

Managing Performance in the Defense Sector: Cases of the Italian Army and US Navy Surface Warfare Enterprise

Armando SUPPA^a, and Natalie J. WEBB^b

^a *Business Administration Department, University of Verona*

^b *Defense Resources Management Institute, Naval Postgraduate School*

Abstract. Over the last twenty years, new public management (NPM) has inspired managerial reforms in public sectors worldwide [1]. The US and Italian governments have embraced one of the main tenets of NPM, managing for results. In this study, we assess the introduction of performance management practice in hierarchical and complex public organizations in both countries, in particular how and to what extent military organizations, the Italian Army (IA) and the US Navy Surface Warfare Enterprise (SWE), implemented performance based management systems (PBMS). We use the first “form” of the framework proposed by Bouckaert and Halligan [2] to compare the two cases with respect to the measurement, integration and use of performance information. We find that both organizations have encountered “benefits”. We contribute to the literature and to the practical use in government organizations by demonstrating some key features or characteristics of hierarchical, complex government organizations that enable or detract from the successful implementation of a PBMS.

Keywords. Results-based management, Performance based management system, Performance measurement, Performance information

Introduction

In the last 20 years, public managers worldwide and corresponding public reforms of OECD countries focused principally on performance management and performance evaluation [1]. Government leaders have shown increasing interest in the measurement of performance to obtain better results [3] in terms of output and outcome. The paradigm of new public management (NPM) inspired this wave of reform [4]. This reform “paradigm” sees in the use of performance measurement and performance management (PMM) one of the means to increase the efficiency and effectiveness of public organizations, under the slogan “value for money” [5]. NPM also motivated public sector reforms in the U.S. (the Government Results and Performance Act, GPRA) and in Italy (Cassese, Bassanini, Brunetta Reforms) [6-7]. Cepiku and Meneguzzo [7] note that the main similarity between the Italian and US approaches of NPM is performance orientation, since the 80’ for US and later 90’ for Italy. Hence, public administrations, including military organizations, developed and introduced PMM. Robinson [8; p. xxxvi] defines performance management as the broad and systematic use of formal information to improve public performance. The above definition underlies the importance of two aspects of performance-based management: performance information availability and the use of performance information in decision making processes. Several factors shown in the literature influenced

components, contextual factors [9], technical factors, other factors (see Fryer et al for a review) [10]. In this study we first compare the performance based management systems (PBMSs) implemented by the Italian Army and US Navy Surface Warfare Enterprise. For this comparison we employ the first form, specification of components of PBMS (measurement, incorporation and use) presented in the Bouckaert and Halligan's framework of comparative analysis [2]. Second, we identify key factors that enable or detract from effective use of a PBMS in both cases. We find that not all factors identified in previous research as necessary determinants of a successful PBMS are significant in the present study. Our findings confirm the relevance of knowledge and training [11], implementation of a new integrated information system [12-13], and sub optimizing behavior [14]. We also find that in both cases, contingent aspects related to organizational culture affect the use of a PBMS. Yet the public management literature scarcely features PBMS in military organizations. Thus, our contribution fills a gap in the existing literature. The research questions that we propose are:

How do the Italian Army and US Navy Surface Warfare Enterprise manage¹ their performance?

What key features of hierarchical, complex government organizations enable or detract from the successful use of a PBMS?

In the first section, we present the theoretical background and the framework for comparing the two cases. In the second section, we discuss our research method, and in the third section, we analyze the cases. Finally, we provide conclusions in the last section.

1. Theoretical Background

Over the past two decades, the resurgence and new approaches to performance management illustrate one of the most "widespread international trends" in public management [15], even if not so new in theory or practice [6].

Results-based management provided the basis for numerous public sector reform initiatives worldwide, stimulating increases in efficiency, effectiveness and quality of public services [16]. Kettl [17] states that the main purpose of managing for results is to improve results and not to produce measures; thus, it is "far more useful" to see this process as performance-based management (PBM) instead of performance measurement. As Bouckaert and Van Dooren [18; p. 151] suggest, performance measurement is the act of measuring, while performance management is the reaction to performance information: "performance management is both measurement and management, [it is] about information and action". Moynihan says [19; p. 78] the principal idea of PBM is "using performance information to increase performance by holding managers accountable for clearly specified goals". It follows that public administrations should produce performance information and use this information to inform decision making, increasing organizational performance [20]. Bouckaert and Halligan [2] offer a slightly different definition, stating that performance management is a management model that incorporates and uses performance information for decision making. They clarify PBM with regard to the measurement, incorporation and use of performance information by stating:

¹ The SWE has started a process changing since the end of 2011.

