

A Study of Performance Information Use and Credibility in Performance Management Systems: Focusing on the Self-Evaluation System*

Kwang-Hee Lee** and Seog-Min Lee***

Abstract: This study aims to identify types of information use in South Korea's performance management system and the level of credibility accorded to these types of information use, along with the factors that affect them. The results show that use of performance measures is passive rather than instrumental and conceptual, while credibility accorded the system is low, and that the main factors that influence the system's use and credibility are leadership commitment and usefulness of information. These findings suggest that the current performance management system in Korea is not being well used and that a change is in order. In particular, external factors, such as reform of the U.S. system, might act as key drivers for change in the near future.

Keywords: Public Performance, performance management system, performance information use

INTRODUCTION

In the 1990s, the main public service evaluation system in the United States was the Government Performance and Results Act (GPRA) and the Program Assessment Rating Tool (PART). However, in 2008, the United States discontinued the use of PART in 2008 and introduced the Government Performance and Results Modernization Act

* This article is a revised version of "A Study on the Use of Evaluation Results and Performance Information" published by the Korean Institute of Public Administration in 2013. The authors wish to thank three anonymous reviewers for their valuable comments.

** Kwang-Hee Lee, first author, is a research fellow at the Korean Institute of Public Administration. E-mail: kwhelee@kipa.re.kr.

*** Seog-Min Lee, corresponding author, is an associate professor in the Department of Public Administration at the University of Suwon. E-mail: newmind68@suwon.ac.kr.

Manuscript received May 8, 2015; out for review June 24, 2015; review completed August 1, 2015; accepted August 6, 2015.

of 2010 (GPRAMA), which revised the GPRA. Among the main reasons for these changes were insufficient use of performance measures and distrust in the performance management system (Moynihan & Lavertu, 2012; OMB, 2011).

In South Korea (Korea hereafter), an assessment system was first introduced in the 1960s, but it wasn't until 2006 when a full-fledged performance management system was developed, after the Korean government adopted the GPRA and PART (Kong, Bang, & Yoon, 2012; M. J. Park & Choi, 2010), following Scotland and Thailand (Government Innovators Network, 2005). There has been substantial research on how performance information is used and what affects its use (Dull, 2008; Gilmour & Lewis, 2006; Moynihan, 2008; Moynihan & Lavertu, 2012; Moynihan & Pandey, 2010; Radin, 2006; Stalebrink & Frisco, 2011).both in Korea and in the United States, but there has been little research on the factors influencing data use, system credibility, and institutional change in relation to a benchmarking model (the U.S. system).

Thus, the primary goal is to identify how performance information is used in Korea and whether the problems that arose in the United States have likewise been found in Korea, which adopted the U.S. system. Secondary goals of include examining the factors that influence both the use and credibility of the system by surveying central government employees. In addition, we discuss the fate of the performance management system in Korea by comparing it with the U.S. system.

To this end, the present study explores types of information use in South Korea's performance management system and the credibility those uses earn along with the factors that affect use and system reliability. In order to do this, we have examined the results of previous studies and have also conducted in-depth interviews with stakeholders as well as administered a survey to a random, nationwide sample of employees in government agencies. Finally, we comprehensively compare and discuss the results derived from the survey and in-depth interviews. In the next section, we review the performance management system of Korea. Then, we describe the types of information use and the factors that affect their use, from which an analysis model is derived. Thereafter, the analysis results are discussed. The final section summarizes the results.

KOREA'S PERFORMANCE MANAGEMENT SYSTEM

A full-scale performance management system in Korea was introduced in 2006 using the U.S.'s GPRA and PART as models(OGPC, 2008; N.-W. Park, 2006). The Korean government enacted the Framework Act on Public Service Evaluation and the National Finance Act in 2006, which have been the basis of an integrated government affairs assessment system that gauges the performance management of the government's

enterprises since 2007. This system is based on self-assessment of departmental programs and outside assessments of cross-departmental items. This study focuses on the self-assessment system, which plays a significant role in Korea's performance management system.

