

Assessing Agency-Level Performance Evaluation Reform in China: Can It Truly Serve as a Management Innovation?

Lin Ye and Xing Ni

Abstract

As in many other countries, performance management has received wide attention in China in order to increase government efficiency and accountability. However, agency-level performance management reforms have been largely overlooked in the existing literature with regard to the process of developing performance measures, stakeholder involvement, and the assessment of these measures. By investigating a provincial departmental case, this article reveals that the latest development of China's performance management reform has become increasingly sophisticated, with consideration for both political legitimacy and management efficiency. However, significant challenges remain before such efforts can be fully implemented as management innovations to truly improve government performance and accountability.

Introduction

Performance management has been considered a mechanism to improve government functions in a variety of ways, including holding bureaucrats more accountable, improving internal management efficiency, enhancing government effectiveness, and increasing government's responsiveness to citizens (Behn, 2003; de Lancer Julnes, 2009; Moynihan, 2008; Schmidle, 2011). As in many other countries, performance management has received wide attention in China. It has been argued that given China's political and administrative system, its performance management reform is utilized as an instrument for building state capacity, improving government accountability, and enforcing top-down inter-governmental political control (Gao, 2009; Burns & Zhou, 2010). Most existing literature studying this subject focuses on general-purpose government practices in China (see, for example, Chan & Gao, 2009; Gao, 2009, 2010; Walker & Wu, 2010). Moreover, most studies analyze written performance management policies and regulations already

in place, with inadequate attention paid to the process of designing performance measures and to the key stakeholders involved in the process.

This article argues that agency-level performance evaluation reform efforts need to be specifically examined in order to discern the political and technical considerations of performance management innovation in China. Three main research questions are: (1) what role does agency-level performance evaluation play in China's performance management reform?; (2) what factors can be identified as affecting the process of developing an agency-level performance evaluation system, in light of the stakeholders involved and their interactions?; (3) how can this process and the performance measures be assessed to discern the implications and drawbacks of the reform?

This article will first provide a brief overview of the existing literature on China's performance management reform. Next, the case of the Department of Ocean and Fisheries of Hainan Province (DOFHP)'s performance evaluation reform will be analyzed and its performance measures assessed. Findings are then discussed in the broader context of whether and how agency-level performance evaluation contributes to performance management reform in China. Finally, conclusions will be drawn based on the case study analysis and its theoretical implications.

Literature Review

Performance management has been considered an integral part of administrative reforms in China since the 1980s. Using Moore's (1995) terminology of managing upward, downward, and outward, Walker and Wu (2010) describe how performance management in China is used in all these directions: "upward from the bureaucracy to provide information to politicians, who in turn use it downward to hold officials to account; it is also used outward to permit citizens to question politicians and their servants" (p. 36S-37S). The structure of Chinese government means that "performance management is concerned not only with service performance but also with maintaining a stable, harmonious society and promoting economic and sustainable development" (Walker & Wu, 2010, p. 37S). The central government translates its macro reform goals into specific policy objectives and assigns the objectives downward through the administrative

hierarchy. These objectives are then converted to a variety of prioritized performance targets for local governments and officials to achieve.

One way to reinforce the political control from central government to local government is by implementing the so-called “target-based responsibility system” (Chan & Gao, 2008) or “objective responsibility system” (Burns & Zhou, 2010). Both terms refer to a system where the central government conveys strong administrative orders by establishing performance contracts with local governments and holds local officials accountable for the accomplishment of the established targets (*mu-biao ze-ren zhi*). The hierarchical nature of government means that performance management is “used chiefly to ensure that local officials comply with higher-level policy priorities” and that “by binding local officials’ target accomplishments to their career future the target-based responsibility system guarantees that local officials will follow the directives that come down from above” (Chan & Gao, 2008, p. 8).

Most existing studies focus on performance management and evaluation reforms in one locality, taking the general government as a whole or evaluating all agencies within a locality. There have been very few studies looking at agency-level performance reforms. Chan and Gao (2008) collectively examined performance contracts from organizations in different functional systems in Zhouzhi county in the inland Xi’an Municipality in Shaanxi province. But it was not shown how these performance measures were initiated, designed, implemented, and monitored by individual agencies and what actors were involved in the process of developing these measures. Dai and Teng (2008) assessed the performance evaluation system of Hangzhou, with performance indicators established and implemented by the municipal government. Liu (1995) investigated the performance evaluation system of state-owned enterprises but the research objective did not cover government agencies.

To fully understand the mechanisms of government management reform, the importance of the agency perspective on performance management has been widely discussed in the western literature (see, for example, Lu, 2007, 2008; Moynihan, 2008; Moynihan & Ingraham, 2003; Willoughby, 2004). When studying performance budgeting,

Lu (2007) argues that agencies “play an important part in the implementation of performance budgeting,” and “are also the place for examining managerial uses of performance measures” (p.3-4). In fact, “agencies are viewed as the center of most efforts to produce performance information in the budget process” (Lu, 2008, p.9). In a broader context Moynihan (2008) contends that “while there is promise in the future of performance management, its potential will best be achieved if governments rethink what it means, offer realistic expectations rather than hyperbole, and focus on the agency level factors that induce performance success rather than government-wide systems” (p. 189). Moynihan & Ingraham (2003) argue that “[t]he complexity of substantive policy areas, the breadth of government-wide responsibilities, the need for detailed performance information, and the advantages of agency ownership of goals point to the necessity of giving agencies freedom in developing goals and performance measures.” (p.483)

Similar principles are true for performance management reform in China. This article postulates that individual agencies’ organizational goals, strategic missions, and performance functions are different from general-purpose government and should be examined more carefully to assess China’s performance management reform.

Although Chinese governments have been considered as running like business firms (Oi, 1992; Walder, 1995), individual government agencies tend to face narrower performance targets and more identifiable service clientele than general-purpose governments. Therefore, agencies can tailor more technically advanced performance management systems to meet their needs. General-purpose governments have broader policy objectives and are accountable to the public, hence holistic performance evaluation methods like citizen satisfaction surveys are more often used (Burns & Zhou, 2010).