- Performance measurement is a bundle of deliberate activities of quantifying performance that results in the “producing” of performance information. Performance measurement is a process in five steps that targets measurement effort, specification of metrics, data collection, analysis and data reporting [6, 10];
- Incorporation is the process of importing performance data into documents and procedures with the intention of using them. The purpose is to incorporate the performance information in the memory and culture of the organization and finally integrate performance information into the policy and management cycles;
- Uses of performance information in decision making for planning, resource allocation, taking corrective action and rewarding. Different managers and stakeholders require information for different uses [21]. Van Dooren et al. [6] proposed three different uses: learning, steering and control, and accountability.

In this study, we use this framework to analyze and compare the PBMS developed and introduced by the IA and the US Navy SWE. A variety of factors determine successful the implementation of PBMS [22]. In our study regarding the factors that were significant in both cases, we find as did De Lancer and Holzer [11] the positive effect of knowledge and training on the implementation of PBMS. Brignall and Ballantine among others [12-13] argue that effective IT/IS is vital to the success of PBMS, while Smith [14] observes that sub-optimization occurs when an optimal condition at the unit level leads to a sub-optimal situation at higher level. This results due to lack of coordination and integration among different organizational functions or areas. In our research we find that some aspects of the “organizational culture” influence the use of PBMS. The IA and the SWE are hierarchical functional organizations, based on the strict observation of rules and procedures. Although they shifted their focus to outputs, one of their main concerns remained the control and “maximization” of appropriations, which is a “feature” of the traditional Weberian bureaucracy [19, 23]. Pollitt [24] observes that the bureaucracy tends to maintain their “memory” through a range of “storage locations” among these he cites the experiences and knowledge of existing staff and the norms and values of the organizational culture [15]. Furthermore, in another study, Pollitt [15; p 29] affirms that “the way major institutions were set up and infused with particular norms casts a long shadow down the years...even when short term instrumental rationality indicates that change would be advantageous”. These institutional norms have an influence on decision making, long after the end of the original reasons that made them [15].

2. Method

To validate our research question and verify our theoretical framework, we use the case study method [25], in particular the multiple case study [26], to analyze observable events and facts in their natural conditions [27]. We selected the IA and SWE cases for two reasons. First, the two organizations provide classical examples of hierarchical, multi-layered organizations exhibiting multiple objectives of many public organizations. Second, we had the access to data and key knowledgeable personnel. With little comparative research on PBMS in military organizations, we have valuable resources with which to begin to fill this gap.

2.1 Case Study Italian Army

We chose the Italian Army for the case study because we have access to the case study information [28]. We use different data collection techniques and sources of evidence to provide information with which to compare the IA to the SWE. We rely on a series of semi-structured interviews using the procedures described by Yin [26], along with internal documents, direct observations and archival records. Using this research material, one author analyzed and inductively coded the evidence to provide the basis of this study [29].

2.2 Case Study US Navy Surface Warfare Enterprise

We also chose the SWE because one author participated in the Webb and Candreva study [30] and analyzed the SWE reporting and use of performance information. Webb and Candreva employed a case study research design to investigate the SWE's activities and decisions. Research material included briefings, notes of briefings, internal documents, meeting minutes and other archival information on the SWE Intranet. The authors conducted approximately 25 hours of interviews with members of the enterprise, representing cross-functional (personnel, maintenance, etc.) teams, class (type of ship) squadrons, and contractor support.

3. Case Presentations

3.1 Organizational environment

The Italian Army, composed of about 400 sub-organizations, functions as a hierarchy for direction and coordination. Sub-organizations perform heterogeneous activities, use different resources and competences, and pursue specific goals [31]. The IA's mission, "to provide the generation and preparation of a land force component with adequate readiness given the available resources for the homeland security and the turnover in international military operations" (OBS 213) [32], results in an output measured by proxies for military readiness (percentage of "ideal" readiness), task force generation (percentage related to a standard) and expenditures (percentage of financial resources allocated versus expended).

