementation phase is "officially" over.

Another important factor that influences the implementation of PMS is the so-called opportunity window. All the municipalities studied by the authors adopted a PMS after a route change, specially the election of new mayors. More generally, route changes consisted in the creation of environments where the political will to follow New Public Management logic or even the pragmatic need of gathering more information were present. Padovani, Yetano & Orelly (2010) have also observed that, in order to obtain success, a PMS needs to kick off as a management control system and, gradually, turn into a strategy formulation one.

Kloot (1999) studied how the performance measurement systems are being implemented by local governments within the state of Victoria, Canada. His research, published much earlier than the work just discussed above, confirms the insights presented by Padovani, Yetano & Orelly (2010) in respect to the opportunity window factor. Pressure made by state government for higher accountability levels in the bidding processes led to the adoption of PMS by Victorian local governments. Also, budget cuts promoted by Victoria government resulted in the recognition by the local governments of the importance of adopting PMS as a path to an adaptation to the new restrictive environment. The difficulties resulted in the PMS' use rise. Both factors can be understood as opportunity windows for the adoption of these performance assessment tools.

Furthermore, Kloot (1999) highlights that, within local governments that adopted the PMS, there were an observable shift regarding its organizational culture. The municipalities' administrations stopped having such a bureaucratic character and have become more similar to market-oriented service providers. PMS adoption, at the same time, induced and was caused by this change. Crantschaninov, Medeiros & Alves (2018), in a study of the determinants of the adoption of performance administration tools in the municipalities of São Paulo and Osasco, have also emphasized the importance of organizational culture.

Part of the PM literature analyses the impact that the use of this tools have on public organizations and the different scenarios and set of elements that augments the probability of changes to happen as a consequence of the adoption of PM. Concerning this first approach, Wang (2002) empirically measured the impact of performance measurement in US local governments. First, he concluded that PM helps public managers identify and specify services goals, strategies, management problems and its possible solutions. Although it had a substantial impact on organizational internal management, the use of PM tools presented a very limited impact in the communications between the administration and stakeholders and also in regard to the resources allocation decisions.

Using Performance Measurement Systems in public administration has a series of objectives. As explained earlier, two of these goals are the improvement of internal management and accountability. With respect to accountability, Wang (2002) found in his inquiry that the studied US local governments were not meeting the goal of accountability because managers are not effectively reporting administration performance informations to different stakeholders, such as the Chief of the Executive Branch, Legislators and, most importantly, to citizens. Although the general public is not interested in more technical measure of performance, citizens may be very interested in broader performance measures and indicators, such as life quality, crime rates, school performance statistics, etc. Hence, when confronted with this weak performance regarding PM's accountability dimension in the US municipalities, Wang (2002) states that one way to improve it is to merge these broader measures that attract more citizen attention into only one index.

Finally, it is also relevant to discuss briefly the relationship between PM and oversight institutions, given that ieg-m is a tool created and implemented by São Paulo's Tribunal de

Contas. Reichborn-Kjennerud & Johnsen (2015) studied the impact of Norway's Supreme Audit Institutions' Performance Audits on public administration. Although these are different jurisdictions, the study provides a series of important insights that are helpful in the context of the present research. They found that, definitely, performance audits leads to changes in norwegian public administration. The researchers also observed that instrumental, institutional and political factors define, in part, the chances of the auditee (public organization) of performing changes in order to adapt to the auditor indications. Reichborn-Kjennerud & Johnsen (2015) also observed that there is a directly proportional relationship between the auditee's belief that performance audit may bring benefits and the probability of changes driven by the performance audits. These changes are more likely to occur if the political leadership pressures the audited entities and if the performance report is used to hold the minister/secretary responsible. Finally, the authors show that the public employees who work closest to service delivery (in service delivery agencies, e.g) are more willing to make changes in response to performance auditing than the ones who work closest to the political power (in ministries or secretariats).

After this brief theory recapture, we can move on to display the research strategy and the methods that were used to build this work.

3. Methods

3.1 Overall methodology

The study relied heavily on a qualitative approach. The main justification for this option is that the first product of the work consists of a structured literature review, which by definition has a strong qualitative character. Secondly, we decided that, at that moment, given the fact that we were trying to make an initial inquiry about the use of IEG-M by São Paulo's municipalities, the best strategy was to perform semistructured interviews with municipal secretaries of education and public managers. By following this line of action we expect to capture the main perceptions of these actors regarding PM and the use of the tool elaborated by the TCESP by the municipal education departments.

3.2 Literature review

The structured literature review was guided by two hypothesis. The main one is that there is a gap between the areas of study of accountability and performance measurement: there are two important literatures, full of vitality, that do not establish a conversation with each other. Also, another hypothesis of ours is that translating performance management from the private sector to government was seen as a natural path to a more effective and efficient public sector. Therefore, the academic literature focused more on how to convert these business techniques into public management ones than dealing with the fact that using this type of dynamic had a direct impact on how public sector was being held accountable.

The literature review was a stepwise process initiated by the selection of seven important public administration, operations management and accounting research journals: *Accounting, Auditing and Accountability Journal*; *Journal of Public Administration Research and Theory*; *International Journal of Operations and Productions Management*; *Public Administration*; *Accounting, Organizations and Society*; *Public Administration Review* and *Journal of Operations Management*. The grade that journals had in the ABS impact list of 2018 - 3 or 4 - (CABS, 2018) was used as a selection criterion. We chose to use the ABS list instead of just the Qualis CAPES system because using the former would allow us to research

papers published in worldwide impact journals. Also, ABS is a British system, a country that has an important tradition of high-quality research in public administration.

Next, we analyzed the titles of all the scientific papers that were published from 2012 to 2017 in all the journals cited above and selected the ones whose titles indicated that the paper had something to do with the public services performance measurement thematic.