As the level of government that implements day-to-day policies and carries out technical projects, individual agencies tend to have more autonomy to initiate, design, and implement sophisticated performance management systems. Even under the “top-down direction” of senior government, individual agencies usually take on the detailed work of designing and implementing performance measures (Lu, 2008, p. 14-15). Because many employees in government agencies in China are civil servants, agency-level

performance management systems evaluate both organizational and individual performance. The individual performance evaluation strictly follows the process required by the Chinese Civil Servant Law and becomes an integral part of the performance management system.

Due to their more technical and less political nature, individual government agencies are more likely, and have more freedom, to engage external or international participants to design and implement performance measures. If their performance measures are in line with general social and economic development goals and if sufficient funds are available, individual agencies are increasingly likely to outsource their performance management functions (Ni, 2008; Ni & Yu, 2009). General-purpose governments in China act more like bureaucratic machines and rely less on engaging external partners in their organizational functions, including performance management.

Therefore, to fill the gap in the existing literature this article analyzes a typical case of a provincial level department. This department initiated, designed, and implemented a highly advanced performance management system with a diverse group of actors participating in the development process of this system. This case reveals the latest development of agency-level performance evaluation reform in China and helps answer the research questions. The findings in this case demonstrate the increasing importance of agency-level performance evaluation in China's performance management reform and how agency-level performance evaluation can serve as an innovative management practice, process, and technique. But significant issues remain. The next section explains the selection of the case.

Introduction to the Case

To answer the research questions, a case study approach is applied. Per Yin (2009), there are five major qualitative research methods: experiment, survey, archival analysis, history, and case study. Each method has its strengths and weakness depending upon the nature of the research questions. The case study method has the advantage of addressing "what, how and why" inquiries, which are appropriate for tackling the research questions raised in this study. Yin (2009) further argues that six sources of evidence, including

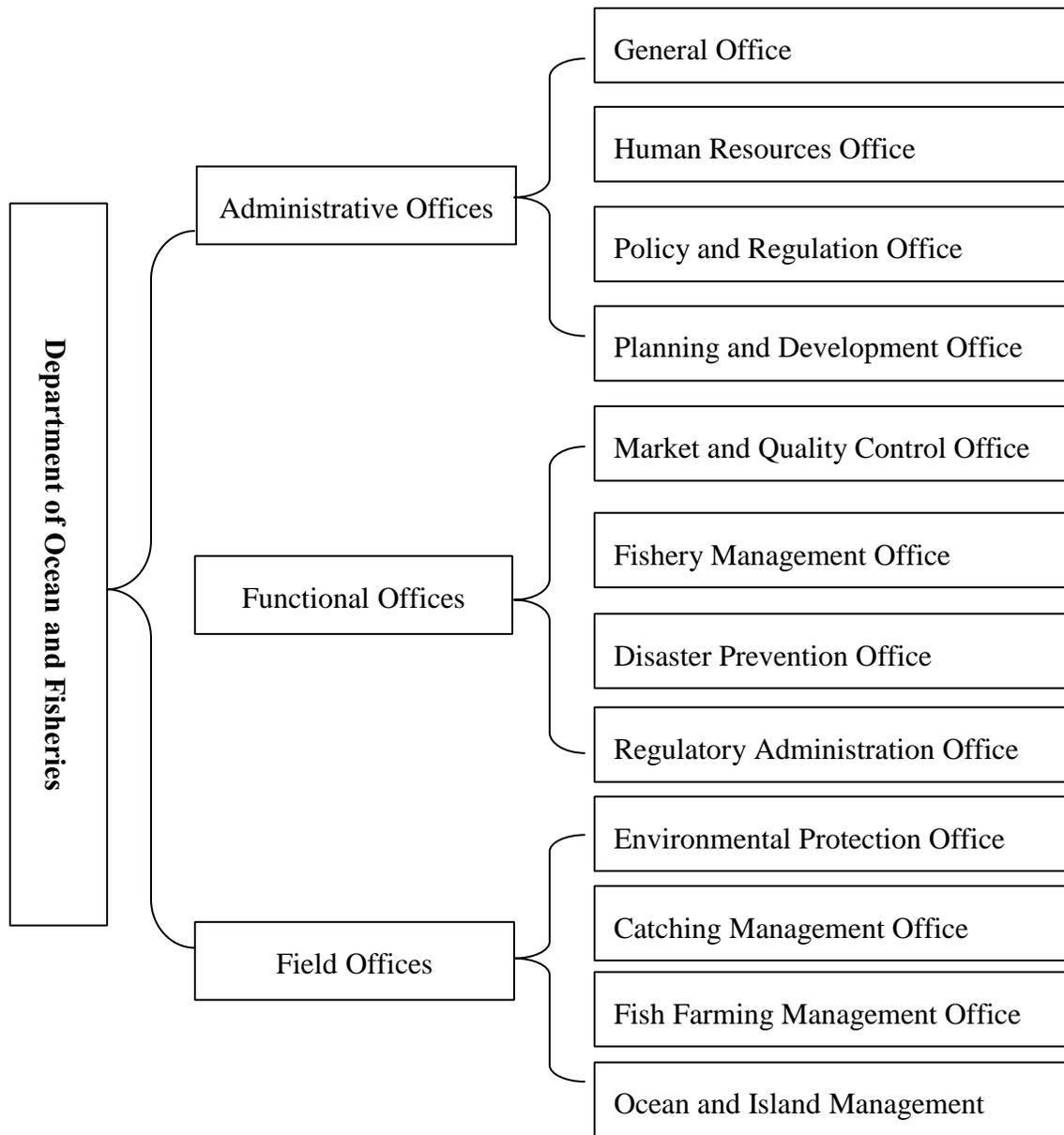
documentation, archival records, interviews, direct observation, participant observation, and physical artefacts, can be used as empirical evidence in case studies. This study primarily adopts documentary analysis, participant observation, and archival records as empirical evidence. Field work was conducted between 2010 and 2012 while the authors led the research team in developing a performance evaluation system for the Department of Ocean and Fisheries of Hainan Province (DOFHP). Evidence was collected from project documents, internal reports, and communications with participants in this project, including department heads, field officials, administrative staff, and external experts. The collected information was analyzed to demonstrate the process, implications, and issues of developing such a system.