Recent budget cuts to the IA resulted in acute difficulties in meeting the IA mission². The IA appropriation for operational expenditures declined by 70% in the last nine years from € 1,028 in 2004 to € 310 million in 2012, not adjusted for inflation [32-33]. The Surface Warfare Enterprise, an organization within the US Navy, commanded by a three-star admiral, supports the 162 surface ships of the U.S. Pacific and Atlantic Fleets. In 2008, SWE personnel managed approximately \$5.2 billion in annual operation and maintenance funds for the readiness of the surface fleet. The SWE responsibilities include providing ready ships, and "optimizing" warfighting readiness of the Navy's surface fleet. The SWE mission, "warship ready for tasking" for multiple possible operational missions, requires SWE personnel to provide ready ships with a given performance measure [34]. Although the SWE did not experience significant budget cuts during the period of study, the US Navy leaders' expectation was to link the PBMS to budget decisions. Webb and Candreva [30; p. 525] report that "Navy leaders express[ed] their desire to drive the budgeting process." While both

² The Minister of Defense, Former Admiral Di Paola, announced that the number of military and civilian personnel will be cut by 20% and almost a third of military bases will be closed or sold over the next five years <http://www.bbc.co.uk/news/business-17042065> 02/15/2012.

IA and SWE function as traditional military-hierarchical organizations, with missions to provide outputs (not outcomes, such as battles won or situations resolved, at least in peace [35]), their leaders face different issues. In the IA case, leaders, focused on internal and external accountability, by attempting to properly communicate performance and increase the organization's chance of survival, while SWE leaders concerned themselves with increasing efficiency, particularly technical efficiency, to free up resources [6].

3.2. Analysis and comparison

Framework of analysis

In this study, we analyze the two cases using one of Bouckaert and Halligan's forms of analysis[2] from their performance management framework. Specifically we use the first form, "specification of the components" of performance management (measurement, incorporation and use). We use only the first form as our research aim is limited only to the analysis of key factors that enable or hinder the implementation of a PBMS, while Bouckaert and Halligan's study focused mainly on comparisons of real country cases to obtain the "dominant performance model of central government". In the following section we describe first how each organization carries out measurement, incorporation and use of the performance information, followed by a comparison of the two organizations.

Performance Measurement

In this section, we present analyses of the main activities of performance measurement in both organizations. We target the measurement efforts, including how managers selected performance indicators, collected data, interpreted the results and reported the performance information [6, 10]. We also assess the quality of the processes and the quality of the process of both organizations, looking at validity (capacity to logically represent the construct measured), reliability (repeatability of measurement), and accuracy (the capacity to measure the actual value) [36].

Italian Army

To begin to manage performance, Italian Army personnel mapped the organization's main internal processes and activities related to operational expenses³. IA managers determined outputs for each activity, using the Goal Question Metric approach [37], selected applicable metrics (performance targets based on outputs; expenses for output units; amounts of outputs provided and impairment thresholds) and indicators, which are combinations of metrics (for details, see Sarcia 2010 [38]). Thus, using the outputs and measures defined then integrating these data into the strategic and financial planning processes, the IA created an output-based budget based on historical data. This budget links IA strategic objectives, operational objectives and operational programs to financial figures. Additionally, IA personnel calculated a composite measure of military readiness for the entire organization by aggregating the percentage achievement of different outputs, weighted by an "impact factor" (the average of five years of appropriations for a particular output as a percentage of five years of total IA

³ IA managers mainly manage the operational expenses of the annual budget because personnel expenses depend on manning levels, which are set by law, and are thus a fixed cost, and investments are in the budget authority of the Ministry of Defense.

appropriations). IA managers can then examine the composite measure, which ranges from 0-100% and recalculate readiness using “what-if” analysis that allows managers to simulate alternative funding levels, providing the rate of performance (% military readiness) obtainable by the whole organization for each amount of allocated resources. This predictive model permits drilling-down into activities and their expenses to highlight ex ante those areas that may experience organizational failures in terms of low performance for a given resource allocation. In practice, the “what-if” analysis and measures of expected performance provide a sort of benchmark, identifying achievable targets for organizational functions or areas with a given level of funding. Data collection supports the measurement of military activities, expenditures and readiness. Data come from internal transactional data bases, especially an internal legacy IT system. Sub-organization managers provide objective data such as logistic and training outputs. Other data come from self assessment. Managers use the data to analyze how well sub-organizations provided their outputs. They compare output generated with the given budget to a performance target set based on historical output (five-year average expenses resulting in some amount of outputs, as noted in the preceding paragraph). Currently, IA managers report quarterly performance. The performance report usually does not exceed 100 pages and depicts graphs and measures depicted using traffic lights (red, brown, yellow, green and pea green). The IA is also currently engaged in developing and testing advanced IT solutions, SiAPS+. These solutions use open source applications based on the field of business intelligence. However, the system is not fully operational because some components need to be further tested before being deployed. Whenever this system will be fully operative, it will supply a powerful tool providing essential and up to date information to all organization levels, through internet access.