The Office for Government Policy Coordination (OGPC) falls under the Office of the Prime Minister and is in charge of the Framework Act on Public Service Evaluation, which requires the government's agencies to produce strategy plans, annual execution plans, and annual performance reports. The agencies then link these plans to organizations, human resources, budgets, and compensation systems. Each administrative agency must revise its strategic performance plan every five years to address to changes in its responsibilities, vision, strategy, and performance targets. This system is very similar to the GPRA (OGPC, 2008). Agencies conduct an annual self-assessment of major policies (SAMP) according to their own self-assessment plan, which has five components: planning, implementation, result, utilization, and legislative management.

The Ministry of Strategy and Finance (MOSF) in charge of the National Finance Act introduced the self-assessment of budgetary programs (SABP). The SABP was introduced to strengthen links between performance information and budget allocation. The MOSF provides a standardized checklist for reporting self-assessment. The checklist contains 30 questions on program design and planning, 20 on management, and 50 on results and accountability. Each agency evaluates their programs by answering the questions provided by the checklist. The MOSF reviews the results of the self-assessment of programs carried out by line ministries/agencies. Its goal is to evaluate

Table 1. Self-Assessment System in Korea

	Self-Assessment of Major Policies	Self-Assessment of Budgetary Programs
managing organization	Office for Government Policy Coordination	Ministry of Strategy and Finance
targets	major policies	government financial program
indicators	selected autonomously by government departments	self-assessment indicators of financial projects
process	first assessment of the self-assessment committee of each government department determination of the assessment results through OGPC	first assessment of the self-assessment committee of each government department determination of the assessment results through confirmation/examination of MOSF
benchmarking model	GPRA	PART

all the budgetary programs by the end of a three-year period, undertaking an assessment of one-third of the programs every year. The SABP is called K-PART because it employs benchmarking against PART (Kong, Bang, & Yoon, 2012; M. J. Park & Choi, 2010). Table 1 compares the two self-assessment systems.

THEORETICAL BACKGROUND

Use Types and Credibility

Types of performance information use have traditionally been classified in terms of instrumental, conceptual, and other models: for example, instrumental • conceptual • persuasive use (Rossi & Freeman, 1985), instrumental • conceptual • symbolic use (Kirkhart, 2000), instrumental • conceptual • process • symbolic use (Johnson, 1998), instrumental • conceptual use (Newcomer, 2007), deterministic • interactive use (Moynihan, 2008), instrumental • noninstrumental use (Julnes, 2008), and passive • purposeful • perverse • political use (Moynihan & Lavertu, 2012). This study describes the instrumental, conceptual and, passive models. According to our interview data, performance measures in Korea are not used for symbolic purposes or for purposes of political persuasion, so these uses are excluded from our analysis.

The instrumental model is also called the “direct model,” “intentional use model,” “deterministic use model,” and “reasonable model,” as it is used to improve policy making (e.g., by identifying changes that need to be made in policy direction, targets, and budget and by effecting work adjustments and functional restructuring) or organizational performance (e.g., by suggesting performance rewards for government workers). Furthermore, this model can be expanded so that it can be used to improve the implementation stage and human resources management (De Kool, 2012; Moynihan & Lavertu, 2012).

The conceptual model takes into account indirect use and long-term effects for users, and thus is called the interactive model (Rossi and Freeman, 1985). This model helps information users increase their understanding of policies and projects through an assessment of the evaluation results and organizational learning (Van Dooren, Bouckaert, & Halligan, 2010; Julnes, 2008).

In addition to these two traditional models, another type has recently been developed that uses performance information to improve performance indicators or revise performance targets according to paradigmatic changes in the evaluation of performance management (Moynihan & Lavertu, 2012; Taylor 2009). The model, termed the passive model, is characterized by use of the performance information only for revising goals

and improving indicators (Moynihan & Lavertu, 2012; Taylor, 2009).