The reasons this case study was chosen are threefold. First, the DOFHP case is a pilot program to design advanced agency-level performance evaluation systems in China. This project was chosen because it is one of the World Bank's Technical Assistance Project of the China Economic Reform Implementation Project (TCC5). As a highly selective program, the TCC5 fund aims to significantly increase the adoption of sensible reform and development strategies, policies, and plans in China. Selected projects are expected to promote the implementation of reform and development through institutional capacity building and by addressing specific challenges facing the government at the national and sub-national levels. The project has a strong focus on developing new knowledge as well as on replication and scaling up successful efforts (World Bank, 2006). The DOFHP project indicates the national significance of performance management innovation of government agencies.

Second, the DOFHP project served as a bottom-up, comprehensive revamping of department-level administration by adopting systematic performance management measures, with the integration of several supporting programs and the involvement of two top Chinese universities. Such an arrangement with multiple stakeholders involved in the management reform process has received wide attention in China (Ni, 2008; Ni & Yu, 2009).

Third, this case provides an opportunity to examine whether the adoption of technical performance measures can develop from exogenous experiment to endogenous enforcement, from constructing technical indicators to carrying out performance management policies. One of the logics of performance management Pollitt (2013) argued is the "logic of escalation," where the proliferation of performance information resulted from the combination of the political will and the agency's need to promote management innovation. Such logic is found in the UK and the Netherlands (Pollitt et al., 2010). Although this article covers the initial period of performance management reform in the DOFHP, it demonstrates the preliminary evidence of whether such logic of escalation applies.

Figure 1. Organizational Structure of Department of Ocean and Fisheries



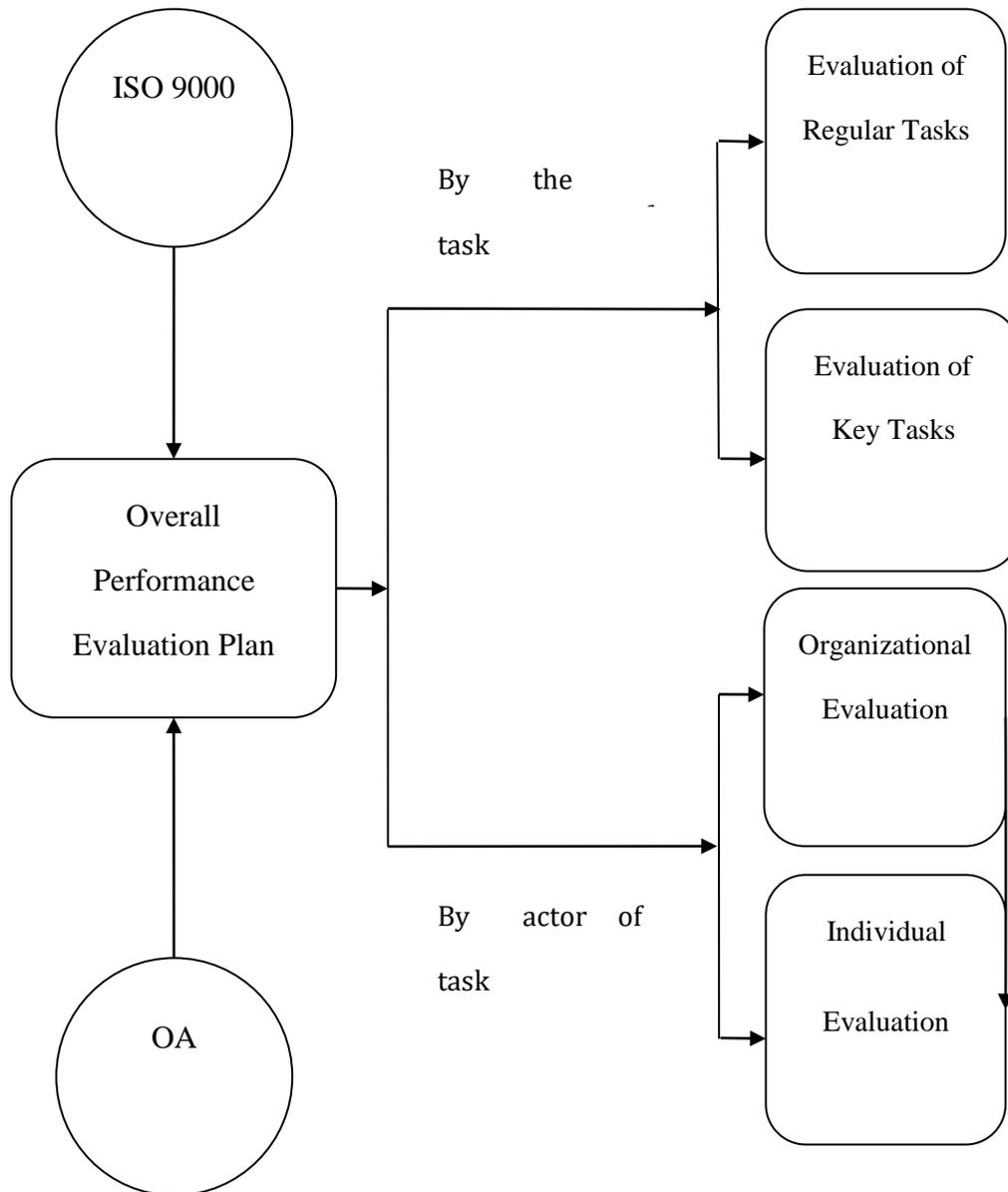
Source: Internal document of DOFHP

DOFHP is one of the provincial level agencies in Hainan Province and its structure and functions serve as a representative model for Chinese government agencies. The Province of Hainan had a population of 8.67 million in 2010, around 7% of which were employed in fishery-related industries. This is to say that DOFHP serves over 600,000 people in the province, with regard to their fishery-related activities, businesses, training, and administration. DOFHP is responsible for monitoring over 2 million square kilometres of water and provides services for fishery industry development in the South China Sea.