After this initial stage, in which we selected 163 papers, we read all the papers' abstracts with the purpose of making a finer selection of the works that ought to be further analyzed. The objective was to focus the literature review only on the papers that had the performance measurement of public services as its core subject. With this new filter, we read the 63 papers that were left in the selection in their full content, excluding around 30 papers that dealt peripherally with our topic of interest. In an effort to include seminal works in our structured review, we looked into the references sections of all the articles that remained after applying the last filter, reading the ones that applied to our criteria in their full content. By the end of all these filters we were left with 37 papers that were going to be specifically analyzed in our literature review.

It is important to clarify that one of the main objectives of the review was to assess whether there were accountability elements and discussions within academic works that dealt specifically with the theme of public sector PM or not. It is obvious that if our search magnifier is directed towards selecting the papers that had this thematic as its core subject, we would not find researches so directly connected to accountability. But, again, the purpose of the present work is to try to capture the accountability elements within a public sector PM literature.

Reading these 37 papers, we identified the new concepts and ideas that were developed in each of the works. Then we assembled all the identified concepts and through a categorization and subcategorization process we were able to create groups of dimensions and concepts that were similar and, sometimes, even the same.

The dimensions were defined as (i) Accountability, (ii) PMS design, (iii) PMS implementation, (iv) PMS use impact, (v) Impact on the results presented by the PMS and (vi) Goals. The concepts linked to dimension (i) are the ones that present an interface between the

performance measurement in public organizations discussion and themes related to accountability, such as public participation, citizen evaluation of government activity and performance-related sanctions on public officers. The concepts listed under the second dimension theoretically express a great myriad of elements that influences the design of performance measurement systems, such as stakeholder, organizational goals, etc. The third dimension represents this same idea, but concerns the implementation process of PMS. It is important to clarify the distinction between dimensions (iv) and (v). While *PMS use impact* refers to the various means by which the use of PMS in a public organization impacts its overall performance and the quality of the public services it provides, *Impact on the results presented by the PMS* concerns the impact that several elements, such as environmental ones, has on the results presented by the PMS, not taking in consideration the impact of the performance system itself.

Finally, the last dimension, *Goals*, refers almost in its entirety to concepts formulated in papers that try to understand the impact that organizational goals definition has on the PMS, mainly in the result it presents but also in other phases such as design and implementation.

3.3 i-Educ/ieg-m analysis

Before entering in the details of how we got the results regarding the second product of this work, it is important to clarify what may be obvious. Chronologically, first we did the literature review, and the review provided us with a number of insights - that will be thoroughly discussed in the results section - that motivated us even more to perform the visits and the interviews at the municipalities. In other words, the results of the literature review helped us build the research questions that drove our exploratory investigation at the municipal departments of education.

Through internet searches we gathered the email contacts of all the 645 municipalities that belong to the Brazilian state of São Paulo. Most of the time we were able to get the contact of the chief of staff of the municipal secretaries of education. But, in some cases, we were only able to get the email contact of lower officers, such as the head of communications

Then, we sent a standardized message presenting us as researchers and students of the São Paulo School of Business Administration at São Paulo's Getulio Vargas Foundation.

Unfortunately, we only received replies from three municipal education departments, even though we fired the emails twice. The responses came from São Caetano do Sul, Franco da Rocha and Embu das Artes. Also, we were able to get an answer from Santo André, as a result of a particular connection between one of the researchers and one of the officers within the department³. Due to a internal staff restructuring process, we were not able to conduct the interview in São Caetano do Sul. Hence, the cities that constitute the research object of the present work are Franco da Rocha, Santo André and Embu das Artes.

In Franco da Rocha and Embu das Artes we interviewed the Municipal Secretary of Education; in the second city we also had the opportunity to talk with the employee that was responsible for gathering the data required by TCESP to produce the i-Educ indicator. In Santo André we interviewed two bureaucrats responsible for the pedagogy area of the education department, who were intimately related to the indicator.

For the interviews we built a question script to guide our conversations with the secretaries and managers (see appendix A). In general, the questions reflected our research yearnings, solidified after the literature review. Vinculated to the research question regarding the functioning of the process of monitoring and assessment of service delivery in the subnational context, we elaborated a set of questions aimed at understanding the internal and pre-existing PM tools⁴ that were used by the municipal education departments. Reflecting the research question of how it is possible to put in practice an effective accountability dynamic through the use of PM, we elaborated a series of questions that tried to understand the relationships established inside the public administration, with other political parties, with the civil society and with the Tribunal de Contas as a result of the compulsory use of i-Educ/IEG-M. Lastly, the research question of how these indicator and index may be used to improve the PM process and service delivery was addressed by several topics of the question script that tried to understand how the municipal education department related to the i-Educ.

³ This fact did not constitute any kind of conflict that could harm the exemption of the research results: we interviewed other officers, without the presence of the officer who had the connection with the researcher. The interviewees had full liberty to express their opinions.

⁴ Here understood in its most broader and wider meaning.

In other words, we tried to understand how the process of gathering the data to fill the reports that needed to be sent to TCESP worked, how this obligatory need changed the internal functioning dynamic of the department, if it improved its internal management capacity, etc.

After each meeting we wrote an interview report that was later used in the construction of the analysis that is presented in the results section. In addition, after the three meetings we interviewed inspectors from TCESP to capture the vision of the Court of Accounts with respect to IEG-M.

4. Results

4.1 Literature Review

4.1.1 Dimensions and concepts

As explained above, the literature analysis was structured through the systematization of novel concepts that were identified in the 37 selected papers. Besides that, we classified these concepts that emerged from the papers into five analysis categories/dimensions: Accountability, PMS design, PMS Implementation, PMS use impact, Impact on the results presented by the PMS and Goals⁵. To identify a paper with one of these categories means that there is a strict convergence between the paper central point of discussion and the category.