The passive model uses the evaluation results and performance data the least and does not meet the aims of Korea's performance management system. However, this model is important because it is premised on the idea of doing the minimum required to comply with the procedural requirements of performance systems (Radin, 2006). The instrumental model goes a step further in meeting the aims of Korea's performance management system because the improvement of policies falls within its scope. The conceptual model that takes into the long-term effects fully meets the aims of the system. Together the three models enable an assessment of information use that moves the passive (short-term effects) through the instrumental (mid-term effects) to the conceptual (long-term effects).

This study also examines system reliability. Managers and employees form judgments of leaders' credibility and rely on those judgments to guide their own behavior. At the center of contemporary reform are questions about the credibility of the commitments policy makers and administrators make to change how government agencies operate (Miller, 1992). Stated succinctly, credibility is a necessary condition for performance management. If managers fear that leadership is not credible, the capacity to produce credible information will not take hold (Dull, 2008).

Factors Contributing to Data Use and Credibility

Previous studies suggest that the characteristics of evaluators/users and the quality of evaluation processes and results are major factors in determining who uses performance information and level of credibility the information is afforded (Alkin, 1985; Forss, Rebien, & Carlsson, 2002; Rossi & Freeman, 1985). Although the factors outlined by these studies vary, the present study suggests the following four independent variables based on the results of in-depth interviews: evaluator characteristics, user characteristics, organizational/cultural characteristics, and the quality of the evaluation process.

Evaluator Characteristics

A number of studies suggest that evaluator characteristics are influential (Alkin, 1985; Forss, Rebien, & Carlsson, 2002; Rossi, 1985). For example, Donald Moynihan and Stéphane Lavertu (2012) found that the use of performance measures in the United States is influenced by the Office of Management and Budget and Government Accounting Office, which measure and manage performance in the U.S. federal government.

User Characteristics

While only a few evaluator characteristics that influence performance information use have been noted since the performance management system was established in the 2000s, user characteristics have been mentioned constantly (Alkin, 1985; De Kool, 2012; Forss, Rebien, & Carlsson, 2002; GAO, 2008; Hatry, 2008; C. Johnson & Talbot, 2008; Julnes & Holzer, 2001; Melkers & Willoughby, 2005; Moynihan & Lavertu, 2012; Moynihan, Pandey, & Wright, 2011; Newcomer, 2007; Rossi & Freeman, 1985). Although some studies have analyzed use by politicians (Askim, Johnsen, & Christophersen, 2008; De Kool, 2012; C. Johnson & Talbot, 2008), most studies have centered on government employees (Julnes & Holzer, 2001; Melkers & Willoughby, 2005; Moynihan & Lavertu, 2012; Moynihan, Pandey, & Wright, 2011; Newcomer, 2007). In particular, the participation in the GPRA and PART by U.S. federal government employees has been found to influence their use of performance data (Moynihan & Lavertu, 2012). In addition, studies have analyzed differences in the degree of usage according to the grade level and employment period of government workers and the type of work they do (Dull, 2009; Moynihan & Lavertu, 2012). When the users are government employees, such factors must be taken into consideration.

Organizational/Cultural Characteristics

Organizational/cultural characteristics are regarded as crucial variables, and they include leadership (Dull, 2008; Julnes, 2008; Melkers & Willoughby, 2005; Moynihan & Lavertu, 2012; Moynihan, Pandey, & Wright, 2011), performance management maturity, flexibility in decision making (Moynihan & Lavertu, 2012; Moynihan, Pandey, & Wright, 2011), organizational structure (Van Dooren, Bouckaert, & Halligan, 2010), and the degree to which the performance management of the organization is results oriented (Melkers & Willoughby, 2005).

Characteristics of the Evaluation Process

Many previous studies have explored factors related to the characteristics of the evaluation itself, such as evaluation methodology, evaluation design, validity of evaluation indicators, difficulty of measurement, data collection, fairness of the evaluation process, and validity and usefulness of evaluation results (Alkin, 1985; Alkin & Taut, 2002; Dull, 2008; Forss, Rebien, & Carlsson, 2002; Laegreid, Roness, & Rubecksen, 2008; Melkers & Willoughby, 2005; Rossi & Freeman, 1985; Van Dooren, Bouckaert, & Halligan, 2010).