DOFHP ran a total budget of over RMB 25 million in 2012 and has 12 offices to perform its administrative, functional, and field duties. Figure 1 shows the organizational structure of DOFHP (Figure 1). DOFHP performs a wide variety of duties, ranging from internal affairs to field work, which makes it a typical Chinese government agency facing a difficult task of developing a comprehensive performance evaluation system. Performance management in DOFHP needs to encompass evaluation at both individual and organizational levels because many DOFHP employees are civil servants. The performance management system also needed to integrate the department's existing ISO 9000 and Office Administration (OA) systems (Department of Ocean and Fisheries of Hainan Province, 2010).

In early 2010, a formal bidding document was promulgated after DOFHP officials visited several other ocean and fishery agencies in China, including Beijing, Dalian, and Guangdong and learned about their performance management reforms. DOFHP intended to use the TCC5 funds to contract out the development of a performance management system. A World Bank Project Implementation Office was established to coordinate the implementation of this program. When DOFHP visited Guangdong province in early 2010, an initial contact was made between DOFHP and the School of Government at Sun Yat-sen University (SOG at SYSU). The school is an influential public administration research institution in China and is home to the Center for Chinese Public Administration Research (CCPAR), the only Ministry of Education Key Research Center in this discipline. A successful bid was submitted by SOG to DOFHP in August 2010. SOG's proposal was accepted and the contract was signed between DOFHP and the SOG project team in October 2010, starting a two-year project to establish a comprehensive performance management system for DOFHP. The project was completed in June 2012.

Figure 2. The Structure of DOFHP Evaluation Plan, ISO 9000, and Office Administration (OA) Systems



Source: Internal document of DOFHP

The final product of the SOG project was a “1+4” scheme performance evaluation plan for the DOFHP, which included an Overall Performance Evaluation Plan, with a Detailed Regulation on Performance Evaluation of Regular Tasks, a Detailed Regulation on Performance Evaluation of Key Tasks, a Detailed Regulation on Organizational Performance Evaluation, and a Detailed Regulation on Individual Performance Evaluation.

This scheme divided the overall evaluation design into two main categories with four detailed sections. Based on the nature of the tasks, the duties are categorized as key tasks and regular tasks. Based on the actor who performs the task, the duties are divided into organizational tasks and individual tasks. Figure 2 demonstrates the overall structure of the evaluation scheme, details of which will be assessed in the following section.

Assessment and Analysis

Developing effective performance measures and implementing these measures are key components of performance management reforms. The nature of performance measures reflects the quality and feasibility of the performance management systems. Many studies have tried to develop a framework for assessing performance measures. The essential criteria an effective performance measurement system should possess include validity, legitimacy, credibility, functionality, and accessibility (Bouckaert, 1993; Streib & Poister, 1999; Moynihan & Ingraham, 2003; Taylor, 2006). Similarly, Lu (2008) summarized four criteria of measurements: validity and reliability, usefulness, administrative feasibility, and political acceptance. Functionality and usefulness both refer to how effectively the performance measures can be used by organizations to achieve multiple objectives. Administratively feasible measures need to be accessible to both internal users and external parties.

Putting these measures together, the following section uses the criteria of validity and reliability, legitimacy and acceptance, credibility, functionality and usefulness, and accessibility to assess the performance measurements presented in the DOFHP case. The information employed here includes both primary source data from questionnaires distributed by the SOG project team and secondary source data from internal reports and third-party evaluations. The assessment traces the implementation of such measures and its political and administrative implications.

Validity and Reliability

The first criterion of an effective performance measurement system is validity and reliability. According to Bouckaert (1993), “validity has been the major, if not the only, concern in the development of measures and performance measurement systems” and

“a valid measure is one that is sound, cogent, convincing, and telling” (p. 31). A reliable measure needs to have consistency and stability when used under different conditions. Technical measures have been widely used to evaluate performance indicators’ validity and reliability. When studying performance-based budgeting (PBB) in American states, Lu, Willoughby, and Arnett (2009) find that state agencies are expected to submit a self-assessment to ensure the validity, reliability, and appropriateness of each performance indicator (p. 279).

In our case, during the program trial stage in January and February of 2012, all offices and agencies of DOFHP were required to conduct a self-assessment of the trial Detailed Regulation on Organizational Performance Evaluation and Detailed Regulation on Individual Performance Evaluation. Table 1 shows the test results of the trial evaluation. Per Nunnally (1978), a Cronbach’s alpha value above 0.7 is considered satisfactory and a value below 0.5 is unacceptable. The Cronbach’s alpha values for the organizational evaluation range from 0.49 to 0.78, with an overall score of 0.60, which can only be considered marginally satisfactory. The Cronbach’s alpha values for the individual evaluation are even more divergent, ranging from 0.17 to 0.78, with an overall score of 0.62. As discussed above, achievement (*ji*) was given a weight of 50% of the total performance output measurement. It had the lowest credibility score of 0.17. The overall Cronbach’s alpha of 0.62 is marginally acceptable. The low credibility level of the most important indicator raises considerable alarm.

Table 1. Test Results of the Trial Evaluation

Organizational Performance Evaluation		Individual Performance Evaluation	
	Cronbach’s alpha*		Cronbach’s alpha**
Functional execution	0.4921	Ability (<i>neng</i>)	0.3994
Managerial operation	0.7832	Diligence (<i>qin</i>)	0.7801
Targeted service	0.6176	Achievement (<i>ji</i>)	0.1731

Innovation and development	0.5327		
Total Score	0.6006	Total Score	0.6245

* Sample=12, there is a total of 12 offices in DOFHP.