The analyzed papers dealt more with specific PMS characteristics than the relationship between this tool and Accountability, what can be clearly observed when the frequency of directed Accountability-linked papers is verified. According to Table I, these types of papers do not exceed 16% of the 37 works that were selected. However, some papers that focused on PMS characteristics also presents dimensions that addresses the issue of accountability such as the relationship between power and local managers performance (Yang and Sven, 2012) and the role of clear goals establishment on citizen expectations towards public services (Chan Su, 2012).

⁵ It is important to make a distinction between the PMS use impact and Impact on the results presented categories. The former concerns the impact that performance measurement systems have on the overall performance of an organization while the latter deals with elements that affects the results presented by an organization's performance measurement system.

Table 01- N° of selected articles per dimension (elaborated by the authors)

Classification	N° of selected articles	%
Accountability (Fowler & Cordery, 2015; Brandsma & Schullemans, 2013; Humphrey & Miller, 2012; Overman, Genugten & Van Thiel, 2015; Reichborn-Kjennerud, 2013; Olsen, 2016)	6	16,22
PMS Design (Andersen, Heinesen & Pedersen, 2015; Barbosa, Pozzebon & Diniz, 2013; Brignall & Modell, 2000; Sutheewasinnon, Hoque & Nyamori, 2016; De Bont & Grit, 2011)	5	13,51
PMS Implementation (Bobe, Mihret & Obi, 2017; Jaaskelainen, Laihonen & Lonqvist, 2014; Meier, <i>et al</i> , 2015; Gerrish, 2016; Yang & Modell, 2012; Nielsen, 2013; Jennings Jr. & Hall, 2011; Destler, 2017; Nutley, <i>et al</i> , 2012; Martin, 2016)	10	27,03
PMS Use Impact (Hvidman & Andersen, 2013; Liang, 2016; De Bont & Grit, 2011)	3	8,11
Impact on the Results Presented by the PMS (Lee & Whitford, 2012; O'Toole, Jr. & Meier, 2014; Walker & Andrews, 2013; Nielsen, 2014; Neshkova & Guo, 2011; Vashdi, Vigoda-Gadot & Shlomi, 2012; Pedersen, 2015; van der Kolk, ter Bogt & van Veen-Dirks, 2015)	8	21,62
Goals (Favero, Meier & O'Toole, Jr., 2016; Rutherford & Meier, 2015; Jung, 2012; Anderson & Stritch, 2016; Jacobsen, Snyder & Saultz, 2014)	5	13,51

Source: Author

Analysing the identified concepts, that were the base for the creation of the dimensions, it is possible to observe that approximately 16% of the 18 identified concepts are directly related to Accountability, 14% related to PMS design, 27% related to PMS implementation, 8% related to PMS use impact, 22% related to the Impact on the results

presented by the PMS and 14% related to Goals. It is also possible to observe that the majority of the concepts identified in the analysis of the papers deals with the elements that impact the results of the performance assessments realized by the PMS. The 18 concepts and the dimensions under which they were agglutinated are summarized in Table II.

Table 02 - The six dimensions and their related concepts (elaborated by the authors)

Accountability	PMS Design
Relationship between sector and accountability	E-gov and performance
Accountability Measurement Method	Stakeholders and the PMS dimension
Transversality of areas of knowledge	Local management and performance
Citizen performance evaluation	
PMS Implementation	PMS Use Impact
Institutional factors	Self-evaluation of performance
Impact on the Results Presented by the PMS	
Social participation and performance	
Performance and organizational resources	
Performance management impact	
Performance and information management	
Organizational environment and performance	
Bureaucratic commitment and performance	
Power	

Management control through management performance measures
Goals
Goals and performance

Source: Author

Concepts that deal with the accountability dimension are also significantly present, even though in a much smaller number than the previous category and, although the majority of the papers are classified under *PMS Implementation*, only one concept (Institutional factors) is vinculated to this dimension. This apparent contradiction shows that the academic production about the implementation of the PMS focus heavily in Institutional factors.

Several of the concepts that were gathered during the analysis contribute to the accountability discussion in a peripheral way. Disconsidering the ones that are under the accountability dimension umbrella, the concepts addresses result-related subjects but also deal with ideas that are part of the accountability discussion. For instance, the concept under the name *Stakeholders and the PMS dimensions* investigates how different constellations of stakeholders may affect the design of the PMS in a public organization. Discussions concerning stakeholders influence are strongly related to the accountability arena but the papers that established this concept do not have as its goal to understand the relationship stakeholder-PMS with an accountability perspective. In fact, the interest is to discover the impact on the PMS design. Again, the focus is strictly on the performance measurement system and not in the *holding public officials accountable* dynamic. But, even though this is not the focus of these researches, there are important contributions that can improve our understanding of the interlinkages between accountability and PMS.

It is worth observing that the closer the analytical dimensions gets from the strictly result-focused dimensions, less frequent are the concepts that also discusses accountability. For instance, on the *Impact on the results presented by the PMS* dimension, only one of the related concepts also discusses the accountability thematic (*Social participation and*

performance). This patterns also corroborates with the hypothesized gap between the two literatures.