RESEARCH DESIGN

Research Model

The present study uses passive, instrumental, and conceptual models of performance data use as dependent variables, and evaluator characteristics, user characteristics, organizational/cultural characteristics, and quality of the evaluation process as independent variables (see Table 2).

Before determining a research model and questionnaires for the survey, we carried out in-depth interviews with government employees and expert officials in the National Assembly of the Republic of Korea. These in-depth interviews were conducted with 11 government employees in charge of government evaluation and 3 officials in the National Assembly. After consulting related theories and previous studies, we designed questions to ascertain the types of use to which performance information is put, the credibility afforded to this information, and factors that influence Korea’s performance management system. The final analysis model and questionnaires were determined after interviews.

Dependent Variables

In this study, the passive use variable takes the average of three adjacent survey items, namely, whether respondents use performance information to set new performance

Table 2. Causal Model of Use and Credibility

Independent Variables			Dependent Variables	
Categories		Variables	Categories	Variables
evaluator characteristics		managing organization's attention	use	conceptual use
user characteristics		position grade employment period work type		↑ instrumental use
organizational/cultural characteristics		leadership of organization head department level		↑ passive use
characteristics of the evaluation process	activities	measurement difficulty trust in the evaluation committee	credibility	credibility (institutional trust)
	outputs	validity of the evaluation results usefulness of performance information		

→

goals or revise existing ones, refine program performance measures, and revise strategy and performance plans of organizations. The instrumental use variable takes the average of four items, namely, whether respondents use performance information to adopt new program approaches or change work processes, set program priorities, allocate resources, and restructure their organization. Finally, the conceptual use variable takes the average of three items, namely, whether respondents use performance information to gain a deeper understanding of their policies, generate news ideas, and identify and share effective program approaches with others. This study measures each use type and employs them as dependent variables in the analysis.

In a 2008 study, Matthew Dull used credible variables derived from surveys administered by the Government Accountability Office (GAO) to gauge if the implementation of the GPRA had improved agency performance up to that time and if it could improve agency performance. This study measures credibility by referencing the GAO survey items and the results of interviews in Korea and seeks to establish to what degree the SAMP and SABP are implicated in, and encourage, morale and improved agency performance.

Independent Variables

For characteristics of evaluators, this study identifies the extent of attention by the OGPC and MOSE, which are in charge of the performance management of government agencies, and in addition, analyzes how differences in this attention affect use of information. User characteristics can be seen as differences in the degree of usage according to the grade level and employment period of government workers and the type of work the government employee does (Dull, 2008; Moynihan & Lavertu, 2012). The impact of organizational/cultural characteristics is gauged by two items: commitment to achieving performance goals and commitment to analyzing the cause of performance results. Similarly, since different types of organization (e.g., ministries, vice ministries, and committees) have different characteristics, we explore whether these variations are an influential factor. Performance management culture and decision-making flexibility are excluded from our measurements because the in-depth interview data show that they are not regularly used in Korea to measurement performance.

The present study measures the characteristics of the evaluation process from two perspectives: evaluation activities and output. The measurement difficulty variable reflects evaluation activities in Korea and takes the average of three items: difficulty developing performance indicators, difficulty obtaining valid data, and difficulty measuring long-term effects. According to the interview results, the fairness of the evaluation process directly influences the acceptance of the evaluation results in

Korea. The fairness of the process is gauged by determining whether assessment committees properly understand the policy of a department, set aside enough time for the evaluation, and fully communicate with government workers.

Four items are used to measure the evaluation outputs, which also derive from the interviews, namely, whether the unique characteristics of the evaluated program have been properly considered, whether the program has achieved its results, whether the evaluation results are consistent, and whether the evaluation results were produced according to a scientific method. Last, the usefulness of the evaluation results are measured by referencing the measurement items of the U.S. Government Accounting Office (2009), which are whether the information obtained from the performance measurement is helpful for identifying implementation processes, service production (output), intended results (outcome), and quality of the policy.