** Sample=332, in the individual evaluation trials 167 employees participated.

Source: Report of DOFHP Evaluation Plan

Facing these results, the performance measures had to be modified after the trial run. For example, as shown in the Appendix, each primary indicator is further composed of several secondary indicators, with score values given to each secondary indicator. After discussing ways to improve the measures' validity with the field offices, the number of measurements in the Operationalized Measurements for Key Tasks Performance Evaluation was increased from 10 to 16 in the final stage. However, no further reliability tests were performed. One of the strategies used in developing performance measures for the DOFHP was to cover as much ground as possible so that the system will serve as a multidimensional evaluation and control tool to improve the “overall government work efficiency” (Department of Ocean and Fisheries of Hainan Province, 2012, p.3) to address the fear that “[if] attention is paid to what is measured, other areas may be neglected” (Pollitt et al., 2010, p.20). Therefore, this case demonstrated that validity was a consideration when the department designed the performance measures. However, the mechanism to improve the validity was limited. Regarding the reliability criterion, the measures developed by the DOFHP had not been tested usable in other agencies. The logic of “a multiplication of indicators” was evident in the process of developing the DOFHP performance measures. This strategy may improve the breadth of the performance measures but it risks becoming a burden to street-level officials, as discussed below.

Legitimacy and Acceptance

The legitimacy criterion of performance evaluation measures the extent to which the development of performance measures involves organizational employees and citizen groups and the level of acceptance received by them (Taylor, 2006). Support from top managers can help ensure sufficient funding, legitimize the implementation of the systems and set new performance expectations, whereas support from lower level managers and employees can minimize resistance to, and sabotage of, the systems (Berman & Wang, 2000). Legitimacy is achieved from either input into the design or approval of the performance measures by the stakeholders, and this legitimacy can also apply to citizens (Streib & Poister, 1999).

Internally, the introduction of the performance evaluation system is considered a “brain change” process in DOFHP. To change the thinking and action modes of the agency, DOFHP and the SOG project team organized several rounds of education and training, including basic knowledge training, key staff training, and internal auditor training for office staff, subordinate agencies, and field offices to enhance the understanding of the system.

Managers’ commitment to the performance evaluation system directly determines the system’s development and sustainability. As observed in this case, managers have a generally positive attitude towards the system and support its functions. However, the level of commitment differs according to the level of managerial position. Managers at the departmental level tend to fully support the system because they believe these systems would help improve the overall management efficiency and service capability. Managers in operational offices have a mixed attitude because these systems have a two-way impact on their offices. On one hand, the systems help clarify office functions and position descriptions, enforce strict management process, and promote employees’ motivation and morale. On the other hand, these systems provide additional pressures, such as an increased workload and escalating internal pressure. According to the new performance measures, all office tasks have to be completed in compliance with the procedure document and work instruction and be collected, recorded, and filed in accordance with the system. All of these detailed duties will increase the workload of

operational offices. These requirements will also increase the internal pressure by requiring supervisors to strictly evaluate their subordinates. The most difficult task is to award the 'annual best performance' because it is hard to develop a performance measure that fits all positions with significantly different duties. The supervisors must decide the award by performance, and, at the same time, maintain overall morale without discouraging other employees. Such a performance system intensifies the competition among offices because evaluation, to some degree, means competition.

As discussed in the previous section and shown in the Appendix, most of the performance measures developed were trying to 'rank' performance and thus make managers' tasks easier. Although these measures are primarily for internal use, the summative nature and the strong connection to incentives (e.g. awarding annual best performance) and sanctions (e.g. linking individual performance to organizational performance as shown in Figure 2) challenge the traditional administration model and management culture in this Chinese bureaucracy. Such a system received more support from high-level managers as a tool for hierarchical control than from mid- or low-level managers, who are burdened with competition among their peers, pressure to perform operational duties from multiple sources, and general red tape.

Credibility

The credibility of a performance evaluation system refers to the criterion that parallel systems are in place to validate the performance information, such as through performance auditing or standards that are similar to those of financial reporting of government organizations (Moynihan & Ingraham, 2003; Taylor, 2006). In this case, the primary tool to ensure credibility is an external review. According to the specific requirements of the project outline, the School of Public Management of Tsinghua University was selected as the third-party evaluation agency through a Single Source (SSS) procurement process after the program trial. The evaluation team reviewed the objectivity and applicability of the system through a variety of methodologies, including document review, in-depth interviews, group meetings, focus group discussions, and survey investigation. A report was submitted to DOFHP after a two-month evaluation process from February to April 2012.

The report states that over 84% of the employees acknowledged the importance of the introduction of a sound performance management system, 9% of the employees indicated no preference for such a system, and the remaining 7% did not respond to the survey question (Tsinghua University, 2012, p.21). However, regarding the techniques used in the evaluation, only 57.3% of the employees who participated in the survey indicated that the criteria were clear-cut and easy to use. About 32.9% of the respondents felt that the evaluation criteria were not easy to understand, while the remaining 9.8% did not respond (Tsinghua University, 2012, p.22). Overall, more than 82.6% of the employees were satisfied with the evaluation process, with a slightly lower satisfaction rate on the evaluation indicators (73.3%) and techniques (74.4%) (Tsinghua University, 2012, p.23-24). These results indicated a generally favourable attitude towards the overall system but less credibility regarding its performance measures.

It should be noted, however, that the 'independent, third-party' external evaluation was conducted by a recruited university team. All public universities, research institutions, and think tanks in China rely heavily on their relationship and connection with the government, in order to exercise their influence on policy-making and agenda-setting (Zhu & Xue, 2007). The extent to which this kind of report can be considered independent and unbiased requires further examination.

Functionality and Usefulness

Evaluating the interaction of measures in a performance management system and determining how useful the measures are gives us the system's functionality (Bouckaert, 1993). It is widely argued that multiple functions are to be fulfilled by implementing performance measures, from guidance in achieving goals to improving accountability, resource allocation, decision making, efficiency, and effectiveness (Behn, 2003). Accordingly, functional measures are believed to provide identifiable benefits for the development or maintenance of the organization (Streib & Poister, 1999).