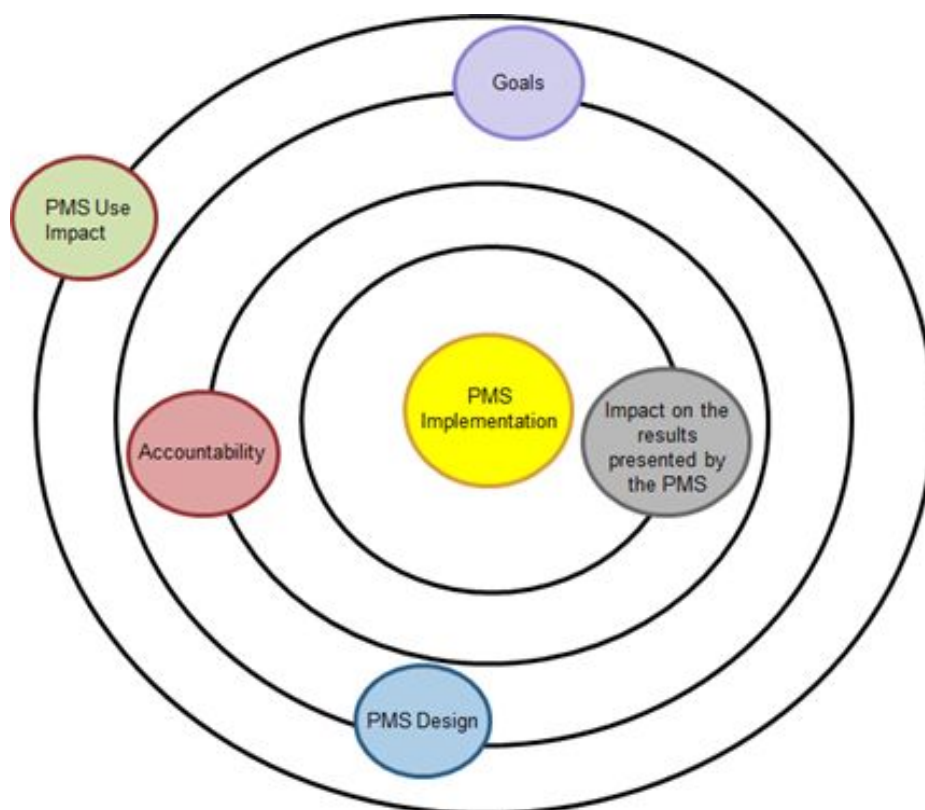
Other evidence extracted from the literature analysis that indicated this lack of dialog between PM and accountability literatures is the fact that while only 8% of the papers pertained to the *PMS use impact* dimension, 22% of them belong to the *Impact on the results presented by the PMS* dimension. Given the fact that the former refers specifically to the impact of the performance measurement tool (also understood as an accountability tool) while the latter concerns the results itself, it is worth noting that the dimension of PMS impact that deals more specifically with results has more papers than the one that also addresses results but with an accountability optics. This observation corroborates the perception that the literature efforts are heavily concentrated on results.

Finally, rather than this absence of accountability-related elements and/or analysis, what stands out the most is the lack of an accountability approach to PM's elements and tools. In other words, the literature does not approach PM as an accountability tool. The academic field is so immersed in the practical side of this discussion that it appears to forget how performance measurement systems, even though their origin is within the private sector, are the translation to the "real life" of public administration of a theoretical accountability dimension. Evidence shows, therefore, that this literature does not have the background knowledge that one of the theoretical foundations of this results-oriented, reporting results-oriented and efficient public administration that is being studied is accountability.

4.1.2 Framework

Investigating the concepts raised by the studied literature, it was possible to develop a framework of the operating dynamic of this literature:

Figure 01 - The framework elaborated by the authors



Source: Author

We believe that a solar system framework is the best way to represent the behavior and the operation of the performance measurement system literature. From inside out, the dimensions are in descending order of quantity of linked papers. In the place of the Sun, as an agglutinating and attraction pole around which other elements orbit, we placed the *PMS Implementation* category. In the next circle, and not only in the next circle but very close to the sun, we placed the *Impact on the results presented by the PMS* category: we decided to place it like that because the majority of the papers that were analyzed raised discussions about how different aspects that are present in a public organization may influence the performance in the provision of public services, both in the sense of the results itself and in the way that the PMS is implemented. Also, we verified that these thematic guides

practically all the papers that were analyzed. Even when these are not the main focus of the work, the dimension of the performance measurement results and implementation, even more than the assessment process in itself, is central. After *Impact on the results presented by the PMS*, the other categories are positioned so that the closer to the Sun the greater its frequency among the identified concepts.

The structured literature review allowed us to confirm the expectation reflected on our first hypothesis. The systematic lack of papers that investigated how PM may improve or worsen accountability and the low number of publications that mixed performance and accountability discussions made evident the lack of dialog between these two literatures. Almost all of the papers under the *Accountability* dimension umbrella do not present the linkage between this theoretical concept and performance measurement.

Our second hypothesis (the conception that translating performance management from the private sector to government was a natural path to a more effective and efficient public sector directed the literature to the mechanisms involving the translation of business practices to public sector, leaving aside accountability), was partially confirmed. The perception that these high-quality public policies could be delivered through this conversion process indeed had an important impact on the academic discussion, what can be seen by the large amount of papers dealing with PMS implementation and results. However, this justification based on the understanding of the relationship between these business practices and good public services is not able to explain the absence of the accountability-PM relationship. Even though this view made the literature focus heavily in the PMS implementation, this does not justify why a connection between these same outcomes and the accountability arena could not have been made.

We have identified that these two literatures function separately, with very little dialog, however, we were not able to identify for sure why this connection has not been made yet.

4.2 i-Educ/ieg-m

4.2.1 The index explained

According to Loureiro, Teixeira & Moraes (2009), since the promulgation of the new Brazilian constitution in 1988 and the creation of the Lei de Responsabilidade Fiscal in 2000, it is possible to observe a trend regarding national and subnational court of accounts. The courts are increasingly emphasizing the so-called modern controls. In other words, instead of only verifying if public entities actions are in accordance to the law, these courts are also assessing the performance of these public organizations regarding the achievement of its expected goals. That is why performance audits are becoming more common in Brazilian Tribunais de Contas.

Following this trend of external control and horizontal accountability modernization, State of São Paulo's Court of Accounts (TCESP) has developed the Municipal Management Effectiveness Index (ieg-m), applied since 2014⁶. As it is informed in its own website, ieg-m is a performance index that assesses the effectivity of the public policies and activities developed by State of São Paulo's municipalities, emphasizing management practices, infrastructure and processes. The goal is to verify if the vision and strategic objectives of the municipalities were effectively achieved (TCESP, 2019).