In the implementation of IA performance measuring tools, managers observed several quality issues including problems related to accuracy in data availability and collection. This results because managers have data on operational expenses for each output rather than actual costs (to include personnel, investment, and other expenses not part of operations expenses), thus the unit of measurement is unit expense, not unit cost. As is the case with many government organizations, the internal information system uses a cash-based accounting system rather than a system based on costs. In addition to problems of accuracy and systematic errors, leaders suspect two other critical factors of performance measurement related to two dysfunctional behaviors [14, 39]. The first behavior, sub-optimization, results when individuals and sub-organizations optimize performance in their own parts of the organization, but do not properly integrate processes among different functions. Interviews revealed that during the year, decisions to allocate resources among functions or sub-organizations may have reflected priorities for a subordinate part of the IA, and sometimes not related to the whole organization due to the lack of an integrated system for financial information. The second behavior, gaming, results when self-assessed data result in distortion or manipulation of reported outputs. As Hood [40; p. 516] suggests, this mismanagement of information can result in “ hitting the target and missing the point”

US Navy Surface Warfare Enterprise

As in the case of the IA, SWE leaders outlined main internal processes and activities related to operations. Using these processes and activities, personnel designed five composite performance measures based on five critical performance algorithms or

“figures of merit (FOM).” One composite measure describes mission readiness relative to each of personnel, equipment (maintenance), supplies, training, and ordnance (or the acronym “PESTO”). One senior officer oversees each of the PESTO areas across ship classes (frigate, destroyer, cruiser and amphibious), and one product line manager oversees each of the ship classes. Each class of ships has unique systems, requirements and capabilities, thus product line managers prepare individual ships according to the ship’s technology and expected mission requirements. To meet the Navy’s goal to project power anytime, anywhere, ships must be ready to function independently and interdependently, complemented by advanced technological reach from other assets. SWE personnel first evaluate ships for mission readiness independently, providing a FOM or composite measure that serves as a proxy for the output, “readiness.” Combatant commanders (at some point) evaluate an individual ship within the group of assets with which it deploys. Navy leaders have an inherent belief that a properly trained and assessed individual ship will be capable of successfully integrating with others for all possible missions.

PESTO algorithms attempt to capture the relationships among the inputs, activities or processes, and outputs in a way that simplifies performance measurement. SWE personnel use composite proxy measures, standardized along a 0-100 scale, to indicate readiness. Using a stoplight chart where colors reflect scores (green=90-100; blue=80-90; yellow=70-80; and red<70), each PESTO indicator provides information on how well a ship can perform a certain type of mission relative to the five PESTO functions. For example, a ship listed as red with respect to equipment cannot perform the mission due to a low maintenance score. Importantly, SWE personnel negotiated and agreed upon each underlying algorithm that provides output (readiness) measures. The maintenance performance indicator, for example, comes from an algorithm that assigns values to repair tasks weighted according to their impact on mission accomplishment. Internal databases provide the data that feed the PESTO algorithms. In particular, financial figures come from legacy IT systems. SWE personnel attempt to use “cost⁴” data to analyze efficiencies. Five analyses exist, each portraying different aspects of “cost⁵”. The first analysis mines data to examine purchases and to assess whether the quantity or price paid for those purchases can be reduced. The second compares the spending on ships of the same class based on (ship) homeport. The third analysis charts cumulative year-to-date spending against rolling averages of performance. The fourth uses the stoplight-coding schema for readiness indicators and attempts to compute the cost to move a ship from one (stoplight) status to the next. The fifth analyzes spending on particular missions in an attempt to understand or manage the “cost⁵” of those missions.