The fundamental goals of the DOFHP performance evaluation system are to "promote the development of an international tourism destination of Hainan Island and the nation's ocean development strategy, implement provincial party and government's policies,

develop a modern government performance management model, and improve the public management and service capacity of the Department of Ocean and Fisheries” (Department of Ocean and Fisheries of Hainan Province, 2010, p.2). This system is also expected to direct performance evaluation for key tasks, regular tasks, agencies and individual employees.

In order to fully utilize the performance management system, DOFHP established a task force for performance evaluation, with the department chief serving as the director. The task force established an operating office to monitor the functions of the performance evaluation system. A committee was set up during the performance evaluation process to be in charge of the organization, coordination, guidance, and oversight of the evaluation. The committee was also responsible for approving the results of key and regular tasks, and organizational and individual evaluations. The committee was made up of the department chief, the human resources director, managers from other offices, and employee representatives. Employee representatives were selected through a nomination or voting process, representing different levels of positions. No more than one representative from each office was allowed. The total number of employee representatives was required to exceed one-third of the committee members. These policies provided ways to support the functioning and to improve the usefulness of the performance evaluation system.

Accessibility

Accessibility of performance information refers to both the internal accessibility for employees and managers and the external public accessibility (Hammerschmid, Van de Walle, & Stimac, 2013). In the DOFHP case, the performance system facilitates the internal accessibility. For example, it is recognized that an individual employee’s achievement, to a great extent, is associated with organizational advancement. Therefore, as shown in Figure 2, the individual evaluation of managers is related to the organizational evaluation, which is based on the four primary indicators of functional execution, managerial operation, targeted service, and innovation and development. The individual evaluation is performed on a basis of quarterly self-evaluation, mid-year evaluation, and year-end review procedure. At the year-end review, all offices scoring in the top 30% are awarded

an “outstanding” rating. Offices that are not among the top 30% but receive a score higher than 60 points will be rated “qualified.” Offices that scored lower than 60 will be rated ‘unqualified’ and reprimanded. The ratings of an office will have a significant effect on the managers. It is required that if an office receives an ‘unqualified’ rating, its manager cannot receive an ‘outstanding’ individual rating. If an office receives an ‘outstanding’ rating, its manager is automatically recommended for an ‘outstanding’ individual rating. With an organizational ‘qualified’ rating, the manager can be recommended for a ‘competent’ individual rating or above. Such information is available to staff and managers through the internal Office Administration (OA) system.

Public accessibility of performance information in government organizations should enable external stakeholders to hold the organizations accountable and help increase the transparency of the organizations’ internal operations (Taylor, 2006). Moynihan and Ingraham (2003) recognize that an effective result-based government needs to blend and balance political accountability with managerial accountability and performance measurement accountability. Among other unresolved issues, the accessibility and external accountability criteria pose significant challenges to the performance management efforts in this case. According to Yu (2011), the Chinese government has gradually opened up more information to the public, including budgetary information, personnel structures and data, and so on. The performance measures developed in this case are primarily intended for internal performance checking and management control. Methods of making this information available to the public have not been on the top of the agenda, which remains a weakness of such efforts. So far the only parties who assume the ‘ownership’ of the performance measures include the DOFHP itself and the academic institutions, (i.e. SOG at SYSU and Tsinghua University, which were involved in the program development process). Therefore, it is difficult to predict whether what Pollitt et al. (2010) describe as a diffusion of regime “ownership” or an external reluctance to trust the performance regime will happen. It can be said that the accessibility of the system remains unclear.

In general, the development of the performance evaluation system in the DOFHP case provided evidence to address many technical issues with significant management

challenges remaining. The following sections will discuss the findings in this case before concluding thoughts are drawn.

Findings and Discussion

As China deepens its administrative reforms and strives for a more service-oriented government, public organizations at different levels are paying increasing attention to developing technically sound and managerially effective performance management and evaluation systems. Such management reforms help Chinese governments achieve performance goals by strengthening the linkage between individual and organizational performance, enhancing organizational capacity of performance evaluation, and improving functional control by adopting advanced evaluation measures (Birkinshaw, Hamel & Mol, 2008; Lorena, Simmonds & Roman, 2012). By analyzing the DOFHP's effort to establish a highly sophisticated performance management system, this case study finds that the advancement of performance management in individual government agencies reflects is of a more technical nature due to agencies' more targeted functions. DOFHP was able to allocate ample financial resources to develop a state-of-the-art performance management system, integrating the ISO 9000 quality control and the Office Administration (OA) systems. The department's operational capacity allows them to invite experts to join the effort to improve performance evaluation and managerial efficiency.

Lu (2013) finds that multiple types of evaluators are involved in China's performance evaluation. This study further identified that domestic and international technical partners, such as the World Bank and two key universities in China, assumed important roles in the process of designing the performance measures by providing the department with leading performance ideologies and models. However, the involvement of universities as technical partners in this process is worth further examination.

On the one hand, universities and research institutions in China have become a major force by participating and assisting in government administrative reforms (Burns & Zhou, 2010; Zhu & Xue, 2007). Universities and institutions introduce the latest international experience in performance management to government and provide expertise in designing innovative policies. As seen in the DOFHP case, two key universities were

involved in the program design, external evaluations, and training activities for the department staff in the trial and implementation stages of the program.

However, in many cases, university teams were hired as consultants in a performance evaluation. In such an arrangement, the university team becomes an “agent” to implement the goals and objectives of government agencies and tends to be limited in its ability to design the system and to specify effective performance evaluation measures. Because of this, although many government agencies claim to have developed performance measurement systems, such systems are often at risk of being rudimentary and consisting merely of difficult-to-improve performance outcomes, such as workload and output measures (de Lancer Julnes, 2009; Thiel & Leeuw, 2002).