Data Collection and Analysis Method

Data collection and analysis were conducted through surveys of government employees. After the interviews, a questionnaire was developed in accordance with the interview data and the analysis model, using a 6-point Likert scale (1=strongly disagree, 2=disagree, 3=somewhat disagree, 4=somewhat agree, 5=agree, 6=strongly agree). The validity of the survey questions was reviewed first at a meeting of performance management experts. Then, a pilot test was conducted with the government employees. The final questionnaire was designed taking into account the results of the pilot test. The survey was administered in 2013, and all survey items targeted government employees' perceptions of performance data use and the credibility of performance data with respect to both the SAMP and SABP.

The survey sample comprised 378 people (an average of 8 people from each of the 41 central administrative organizations). Of the respondents, 68 (18%) were in charge of performance management accounting and the remainder managed their own programs. As to level of employment, 44.2% worked at grade 6 or lower, 39.7% at grade 5, and 16.2% at grade 4 or higher. With regard to employment period, 19% had less than 5 years' experience, while 14.8% had 25 years' experience or more. Finally, 40.5% worked for ministries, 16.9% for organizations/committees at the vice ministry level, and 42.6% for agencies.

This study employs ordered logit regression analysis, in which each item is averaged into indicators for dependent variables, and then the indicators are transformed into positive integer values. In addition, this study investigates the linear path relationship between variables using hierarchical regression analysis.

RESULTS AND DISCUSSION

Information Use Types and Credibility in System

As summarized in table 3, Korean government employees show passive use (4.35), instrumental use (4.19), and conceptual use (4.11). The mean difference between passive use and conceptual use is statistically significant.

In the interviews, government employees responded positively to the passive use of information. In particular, they all were concerned about the results of evaluations of their organizations, as these influence their individual evaluations; moreover, they were used to the work they were doing, as they had been in their jobs for several years, and organizations they worked for were also engaged in the monitoring and managing of work.

Table 3. Data Description for Use and Credibility

Variable	Description	Mean	Mean(SD)/ Cronbach α
passive use	setting new or revising existing performance goals	4.35	4.35(0.84)/ .94
	refining program performance measures	4.39	
	revising strategy and performance plan of organization	4.32	
instrumental use	adopting new program approaches or changing work processes	4.30	4.19(0.84)/ .91
	setting program priorities	4.31	
	allocating resources	4.17	
	restructuring the organization	3.98	
conceptual use	developing a deep understanding of policies	4.19	4.11(.087)/ .93
	creating new ideas	4.03	
	identifying and sharing effective program approaches with others	4.13	
analysis of variance: multiple comparison a-b=.11/0.24 (mean difference / p-value), a-c=.24/0.00, b-c=.13/0.12			
credibility	is embedded within government employees and agencies	3.59	3.73(1.01)/ 0.92
	boosts morale among government employees	3.39	
	contributes to improving agency's performance	3.57	

However, with regard to instrumental use, even though there were some cases of policy measure improvements, respondents replied that only a little information was reflected in the budget or in the organization's operations. In addition, with regard to conceptual use, they responded that although some organizations had introduced advanced performance management techniques, it was unusual to use the evaluation results for daily interactive learning. According to a staff member in charge of performance management, "The evaluation opinions were not all reflected; they were partially reflected. Realistically, there are many limitations when linking the evaluation results to the organization, human resources, and budgets."

As shown in table 3, the credibility score for Korea's performance management system is only 3.50 (50%). The officials of the National Assembly doubt the authenticity of government performance information and are indifferent about using it, doing so only as a reference. Another National Assembly staff member stated, "I have not seen any member of the National Assembly ask a question by using the self-assessment report. The ruling party and the opposition parties as well as the standing committees share this viewpoint."

Factors Affecting Use

Table 4 shows that the managing agencies' attention (4.50) and the leadership of organization heads (4.87) is relatively high regarding the use of performance information. Measurement difficulty is high (4.65), and trust in the government's departmental self-assessment committee is relatively low (4.05). In particular, respondents point to poor communication between evaluators and government employees (3.87). It is not surprising, therefore, that the respondents did not find the evaluation results to be very useful (4.22).