SWE personnel measure and report readiness consistent with the overarching Defense Readiness Reporting System (DRRS), a defense-wide IT system for reporting military unit readiness for a given mission. The DRRS is very close to an internal BI solution, which focuses on military readiness rather than the “bottom line”. Performance measures change daily through the recording of maintenance actions, personnel changes, training events, etc. The DRRS captures activities that change the readiness measures in a way that anyone with access to the database can drill down to find reports on individual action items by ship. As in the case of the IA, SWE performance measures exhibit quality problems due to data inaccuracy, and systematic errors and biases. As of 2008, the SWE had not made the leap from its longstanding

⁴ Expenses, obligations not cost.

encumbrance-based budgeting and accounting systems to systems of cost accounting that could provide performance-based cost information. Similarly, SWE personnel note that sub-optimization affects performance measurement. Rather than being solely a behavioral issue, SWE personnel recognize that sub-optimization may be driven by the use of composite measures. They recognize that managers cannot aggregate the PESTO performance indicators to calculate a single measure of “warships ready for tasking.” SWE leaders know they cannot aggregate scores calculated from different composite algorithms. Thus, managers can optimize resource allocation within a functional area but cannot calculate an optimal mix of resources for the organization as a whole. As in many cases, SWE performance management requires subjective use of available performance information. Finally SWE managers noted that several PESTO measures added nothing to the understanding of performance. Van Dooren et al. [6] describe this as a problem of performance measurement dysfunction, information overloading or indicator mushrooming.

IA and SWE performance measurement comparison

The two cases, even with contextual differences, share some commonalities. The organizational “culture” and overarching concerns about appropriations result in the use of obligations and expenditures as proxies for costs in both organizations. These issues are twofold. The first is technical. The introduction of a new information system that can support newly-recognized information needs is crucial for effective performance measurement. As Kaplan and Norton [12] note, an unresponsive information system can be the Achilles’ heel of a performance measurement system. Information technology plays a critical role in making performance measurement possible [41]; thus implementing an enterprise resource planning (ERP) and business intelligence (BI) solution gives a more realistic picture of the organization. The second aspect is related to the culture of an organization, its internal knowledge. In both the IA and SWE managers lack some performance measures, especially cost-related measures; their knowledge and focus is concentrated mainly on annual appropriations and obligations. As a consequence sub-optimal behavioral dysfunction arises because managers cannot properly integrate across functional areas and sub-organizations’ performance information, and, because they focus on expense data rather than costs, their views are biased.

Incorporation of Performance Information

Incorporation means to extend performance information by importing it into the management cycle (planning, monitoring, evaluation, feedback) with the intention of using the information. In other words if the performance related data are integrated in organizational processes and documents, they will ultimately be incorporated in its “culture and memory” [2]. We considered only the management cycle because the political and contractual cycles are outside the scope of our investigation. We mean the management cycle as ex ante planning and programming, ex nunc monitoring, ex post evaluation and finally feedback [6].

Italian Army

We observed that the IA leaders introduced the performance information (average expense per output and the target output for 100% of military readiness, etc.) in the planning process of the output-based budget. Additionally a report from each IA

function fed this activity indicating in detail three possible resource allocation levels with corresponding output levels and issues (risks or pitfalls). Furthermore, “what if analysis” supported this process predicting the possible areas where lack of resources could cause organizational failures (e.g., military readiness would fall short due to an established impairment threshold⁵). The quarterly performance report provided performance data on the monitoring process using graphs and traffic lights. Additionally, if more resources became available during the year, e.g., for international peacekeeping operations, an ad hoc performance report and “what if analysis” could support in the decision making for resource allocation the Chief of Army General Staff. Performance-related data were introduced partially in the evaluation process. In fact the annual performance report and its performance information fed only organizational evaluation; these data did not support the evaluation of military personnel. Finally the performance information of previous years were used in the procedures and processes for reviewing the PBMS for the next year, principally for updating output unit expenses and changing metrics and/or indicators.

Surface Warfare Enterprise

We noted that the PESTO performance information supported the four processes. In particular the DRRS and PESTO five pillar data fed the monitoring process almost in real time. SWE leaders did not include performance-related data on individuals due to an existing personnel evaluation system and mandatory procedures.

IA and SWE comparison incorporation of performance information

We observed that both the IA and the SWE introduced performance information in all four components of the management cycle. However, we noted different degrees of incorporation in the two organizations. The IA use performance information in a more “institutionalized” way in the planning phase than the SWE. In fact different performance-related data fed this process in the IA (unit expense for output, target for 100% military readiness, target for allotted resource, what if analysis, possible pitfalls for lack of resource). The SWE used performance-related data in a more integrated use of performance related data in the monitoring process. The PESTO measures provided timely performance information on current activities. These findings are consistent with the different informational uses of both organizations, accountability for the IA, and steering and control for the SWE. Finally we observed that in both cases the processes of evaluation of military personnel were not well integrated with organizational performance information because both organizations have other systems that measure performance of individuals under mandatory service requirements.