The implementation of such measures may create conflicts. As Dubnick (2005) and Yang (2011) point out, conflicts may exist between the two broad purposes of government-wide management reform: internal management (or performance improvement) and external political accountability. On the internal management side, the purposes of improving management efficiency sometimes conflict and can often raise tensions. The mixed attitudes towards the implementation of the comprehensive performance measures from different levels of managers and employees found in the DOFHP case reflected such conflict and tensions. Externally, performance management should provide information through reporting and compliance mechanisms for stakeholders and citizens to hold agencies accountable (Behn, 2003). The DOFHP case shows limited effort from the department to make the evaluation outcomes available to the public to allow them to monitor the agency’s performance. How to improve external accountability with such a highly technical system remains problematic.

After these findings, it is yet to be determined if performance evaluation serves as a political task or a management innovation in China. As in the DOFHP case, one of the key objectives of the project is to “fully implement the harmonious society goals of the Hainan province and build a service-oriented government” (Department of Ocean and Fisheries of Hainan Province, 2012). The DOFHP tried very hard to use this project to showcase their achievement in building a more effective and efficient department and to receive a

favourable commendation from the provincial government. It seems that a 'good looking' technical system is more likely to serve as the main goal of such an effort, rather than a *de facto* external control mechanism to actually improve accountability. Such a system may not be easily institutionalized to improve performance management practices. As there are few organizational or personnel-related mechanisms to make good use of performance information, it is difficult to build the link between the use of performance indicators and decision-making (Hammerschmid et al., 2013). Performance measures tend to remain formative, less implementable and uncertain of outcome (Pollitt & Dan, 2013).

Conclusions

As performance management remains a central concern in the field of public management research and practice, it should be understood within specific political and ideological contexts. This article fills the gap in the existing literature that largely overlooks agency-level performance management reform in China and responds to the western theories of the agency perspective of performance management. The DOFHP case represents a front-runner agency-level performance evaluation effort in China, owing to its support from the World Bank, ample financial resources, and state-of-the-art technical design. Findings in this case pose questions about how well such a reform can assist Chinese bureaucracies in improving efficiency and accountability.

It must be admitted that duplicating such a program for adoption in other settings may be challenging. Nevertheless, as in many other countries, agency-level performance evaluation is being adopted as an innovative tool to improve performance management in China (Daglio, Gerson & Kitchen, 2014). This trend is likely to continue. However, the present situation is reminiscent of Lagerkvist's (2005) analysis of China's e-government movement: "it is hard to avoid the idea that e-government is all about installing a machine that will enhance the productivity of a slow-moving bureaucracy" (p.197). Performance management itself cannot guarantee a more efficient and accountable government. Without adequately addressing the intrinsic technical and political issues of the reform, the goal of adopting performance evaluation reform as a genuine management innovation in China will remain difficult to achieve.

References

- Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review*, 63(5), 586–606.
- Berman, E., & Wang, X. (2000). Performance measurement in U.S. counties: Capacity for reform. *Public Administration Review*, 60(5), 409–420.
- Birkinshaw, J, Hamel, G., & Mol M. (2008). Management innovation. *Academy of Management Review*, 33(4), 825–845.
- Bouckaert, G. (1993). Measurement and meaningful management. *Public Productivity and Management Review*, 17(1), 31–44.
- Burns, J. P., & Zhou, Z. (2010). Performance management in the government of the People's Republic of China: Accountability and control in the implementation of public policy. *OECD Journal of Budgeting*, 2, 1–28.
- Chan, H. S., & Gao, J. (2008). Performance measurement in Chinese local governments. *Chinese Law and Government*, 41(2–3), 4–9.
- Chan, H. S., & Gao, J. (2009). Putting the cart before the horse: Accountability or performance? *The Australian Journal of Public Administration*, 68(S1): S51–S61.
- Daglio, M.; Gerson, D. & Kitchen, H. (2014, November), *Building organisational capacity for public sector innovation*. Paper presented at the OECD Conference “Innovating the Public Sector: from Ideas to Impact”, Paris, France.
- Dai, G., & Teng, W. (2008). *Performance evaluation at local government department level in China: A case study of People's Municipal Government of Hangzhou (PMGH)*. Nottingham, UK: Nottingham Trent University.
- de Lancer Julnes, P. (2009). *Performance-based management systems: Effective implementation and maintenance*. Boca Raton, FL: CRC Press.
- Department of Ocean and Fisheries of Hainan Province. (2010). *Bidding document of performance management system*. Haikou, Hainan: DOFHP Office.
- Department of Ocean and Fisheries of Hainan Province. (2012). *Completion report on performance evaluation of Department of Ocean and Fisheries of Hainan Province (Subproject D05–09)*. Haikou, Hainan: DOFHP.
- Dubnick, M. (2005). Accountability and the promise of performance: In search of the mechanisms. *Public Performance & Management Review*, 27(3), 376–417.

- Gao, J. (2009). Governing by goals and numbers: A case study in the use of performance measurement to build state capacity in China. *Public Administration and Development*, 29, 21–31.
- Gao, J. (2010). Hitting the target but missing the point: The rise of non-mission-based targets in performance measurement of Chinese local governments. *Administration & Society*, 42, 565–765.
- Hammerschmid, G., Van de Walle, S., & Stimac, V. (2013). Internal and external use of performance information in public organizations: Results from an international survey. *Public Money & Management*, 33(4):261–268.
- Lagerkvist, J. (2005). The techno-cadre's dream administrative reform by electronic governance in China today?. *China Information*, 19(2), 189–216.
- Liu, K. C. (1995). State control and performance evaluation: The case of a state-owned enterprise in China. *International Journal of Public Sector Management*, 8(6), 29–50.
- Lorena R. L., Simmonds, P. & Roman, L. (2012). *Trends and challenges in public sector innovation in Europe*. Brussels, Belgium: Technopolis Group.
- Lu, Y. (2007). Performance budgeting: The perspective of state agencies. *Public Budgeting & Finance*, 27(4), 1–17.
- Lu, Y. (2008). Managing the design of performance measures: The role of agencies. *Public Performance & Management Review*, 32(1), 7–24.
- Lu, Y., Willoughby, K., & Arnett, S. (2009). Legislating results: Examining the legal foundations of PBB systems in the States. *Public Performance & Management Review*, 33(2), 266–287.
- Moore, M. (1995). *Creating public value: Strategic management in government*. Cambridge, MA: Harvard University Press.
- Moynihan, D. P. (2008). *The dynamics of performance management: Constructing information and reform*. Washington, DC: Georgetown University Press.
- Moynihan, D. P., & Ingraham, P. W. (2003). Look for the silver lining: When performance-based accountability systems work. *Journal of Public Administration Research and Theory*, 13(4), 469–490.
- Ni, X. (2008). Reflections on recent performance measurement practices in Chinese government. *Journal of Sun Yat-sen University (translated from Chinese)*, 48(3), 134–141.