Table 5 shows which factors affect each use type in the Korean performance management system. Passive, instrumental, and conceptual uses are the dependent or response variables in the ordered logit regression. Coefficients of variables are the predictors in the model and the cut points for the adjacent levels of the response variables. In terms of evaluation characteristics, the attention of the institutions in charge of performance management, the OGPC and MOSF, has no influence over data usage by frontline government employees. In other words, ordinary government employees do not care about the attention of the performance managing agencies; indeed, the negative regression coefficient implies that the employees may not regard these agencies favorably.

With regard to user characteristics, the position level and employment period of respondents are not significant variables for data use. According to the organizational/cultural characteristics, when there is high commitment to data use by the organization

Table 4. Data Description for Factors Affecting Use and Credibility

Variable		Description	Mean	Mean(SD)/ Cronbach α	
managing organization's attention		OGPS's attention in the use of the performance information	4.60	4.50(0.96)/ 0.85	
		MOSF's attention in the use of the performance information	4.39		
leadership commitment to results		commitment to achieving performance goals	5.05	4.87(0.77)/ 0.80	
		commitment to analyzing cause of performance results	4.74		
process quality	measurement difficulty	difficulty generating meaningful measures	4.74	4.65(0.87)/ 0.87	
		difficulty obtaining valid or reliable data	4.50		
		difficulty measuring long-term effects	4.72		
	trust in the evaluation committee	has an in-depth understanding the policy of the department	4.26	4.05(0.97)/ 0.93	
		sets aside enough time to provide accurate evaluation results	4.03		
		communicates properly with people in charge of programs	3.87		
output quality	validity of evaluation results	due consideration of the unique characteristics of the evaluated program	4.12	4.06 (0.92)/ 0.95	
		proper consideration of the question of whether the program has achieved its goals	4.17		
		generation of consistent evaluation results for the same program	4.04		
		use of a scientific method to gauge evaluation results	3.92		
	usefulness of performance information	aids in identification of implementation processes	4.20	4.22 (0.83)/ 0.93	
		aids in identification of service production (output)	4.24		
		aids in identification of intended results (outcome)	4.37		
		aids in identification of quality of policy	4.07		

Table 5. Factors Affecting Use

		Variable	Passive	Instrumental	Conceptual
evaluator characteristics		managing agency's attention	-.15(.13)	-.07(.13)	-.11(.13)
user characteristics	position grade (basis: grade 6)	grade 5	-.11(.24)	-.45(.24)	-.43(.25)
		grade 4 or higher	-.06(.34)	-.57(.33)	-.17(.35)
	employment period	employment period	.01(.01)	.02(.01)	.01(.01)
	work type (basis: performance management)	program management	.41(.29)	.49(.28)	.57(.29)
organizational/cultural characteristics	leadership of the organization head	leadership of the organization head	.74(.16)**	.84(.16)**	.80(.16)**
	department level (basis: agencies)	vice ministry • committee	.13(.32)	.94(.32)**	.48(.34)
		ministry	-.33(.25)	.05(.24)	-.03(.26)
evaluation process characteristics	self-assessment activities	measurement difficulty	.33(.13)**	.13(.12)	.07(.13)
		trust in self-assessment committee	.68(.18)**	.56(.17)**	.58(.18)**
	information quality	validity of evaluation results	.26(.19)*	.44(.19)*	.53(.20)**
		usefulness of evaluation results	1.53(.19)**	.96(.17)**	1.53(.20)**
/cut1		6.30(1.09)	5.45(1.08)	4.73(1.02)	
/cut2		8.65(1.03)	7.85(0.99)	7.61(0.98)	
/cut3		11.76(1.01)	10.79(1.05)	10.45(1.05)	
/cut4		15.33(1.22)	14.21(1.16)	14.37(1.17)	
/cut5		19.60(1.40)	18.07(1.32)	17.89(1.31)	
N		374	374	374	
LR χ^2		331.17**	271.20**	338.32**	
Pseudo R ²		.34	.28	.37	