The index is composed by seven dimensions, or seven indicators: i-Educ (municipal public education area), i-Saúde (municipal public health area), i-Planejamento (municipal planning area), i-Amb (municipal ambiental management area) and i-Gov TI (municipal information technology governance). Each of the indicators and the aggregate index presents one of the following grades: A (highly effective), B (effective), C+ (in phase of adaptation) and C (low adaptation level).

Annually, municipalities are obliged to fill a series of forms, that can also be understood as questionnaires, with municipal management data regarding each one of the seven areas. Usually, a central department within the city hall receive the forms and distribute them to the respective areas; these areas fill the forms with the informations needed and send them to the same central department, responsible for gathering all these documents and submitting them to TCESP. Every year the inspectors from the Court of Accounts visits all municipalities to confirm if the informations presented in the ieg-m forms are truthful.

⁶ The use of ieg-m has spread and, nowadays, all Brazilian state court of accounts apply the index to its jurisdictional municipalities

Besides this visit, the grades that each city receive are one of the tools used by the inspectors to decide whether this municipality will receive an official visit from the court's auditors or not. That is, ieg-m is also important for the surveillance planning process. One year later, the index grades - in its aggregated and disaggregated forms - are published online. This delay is due to the confirmation of the accuracy of the information process that happens with every municipality.

Before getting into the analysis of the interviews it is important to summarize the main and basic informations regarding the three municipalities that are part of this study. In Table 03 below we present the Municipal Human Development Index (MHDI), population and the results of the IEG-M and i-Educ available for all three cities.

Table 03 - Basic data from the three municipalities

Municipality	MHDI (2010)	População (2018)	IEG-M	i-Educ
Franco da Rocha	0,731	152.433	B+ (2014)	A (2014)
			B (2015)	B (2015)
			B (2016)	B (2016)
			C+ (2017)	C+ (2017)
Embu das Artes	0,735	270.843	C+ (2014)	B (2014)
			C+ (2015)	C+ (2015)
			B (2016)	B (2016)
			C (2017)	B (2017)
Santo André	0,815	716.609	C+ (2014)	B+ (2014)
			C (2015)	C+ (2015)
			C (2016)	C+ (2016)
			C (2017)	C+ (2017)

Source: Author, with data from Atlas do Desenvolvimento Humano no Brasil, IBGE, TCESP

Below, we present the main topics that were raised during the interviews realized in Franco da Rocha, Embu das Artes and Santo André.

4.2.2 Franco da Rocha

The interview allowed us to observe that, for the department of education, there is a strong disconnection between the assessment that is made by i-Educ and the reality of the municipal education system. One single grade is insufficient to explain the complexities involved in that reality. Also, the aggregated character of i-Educ does not allow it to have an evolutionary perspective that demonstrates the process of changes of the educational authority. *A to C+ to B*, for example, does not say a lot about what has really happened. Besides, TCESP does not provide any guidance as to how the secretary of education should act to improve the performance evaluation.

Another perception is that, sometimes, there are differences between what the education authority understands as good quality and what i-Educ values as quality. And these differences results in bad evaluations at i-Educ. Furthermore, there is also a political dimension concerning the publication of the results because bad performance indicators are often used by opposition parties to attack the entire municipal government. This possibility frighten the staff.

Differently than with i-Educ, the secretary of education is able to connect the Basic Education Development Index (IDEB) results with the reality of what is being done in the department. Also, there are a series of managerial actions that are activated as a response to the IDEB results, especially when they are poor. Therefore, IDEB, on the opposite of i-Educ, is central both for the day-to-day practice of the secretary and for the long-run strategic planning. Given the fact that IDEB measures the finalistic education outcomes and i-Educ education management, maybe what is needed in the department is the improvement of the analytical capability to make the connection between management, infrastructure and processes with educational and pedagogical outcomes.

It is possible to realize that, for the staff of the education department, i-Educ is not completely useful, given that it does not present a clear connection with the reality of the actions being performed. The main reason for this disconnection is the way by which the performance assessment results are presented - an simplistic, non evolutionary way that does not express any sense of process. This feeling is corroborated by the absence of feedback

from TCESP, which does not present an explanation for the results and neither gives instructions as to how it is possible to develop. Therefore, i-Educ, in this case, is only being used by São Paulo's Court of Accounts as a way of monitoring the performance of the municipality's public education; it is not absorbed by the education department as an internal assessment tool and its compulsory application is not generating any kind of structured and systematic managerial actions.

4.2.3 Embu das Artes

The officer who talked to us only participated in last year's filling of the i-Educ forms. The officer emphasized how difficult it was to gather the required information. The secretary did not even have some of the needed information and several informations were spread across the department and school units. Other informations that are required by the Court of Accounts does not belong to the secretary jurisdiction. This lack and disorganization of informations, that is due to structural problems and management issues, had a negative impact on the performance assessment.

According to the interviewee, i-Educ has not changed what the department already does in terms of management because it is just a formal procedure of inspection realized by TCESP. Besides that, the filling process evidenced the need of better information systematization procedures. Although the department of education has a lot of informations, they are displayed in a series of different systems. There was a clear incompatibility between the secretary's systems and the way the forms were organized. This incompatibility made the process even more difficult. According to Hoontis & Kim (2012), this type of incompatibility is considered a main PMS implementation cost. The officer informed that this difficulties resulted in a kind of trauma within the department; that is why new actions regarding the information systems are being planned to facilitate the next form filling process.

Therefore, it is evident that the main issue regarding the relationship between Embu das Artes and i-Educ is the lack and disorganization of informations, which made it so difficult for the managers to answer court questions. By the answers we received, it is possible to imagine that the secretary of education will improve its methods of gathering and organizing information. This possible set of actions has a clear managerial character.

Although the interviewee expressed a view that i-Educ had no managerial impact over the administration, it is possible to see that, ironically, the difficulty in providing TCESP with the needed information to calculate i-Educ may have indirectly enhanced the managerial capacity of the education department. Despite of the fact that its improvement may happen, there is no evidence that i-Educ can be used as a traditional performance measurement tool by Embu das Artes.