Use of performance information

In this section we analyzed why and how both organizations use performance information in different components of the management cycle. To answer the “why” question, we considered three possible uses of performance information: learning; steering and control; and accountability. These do not exclude each other, but one use

⁵ Organization managers define an impairment threshold as a level of financial resources under which the whole organization or a function provides insufficient or no outputs or services.

tends to prevail over the others. Indeed, to answer the “how” question, we studied the use of performance information in the different components of the cycle.

Italian Army

The main use of performance information for the IA was for external and internal accountability reasons. Due to the global economic crisis and the consequent fiscal crisis, the IA budget suffered a severe cut: “everyone must cut back by... %”[42]. Hence, to have more resources, the Ministry of Economy and Finance requested more transparency and required showing outputs against resources. This, combined with a new enacted law (150/2009) prescribing introduction of the performance management cycle in all public organizations, provided the stimuli for the implementation of IA performance-based management. In fact almost all of the officers interviewed pointed this crucial moment for survival of the organization. Moreover the commanders of the operative units, being accountable for the full “readiness” of their organizations, saw in the system a way to report their accountability.

The IA Army principally used performance information for planning the annual output-based budget. During the ex ante phase the performance-related data intensively supported and were used for resource allocation. Performance information only slightly influenced decisions about resource allocation and corrective action during the year. At the end of the year the organizational evaluation resulting from performance data mainly served for external accountability. However, the evaluation process of military personnel used another set of measures and procedures, while the evaluation of civilian managers is based on them. Finally performance information fed the feedback process. Interviewees affirmed that after the annual report and before budget planning, they updated the performance system, i.e. updating a unit expense of an output or developing new metrics and indicators. Furthermore they observed that officers in charge of different functions had an active part in this phase and usually suggested new developments in the system.

IA managers intensively used performance information for planning of the output-based budget. This is consistent with the organizational purpose of communicating IA output performance (external accountability). The slight use of performance information during the year for taking corrective actions was principally due to lack of an integrated information system that would allow obtaining an optimum resource allocation (or reallocation) across organizational functions.

Surface Warfare Enterprise

The SWE managers used performance information primarily for steering and control or internal processes, increasing internal (technical) efficiency. In fact, SWE leaders faced a challenge of continuously improve technical efficiency in the core areas of maintenance, logistics, manning and training to create budget slack so the Navy could increase its purchase of ships, ammunition, and fuel. We noted that SWE leaders attempted to create a performance budget using the information supplied by PESTO. In practice, however, PESTO only served to direct corrective actions and resource allocations based on PESTO information. At every operational level within the SWE (e.g., maintenance, crew training, etc.), PESTO provided performance information to assess readiness of a ship in one of the five PESTO dimensions. Overall organizational evaluation was related to this. Finally expense analysis employed the

information data for resource allocation. We observed that the SWE principally used performance information for taking corrective actions during the monitoring phase. This is consistent with the main steering and control use performance information by SWE managers.

IA and SWE comparison of performance information use

We found that both organizations use performance information for evaluation only in part. Personnel evaluation occurs through different processes and systems due to legal and other mandatory reasons; thus no personal rewards or incentives exist within the PBMSs. Similarly, the evaluation systems of the organizations have no provision for rewarding organizations for meeting performance targets (e.g., increasing funding, allowing carrying over of efficiency gains, providing bonuses, etc. [43; p. 174]). We noted that three main factors, somewhat related each other, hindered the functional use of performance information in both organizations. The first was a technical factor, the absence of an integrated financial information system, “where... [the] accounting system shift[ed] to [a] cost accounting system...in a coherent a systematic way”[2; p. 112]. In fact both organizations did not have a holistic system (i.e. enterprise resource planning system integrated + a business intelligence system) that permitted them to identify the priorities across different organizational functions and to calculate costs (consumption of inputs) of the output produced rather than calculating expenses (purchase of inputs). The second was a lack of knowledge and general understanding of on performance management [11]. We noted that obligations (promises to pay) were normally used as proxies of cost. Costs of personnel (which were and are the largest costs for both organizations) were not included due to the inability to integrate personnel costs and systems into the performance management system.. The third was related to “organizational culture”. Both organizations tried to control and increase their budget appropriations using performance information. The IA uses the PBMS for external accountability, and the SWE tried to “link” the PBMS to budget decisions. This focus on appropriations is rooted in the “memory” of the traditional bureaucracy [19, 23].