- Ni, X., & Yu, Q. (2009). Performance measures for local governments: The utilization of BSC, KPI, and performance diamond model. *Journal of Wuhan University (in Chinese)*, 62(5), 702–710.
- Nunnally, J. C. (1978). *Psychometric theory*, 2nd ed. New York, NY: McGraw-Hill.
- Oi, J. C. (1992). Fiscal reform and the economic foundations of local state corporatism in China. *World Politics*, 45(1), 99–126.
- Pollitt, C. (2013). The logics of performance management. *Evaluation*, 19(4), 346–363.
- Pollitt, C. & Dan, S. (2013). Searching for impacts in performance-oriented management reforms: A review of the European literature. *Public Performance and Management Review*, 37(1):7–32.
- Pollitt, C., Harrison, S., Dowswell, G., Jerak-Zuiderent, S., & Bal, Roland. (2010). Performance regimes in health care: Institutions, critical junctures and the logic of escalation in England and the Netherlands. *Evaluation*, 16(13), 13–29.
- Schmidle, T. P. (2011). Performance management: Laudable objectives, limited usage, lowered expectations. *Public Performance & Management Review*, 35(2), 370–389.
- Streib, G.D., & Poister, T.H. (1999). Assessing the validity, legitimacy, and functionality of performance measurement systems in municipal governments. *American Review of Public Administration*, 29(2), 107–123.
- Taylor, J. (2006). Performance measurement in Australian and Hong Kong government departments. *Public Performance & Management Review*, 29(3), 334–357.
- Thiel, S., & Leeuw, F. L. (2002). The performance paradox in the public sector. *Public Performance & Management Review*, 25(3), 267–281.
- Tsinghua University. (2012). *Evaluation report on performance evaluation of Department of Ocean and Fisheries of Hainan Province*. Beijing, China: Author.
- Walder, A. G. (1995). Local governments as industrial firms: an organizational analysis of China's transitional economy. *American Journal of Sociology*, 101(2), 263–301.
- Walker, R. M., & Wu, J. (2010). Future prospects for performance management in Chinese city governments. *Administration & Society*, 42, 345–555.
- Willoughby, K. G. (2004). Performance measurement and budget balancing: State government perspective. *Public Budgeting & Finance*, 24(2), 21–39.

- World Bank. (2006). China economic reform implementation project. Retrieved from <http://www.worldbank.org/en/news/loans-credits/2006/04/11/china-economic-reform-implementation-project>
- Yang, K. (2011). The Sisyphean fate of government-wide performance accountability reforms: Federal performance management efforts and employees' daily work, 2002–2008. *Public Performance & Management Review*, 35(1), 149–176.
- Yin, R. K. (2009). *Case study research: Design and methods (4th Edition)*. Los Angeles, CA: Sage.
- Yu, W. (2011). Open government information: Challenges faced by public human resource management in China. *International Journal of Public Administration*, 34(13), 879–888.
- Zhu, X., & Xue, L. (2007). Think tanks in transitional China. *Public Administration and Development*, 27, 452–464.

Appendix

Sample Score Sheet of Key Tasks Performance Evaluation Measures

Primary Indicators	Secondary Indicators	Operationalized Measurements	Technical Instructions
	Human Resources	Personnel 50%	The availability of personnel meets the demand of the key task.
	50%	Professionalism 50%	The personnel have the professional skills to fulfill the key task.
Input	Financial Resources	Fund appropriation rate 60%	Appropriated fund is sufficient to fulfill the key task.
20%	30%	Fund utilization rate 40%	The utilization of the fund is according to the plan of key tasks.
	Material Resources	Office equipment 100%	The availability of office space, equipment, and vehicle is sufficient.
	20%		
		Clear objective 35%	There is a clear target, with an explicit quality control standard.
Activities	Plan	Task allocation 30%	The responsibilities are effectively allocated to individual positions, with clear job descriptions and assessment standards.
30%	25%	Management duties 35%	There is an established job requirement and cooperation mechanism, with clear-cut working rules and administrative procedures.
	Do	Policy implementation 60%	Policy implementation follows laws, regulations, and procedures to make steady progress.
	35%		

	Resource management 40%	There are strict financial management regulations, electronic records, and work proceedings.
Check 25%	Measurement and supervision 50%	The collection of quality information is clear, with the implementation of streamlined control procedures.
	Analysis and evaluation 50%	Mechanisms are available to consistently analyze ongoing problems and to systematically report important issues.
Improvement 15%	Correction 40%	Mistakes and deviation from the task objectives can be remedied.
	Improvement 60%	Mechanisms are in place to improve policies and optimize operation.
Output 20%	Planning development, regulation implementation, administration, enforcement, public service, etc.	Offices report output indicators to the department, with an approval by the departmental level performance management committee.
Outcome 30%	Economic development, social benefit, administrative efficiency, environmental protection, and natural reservation, etc.	Offices provide outcome evidence, with an approval by the departmental level performance management committee.

Source: DOFHP Evaluation Plan