4.2.4 Santo André

The officers stated that the feedback provided by the Court of Accounts (a “cold” grade) does not help to develop the administrative capacity of the secretary of education. However, the process of gathering the information needed and answering the form questions is very important because it promotes a series of improvements: just observing the final grade does not help, but filling the forms clarify what is being made and its problems. This process works out as an action targeting dynamic. For instance, answering the questions made the staff realize that there were flaws in the literacy teacher training. As a reaction, one of the pedagogical teams started to focus on this thematic. The questionnaire is, therefore, qualifying administrative action through alert points in day-to-day management. Occasionally, the secretary personnel found out that they were simply unable to answer some of the questions from the form, what indicated that the department was not fulfilling all of its duties and functions.

It is clear for the department where it needs to go in the mid and long terms because these goals are already defined in the municipal education plan. Thus, answering the questionnaire works as a way of qualifying the managerial path towards the achievement of these goals. According to the interviewees, the daily life of the organization is marked by the constant demand to “put out fires”, to resolve urgent needs. The moment of the form filling and informations gathering forces the staff to think, reflect, plan and escape from this extremely agitated routine. Also, the fact that these important reflections are made during the filling of the forms and not when the grade is received - even because the grade is insufficient to promote such reflections - represents a way out of the i-Educ inadequate periodicity (grades published only every two years).

Unlike in the other two cities, the interview made in Santo André emphasized positive aspects in respect to the application of i-Educ. Mainly, the officers showed that answering to the TCESP questionnaires helps a lot to identify internal management problems that otherwise would not be seen. It is interesting to reflect to what extent these positive aspects only exists because of the greater administrative capacity of Santo André - a relatively bigger city than Franco da Rocha and Embu das Artes.

4.2.5 Some notes regarding the interviews

The interviews with the officers from the three cities made clear that i-Educ, when understood only as the final grade published by TCESP, cannot work as an internal performance measurement and management tool for the municipal education departments. A simple grade does not provide public managers with the input that is needed to make changes or to perform different managerial actions. It seems to be unanimous among the interviewees that the feedback provided by the Court of Accounts - both through the grades and through the recommendations - is completely insufficient to stimulate in a structured way differential set of responses and managerial reactions. In Franco da Rocha, this could be observed by the constant reminder of the distance between the i-Educ assessment and the reality of the administration. In Embu das Artes and Santo André this was observed by the quick discard of the possibility of any improvement led by the publication of the grades. As already discussed, Bourne and Mills (2000) show the importance of feedback.

This discard makes a lot of sense if we take into consideration that an usable performance measurement system must have at least a set of indicators and not only one grade. Hence, the discussion of the impact of the i-Educ for internal management is left with the discussion regarding the assumption that filling the forms is *per se* an important managerial action. This could be clearly seen in the interview made in Santo André. Although the interviewee from Embu das Artes not seemed to realize that the actions taken as a response to the great difficulties faced by the department's staff to gather all the informations had a heavy managerial background, the importance of filling the forms was also observed in this case.

Therefore, it is possible to state that i-Educ is not being absorbed by the municipal departments of education as a tool that compose a performance measurement system because the way that it is made and displayed does not allow this to happen. However, the fact that the making of this indicator involves the departments answering a wide questionnaire represents a significant potential of performance self-evaluation and change by municipalities. Although not in the most anticipated way, it was possible to observe in the study that i-Educ, when seen as the process of gathering informations and answering questions and not as an indicator, enhanced the managerial capacity and the departments' self assessment of performance in two of the three case studies. Thus, it presents an important performance measurement potential.

The conversation with one of the inspectors from TCESP showed us that the limitations we found on the interviews were compatible with the purpose of the Court of Accounts for creating the IEG-M. The index was elaborated with the goal of evaluating the management performance of São Paulo's municipalities, given the growing understanding that the role of an public audit institution is to assess the performance of public organizations. TCESP was not necessarily planning to create a tool that could be used by municipalities to improve their management capacity; the target was to give a tool to the court that enlarged its scope of fiscalization. That, and the fact that ieg-m is mainly used to define which municipalities will receive audit visits, explains, in part, the simplistic feedback that is provided from TCESP to the respective prefectures. Another indication that the court mindset concerning the creation and use of the index did not took into consideration its potential of increasing municipalities' internal capacity management is the fact that the ieg-m results of the 2016 exercise, for example, are only published in 2018. This periodicity unables a performance measurement system and an effective administration of performance. Initially, the time interval was smaller, but even an one year gap is too long to constitute a performance measurement system.

After analysing the interviews, it is important to reflect on the insights generated by the study in relation to the research questions. The case studies allowed us to understand more deeply how the process of monitoring and evaluating public service delivery works in the municipal level. First of all, municipalities presents different maturation levels of monitoring systems. While some have a diversity of internal indicators, indexes and methods of assessment, others don't have almost none measurement tool, depending completely on

external evaluations, such as IDEB. A comparison between Franco da Rocha and Embu das Artes and Santo André shows these inequalities. While in the first there is a strong absence of even simple evaluation methods, in the second and third cases there are a much wider and more complex set of evaluation tools and strategies. However, these tools are more concentrated in evaluating the quality of the education and not the quality of management.

There is a clear accountability dynamic that involves the departments of education and TCESP. Even though there are a series of flaws and difficulties regarding this fiscalization relationship, definitely *ieg-m* has qualified the court's inspection over municipalities, bringing it closer to the reality of public management. This is positive because, by doing this, the Court of Accounts is showing that it understands that one of its main roles is to hold public management accountable for delivering what is essential to citizens, that is, public services. Only assessing the legality of public actions is not enough.