These three factors relate to each other. The focus on appropriations has as a consequence the focus on obligations; as Webb and Candreva [30; p. 545] note about the SWE case, “[i]nstitutional norms and overarching concerns about preserving, consuming, and expanding appropriations result in the use of inappropriate proxies for cost”. Thus it is logical that if appropriations and obligations are important there is no need for a cost-based accounting system. Finally, these factors resulted in organizational sub optimization and a dysfunctional allocation of resources because decision making relied on inaccurate performance information.

Conclusion

The IA and SE made strong commitments to introduce PBMSs, which they accepted as a rational modern technique. However, to assess whether the systems are functional requires evaluating them against three factors:

- introduction of an integrated financial information system that tracks and records the costs of the outputs;
- improvement of technical knowledge and training on performance management in the whole organization to increase the accuracy of PBMS; and

- changing organizational “memory” that still “remembers” decision making based on inputs, even if managers or leaders make decisions based on outputs;

The present research provides a starting point for understanding the factors that managers of large, hierarchical government organizations should assess in their attempts to manage performance. The main limitations of this study are two. First we did not consider the institutional aspects of both organizations, specifically national culture [44], national systems, [15] policy cycles of PBM [2], and national budgeting processes. Second, both organizations are still involved in reviewing and developing their PBMS. Future research could examine similar cases of the implementation of a PBMS by public administrations, based on the factors this study provided on performance management integration in military-bureaucratic organizations.

References

1. OECD, *Modernising Government*. 2005, Paris: PUMA OECD
2. Bouckaert, G. and J. Halligan, *Managing Performance. International Comparisons*. 2008, London: Routledge.
3. Micheli, P. and A. Neely, *Performance Measurement in the Public Sector in England: Searching for the Golden Thread*. *Public Administration Review*, 2010. **70**(4): p. 591-600.
4. Hood, C., *A PUBLIC MANAGEMENT FOR ALL SEASONS?* *Public Administration*, 1991. **69**(1): p. 3-19.
5. Osborne, D. and T. Gaebler, *Reinventing government*. 1992, Reading, MA: Addison-Wesley.
6. Van Dooren, W., G. Bouckaert, and J. Halligan, *Performance Management in the Public Sector*. 2010, New York: Routledge.
7. Cepiku, D. and M. Meneguzzo, *Public Administration Reform in Italy: A Shopping-basket Approach to the New Public Management or the New Weberianism?* *International Journal of Public Administration*, 2011. **34**(1-2): p. 19-25.
8. Robinson, M., *Performance Budgeting: Linking Funding and Results*. 2007, New York: International Monetary Fund.
9. Lüder, K.G., *The 'contingency model' reconsidered: Experiences from Italy, Japan and Spain*, in *Perspectives on Performance Measurement and Public Sector Accounting*, E. Buschor and K. Schedler, Editors. 1994, Haupt: Bern.
10. Fryer, K., J. Antony, and S. Ogden, *Performance management in the public sector*. *International Journal of Public Sector Management*, 2009. **22**(6): p. 478 - 498.
11. de Lancer Julnes, P. and M. Holzer, *Promoting the Utilization of Performance Measures in Public Organizations: An Empirical Study of Factors Affecting Adoption and Implementation*. *Public Administration Review*, 2001. **61**(6): p. 693-708.
12. Kaplan, R.S., . and D.P. Norton, *The Balanced Scorecard - Measures That Drive Performance*. *Harvard Business Review*, 1992. **70**: p. 71-79.
13. Brignall, S. and J. Ballantine, *Performance measurement in service businesses revisited*. *International Journal of Service Industry Management*, 1996. **7**(1): p. 6 - 31.
14. Smith, P., *On the unintended consequences of publishing performance data in the public sector*. *International Journal of Public Administration*, 1995. **18**(2-3): p. 277-310.
15. Pollitt, C., *Performance Management in Practice: A Comparative Study of Executive Agencies*. *Journal of Public Administration Research and Theory*, 2006. **16**(1): p. 25-44.
16. OECD, *Distributed Public Governance: Agencies, Authorities and other Government Bodies*. 2002, Paris: PUMA OECD.
17. Kettl, D.F., *The global revolution in public management: Driving themes, missing links*. *Journal of Policy Analysis and Management*, 1997. **16**(3): p. 446-462.
18. Bouckaert, G. and W. Van Dooren, *Performance Measurement and Management in Public Sector Organizations*, in *Public Management and Governance, 2nd ed.*, T. Bovaiard and E. Löffler, Editors. 2009, Routledge New York. p. 151-164.
19. Moynihan, D.P., *Managing for Results in State Government: Evaluating a Decade of Reform*. *Public Administration Review*, 2006. **66**(1): p. 77-89.
20. Moynihan, D.P., *The Dynamics of Performance Management: Constructing Information and Reform*. 2008, Washington D.C.: Georgetown University Press.

21. Behn, R.D., *Why Measure Performance? Different Purposes Require Different Measures*. Public Administration Review, 2003. **63**(5): p. 586-606.
22. Chenhall, R. and K. Langfield-Smith, *Factors influencing the role of management accounting in the development of performance measures within organizational change programs*. Management Accounting Research, 1998. **9**(4): p. 361-386.
23. Rainey, H.G., R.W. Backoff, and C.H. Levine, *Comparing Public and Private Organizations*. Public Administration Review, 1976. **36**(2): p. 233-244.
24. Pollitt, C., *BUREAUCRACIES REMEMBER, POST-BUREAUCRATIC ORGANIZATIONS FORGET?* Public Administration, 2009. **87**(2): p. 198-218.
25. Creswell, J.W., *Research design: Qualitative, quantitative, and mixed methods approaches (2nd ed.)*. 2007, Thousand Oaks, CA: Sage.
26. Yin, R.K., *Case study research: design and methods*. Fourth edition ed. Applied Social Research Methods, ed. L. Bickman and D.J. Rog. Vol. 5. 2009, Los Angeles: Sage.
27. Benbasat, I., *An Analysis of Research Methodologies*, in *In The Information Systems Research Challenge*, F. Warren, Editor. 1984, Harvard Business School Press: Boston.
28. Eisenhardt, K. and M. Graebner, *Theory Building From Cases: Opportunities and Challenges*. The Academy of Management Journal ARCHIVE, 2007. **50**(1): p. 25-32.
29. Miles, B.M. and A.M. Huberman, in *Qualitative Data Analysis: An Expanded Sourcebook B2 - Qualitative Data Analysis: An Expanded Sourcebook*. 1994, Sage Publications: Thousand Oaks, CA.
30. Webb, J.N. and P.J. Canderva, *Diagnosis Performance Management and Performance Budgeting Systems: a Case Study of the U.S. Navy*. Public Finance & Management, 2010. **10**(3): p. 524-555.
31. IAGS, *Army Report 2010*, in *Rivista Militare*. 2010, Italian Army General Staff: Rome.
32. MEF, *Nota Integrativa alla Legge di Bilancio per l'anno 2012 e per il triennio 2012-2014 del Ministero della Difesa*, MEF, Editor. 2012, MEF: Rome.
33. MoD, *Nota Aggiuntiva allo stato di previsione per la Difesa per l'anno 2004*, MOD, Editor. 2004, MoD: Rome.
34. Robinson, M., "Performance Budgeting Models and Mechanisms", in *M. Robinson (Ed.), Performance Budgeting: Linking Funding and Results 2007*, International Monetary Fund: New York.
35. Wilson, J.Q., *Bureaucracy - What Government Agencies Do and Why They Do It*. 1989, New York: Basic Book.
36. Bouckaert, G., *Measurement and Meaningful Management*. Public Productivity & Management Review, 1993. **17**(1): p. 31-43.
37. Basili, V., G. Caldiera, and H. Rombach, *The Goal Question Metric Approach*. Encyclopedia of Software Engineering, 1994. **2**: p. 528-532.
38. Sarcia, S.A., *Is GQM+Strategies really applicable as is to non-software development domains?*, in *Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*. 2010, ACM: Bolzano-Bozen, Italy. p. 1-4.
39. Pidd, M., *Perversity in public service performance measurement*. International Journal of Productivity and Performance Management, 2005. **54**(5/6): p. 482 - 493.
40. Hood, C., *Gaming in Targetworld: The Targets Approach to Managing British Public Services*. Public Administration Review, 2006. **66**(4): p. 515-521.
41. Eccles, R.G., *The Performance Measurement Manifesto*. Harvard Business Review, 1991. **69**(1): p. 131-137.
42. Pollitt, C. and G. Bouckaert, *Public Management Reform: A Comparative Analysis-New Public Management, Governance, and the Neo-Weberian State*. 3rd ed. 2011, Oxford New York: Oxford University press.
43. OECD, *OECD Public Management Reviews: Ireland*. 2008, Paris: OECD.
44. Hofstede, G., ed. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. 2nd ed. ed. 2001, Siege: Thousand Oaks, CA.