Also, the interviewees informed us that they fear the possible bad public repercussion of the publication of a negative *i-Educ* evaluation. Although this fear represents some bad incentives for the assessment process, it is a good sign that the performance measurement by TCESP is also overflowing a strict performance measurement framework because it has a clear citizen-related-accountability dimension, following Wang's (2002) discussion. The fact that managers expect the publication of the *ieg-m* grades to have real impact over citizens' government perceptions goes according to the author's expectation that merging indicators into expressive indexes may cause more impact to citizens.

Certainly there are several changes that TCESP could make to improve its measurement of municipalities' management. Given the fact this study focused almost completely in the municipal department of education and not in the court, it is important to weave some comments on how *i-Educ* can be used by these departments to improve both their own measurement of performance and the delivery of education public services. The interviews emphasized that the potential of developing these two factors cited above is almost entirely concentrated in the moment when the departments answer the court's questionnaire. Along other motivations, this moment may show to the administration the flaws of their internal information systems. Knowing that next year the staff will have to go through the same process again may force the department's personnel to prepare new strategies that

improves the knowledge that the secretary has over its own informations, what may include novel forms of data collection and gathering. Besides, as shown by Santo André, having to prepare the data and answer questions may compel the staff to stop and rethink their own management practices and performance. If it is in the interest of the court, TCESP should engage the municipal departments to pay attention to the potential that this moment represents for their own management capacity improval.

5. Conclusion

From the point of view of municipalities, there are a series of limitations involving i-Educ and ieg-m. The index as it is will not be transformed in a performance measurement tool used by prefectures and its departments. The chances of using this index as a central component of a performance measurement system are even lower. The index was not built for this goal and all of its main characteristics does not allow it to be used by municipalities for this purpose. However, this doesn't mean that ieg-m don't improve or won't help to develop the management practices within public organizations. As observed above, the mere fact that the secretary's staff needs to mobilize to answer the questionnaires may have a positive influence over public management.

This study also showed that there are performance measurement tools that used in the public departments. The presence and use of these tools varies intensely from municipality to municipality. However, the absence of a strategic coordination between these assessing mechanisms is almost equally spread across the studied departments, altogether with no system of indicators that is recurrently used for management purposes. Administration of performance is a virtually non existent practice. When it exists, it is of poor quality and in a non-systematic way. Given this negative picture, the Court of Accounts and other governmental entities such as the state government could help municipalities implement these performance measurement systems.

To answer the i-Educ questionnaire, the secretaries have to gather a series of informations that would, for sure, be part of a performance measurement system. Hence, the basis of and the pathway towards a PMS already exists. Very briefly, what is needed for the construction of such systems and its effective use is training, political will and the support and help from other institutions, like TCESP. These elements are very hard to get, of course, but implementing performance assessment systems is extremely important for the modernization of the administration and for the delivery of better public services. To improve, public management must know exactly where it is failing and, to act over this failures, it needs, along other things, to have the informational basis provided by the PMS.

Regarding the literature review, it was possible to observe that the literature is already focusing on the so-called "factory floor" given the fact that it dedicates itself intensely to the results of the PMS. In addition, the literature does not focus on a deeper reflection about the

design of the PMS and on how it can be more effective. Instead of doing that, the literature already advances to the implementation and results of the performance evaluation and to the elements that influence them. The theoretical framework that is being built is, therefore, increasingly implementation-centered and results-centered.

Thus, the state of the art of the academic production related to PMS in the public sector is not worried in showing how the design of different measurement models may influence in a broader way the accountability process. The central worry is with the results presented by the already existing measurement systems in the institutions, and the relationship between these same results and the strategies used to adopt and implement the performance measurement systems. If the literature were predominantly centered in the transition from the theoretical dimension of accountability to the public administration “factory floor”, the central theme would be the different PMS designs and how they respond to the holding government agencies accountable demand and not the performance results itself, as it was possible to verify in this study

Hence, it is possible to conclude that the academy is studying a lot the implementation and results of the PMS and how they are affected by different realities, but it is leaving aside the discussion about how it fits in the accountability thematic. This lack of dialog between the two literatures ends up leaving aside the discussion of holding the state accountable towards the civil society regarding the results of the performance assessments. Moreover, the literature demonstrated a certain flaw on the construction of a perception of conjuncture by not clarifying the role of the oversight institutions responsible for executing performance evaluations and also what is the role of PM inside a broader public administration system. Most of the literature pays attention to what already exists in terms of performance measurement and tries to gauge if it is bringing results or not but it does not pay attention in a systematic way to the oversight institutions neither to the construction of a PMS as part of a broader and deeper accountability process.

When the insights from the literature review and from the case study are combined, it is possible to conclude that there is no such thing as a vacuum in reality. If the PM literature presents a gap between the concept of performance measurement and accountability, in reality, this connection has already been made. Accountability is already being translated to

the public administration factory floor by performance measurement. Ieg-m in general and i-Educ in particular are operationalizing the theoretical discussion of accountability in the Brazilian subnational administrative environment. The reality shows us that both trends that were theoretically described by Dahl (1997) - democratization and public services delivery - are converging. Even with all its problems, when TCESP applies ieg-m and assesses the delivery of public services, it is enhancing democracy and approximating the normative aspects of democracy with the reality of society and public administration.

Therefore, the literature needs to advance and explain in a satisfactory way the accountability elements that compose the practice of measuring government performance, both internally and externally. In the case of the index studied in this work, this could be done by investigating more deeply the effects of the publication of ieg-m on the public perception of governmental performance, by understanding the interaction of public administration and oversight institutions regarding the index, by assessing the incentives generated by ieg-m for a greater performance by municipal public management, etc.

Lastly, it is important to remember that this study has some explanatory power limitations. First, the literature review analysed only the PM “side” of the academic production. That is why future research could engage in the same literature review but, instead of diving into PM papers, it could look for performance measures discussions within an accountability-focused literature. Secondly, the results of the study conducted in the municipalities are limited because only three departments of education were visited. Future researches could enlarge this scope of assessment and integrate interviews with data analysis from other indexes and databases. As expected, the suggestion for a future research agenda is to investigate better and more thoroughly the accountability implications of PM.

6. References

BITITCI, Umit S.; TURNER, U. Trevor. BEGEMANN, Carsten. Dynamics of performance measurement systems. *International Journal of Operations and Production Management*. v. 20, n.6, p. 692-704, 2000.

BOURNE, M.; WILCOX, M.; NEELY, A.; PLATTS, K. Designing, implementing and updating performance measurement systems. *International Journal of Operations and Production Management*, v. 20, n. 7, p. 754-771, 2000.

CHARTERED ASSOCIATION OF BUSINESS SCHOOLS. Disponível em: <<https://charteredabs.org/academic-journal-guide-2018/>>. Acesso em: 30 de junho de 2019.

CRANTSCHANINOV, T. I.; MEDEIROS, A. K.; ALVES, M. A. Elementos Determinantes e Impactos da Adoção de Instrumentos de Administração da Performance no Setor Público: Estudo Comparado entre os Municípios de São Paulo e Osasco. *Revista Gestão & Planejamento*, v. 19, n. 1, p. 20-33, 2018.

DAHL, R. **Poliarquia: Participação e Oposição**. São Paulo: Editora Universidade de São Paulo, 1997.

GROISMAN, E.; LERNER, E. Responsabilização pelos controles clássicos. In: Bresser-Pereira, L.C. Grau, N.C (Ed.). **Responsabilização na administração pública**. São Paulo: CLAD/Fundap, 2006.

HOOD, C. “The ‘New Public Management’ in the 1980’s: variations on a theme”, *Accounting, Organizations and Society*, Vol. 20, N. 2/3, p. 93-109, 1995.

HOONTIS, P.; Kim, T. Antecedents to Municipal Performance Measurement Implementation. *Public Performance & Management Review*, v. 36, n.1, p. 158-173, 2012.

KAPLAN, R.S.; D. P. Norton. The Balanced Scorecard: Measures that Drive Performance. *Harvard Business Review*, p. 71-79, 1992.

KLOOT, L. Performance measurement and accountability in Victorian local government. *International Journal of Public Sector Management*, v. 12, n. 7, p. 565-584, 1999.

LOUREIRO, M. R.; TEIXEIRA, M.A.C.; MORAES, T.C. Democratização e reforma do Estado: o desenvolvimento institucional dos tribunais de contas no Brasil recente. *Revista de Administração Pública*, v. 43, n. 4, p. 739-772, 2009.

MORA QUIRÓS, M. Responsabilização pelo controle de resultados. In: Bresser-Pereira, L.C. Grau, N.C (Ed.). **Responsabilização na administração pública**. São Paulo: CLAD/Fundap, 2006.

NEELY, A.; GREGORY, M.; PLATTS, K. Performance measurement system design: a literature review and research agenda. *International Journal of Operations and Production Management*, v. 15, n. 4, p. 80-116, 1995.

O'Donnell, G. Accountability horizontal e novas poliarquias. *Lua Nova*, v. 44, p. 27-54. 1998.

OECD. Towards Better Measurement of Government. OECD Working Papers on Public Governance. *OECD Publishing*. 2007.

PADOVANI, E.; YETANO, A.; ORELLI, R.L. Municipal Performance Measurement in Practice: Which Factors Matter? *Public Administration Quarterly*, v. 34, n. 4, p. 591-635, 2010.

RANDMA-LIIV, T.; KICKERT, Walter. The Impact of the Fiscal Crisis on Public Administration Reforms: Comparison of 14 European Countries. *Journal of Comparative Policy Analysis: Research and Practice*, v. 19, n. 2, p. 155-172, 2017.

REICHBORN-KJENNERUD, K.; JOHNSEN, A. Performance Audits and Supreme Audit Institutions' Impact on Public Administration The Case of the Office of the Auditor General in Norway. *Administration & Society*, v.50, n. 10, p. 1422-1446, 2015.

ROMZEK, B.S.; DUBNICK, M.J. Accountability in the Public Sector: Lessons from the Challenger Tragedy. *Public Administration Review*, v.47, n.3, p. 227-238. 1987.

WANG, XiaoHu. Assessing Performance Measurement Impact: A study of US local governments. *Public Performance & Management Review*, v. 26, n. 1, p. 26-43, 2016.

7. Appendices

APPENDIX A - question script for the interviews in the municipalities

PART ONE: SECRETARY'S OWN TOOLS AND METHODS OF ADMEASUREMENT

1. Is there any kind of tool that measures performance or public services indicators in this secretary?

If so

2. Which tools are these?

3. Can you give some examples of the use of these tools?

4. Since when these tools exists?

5. How do these tools work?

6. Who built it?

7. What is the frequency of use?

8. How is the data collected and updated? Who is responsible for doing this?

9. What is your assessment of the tool?

10. Is there any meeting to evaluate the indicators results?

PART TWO: the presence of i-Educ in the secretary of education

1. Is there any secretary employee who is responsible for dealing with this question?

2. How the school data collection is done?

3. Is there any employee responsible for data consolidation? How this consolidation is done? How often does this consolidation occur?
4. What is the structure and the dynamic used by the secretary to set the data? The secretary possess any kind of structure of its own or the mayor's cabinet is the responsible for this articulation?
5. Besides i-Educ, is there any other state or federal government demand to feed systems with data?
6. When the staff finishes answering the questionnaire, these informations are sent to the mayor's cabinet or directly to the Court of Accounts?
7. What are the main difficulties to obtain the necessary data for the filling of the i-Educ reports?

PART THREE: i-Educ influence and impact

1. The receipt of the i-Educ results presents any kind of novelty or new inputs?
2. How good is i-Educ to evaluate the performance of this secretary of education?
3. Do you miss having more indicators that are specific to this municipal reality? Would have any improvement suggestions?
4. The secretary uses i-Educ results to improve internal management and services provision? It is really a support for decision making?