Note: Significance levels are based on two-tailed or chi-square tests: ** p < 0.01, * p < 0.05

head, government employees are more concerned about performance information use, and this type of leadership has greater impact on data use. In other words, a one unit increase in leadership scores would result in a 0.74, 0.84, and 0.80 unit increase respectively in the ordered log-odds of being in a higher passive, instrumental, and conceptual category respectively when the other variables in the model are held con-

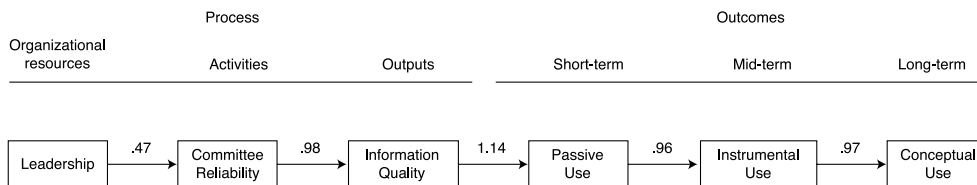
stant. However, there is no difference in use between ministries, vice ministries, and agencies.

When it comes to the self-assessment activities related to the evaluation process, measurement difficulty does not influence the use of performance data except passive use. It seems that the more employees recognize measurement difficulty, the more they focus on only passive use (0.33) and the less concerned they are about instrumental and conceptual use. Trust in the self-assessment committee positively influences each use (0.68, 0.56, and 0.58 respectively). With respect of output quality, the validity and usefulness of the evaluation results also positively impacts every type of use; in particular, of all other variables, the usefulness of the performance data has the largest impact (1.53, 0.96, and 1.53 respectively).

This study adopts a logic model as an evaluation utilization process model. Figures 1 and 2 demonstrate the path diagram of use. Following the model of figure 1, we conduct a path analysis using a hierarchical regression method in order to investigate whether the variables influence one another, especially whether they affect passive, instrumental, and conceptual use. Table 5 shows that leadership, committee reliability, and information quality variables are significant influential variables when it comes to the use of performance information.¹ According to the logical model, these variables are designated factors: leadership as an organizational resource factor, committee reliability as an evaluation activity factor, and information quality as an output. The model fit for path analysis is fine (goodness of fit index = 0.974, adjusted goodness of fit index = 0.923, comparative fit index = 0.977, normed fit index = 0.968, and root mean square error of approximation = 0.079).

Figure 1 and table 6 demonstrate the results of the path analysis: all paths demonstrate a positive significant relationship and, in particular, passive use impacts positively on instrumental use, which, in turn, impacts on conceptual use. Figure 1 shows a linear

Figure 1. Path Diagram of Use



Note: arrows (→) indicate direct effect.

1. The information quality variable is generated by summing up the validity and usefulness of the evaluation results.

Table 6. Standardized Total Effects (Use)

	Leadership	Committee Reliability	Information Quality	Passive Use	Instrumental Use
committee reliability	.47*				
information quality	.46*	.98*			
passive use	.53*	1.12*	1.14*		
instrumental use	.50*	1.07*	1.09*	.96*	
conceptual use	.49*	1.04*	1.06*	.93*	.97*

** p < 0.01, * p < 0.05.

direct positive effect from leadership to conceptual use. The columns of each variable in table 6 reveals the total effect: for example, leadership has a direct and total effect (.47) on committee reliability, a total effect (.46) on information quality, a total effect (.53) on passive use, a total effect (.50) on instrumental use, and a total effect (.49) on conceptual use.

Factors Affecting Credibility

This subsection attempts to identify the impacts of each factor on the reliability of the performance management system. Table 7 indicates that respondents do not correlate system reliability with the attention of the OGPS and MOSF.

With respect to leadership Government workers believe in the system (0.35) when the leadership of the organization head appears strong to them. With regard to evaluation process characteristics, the impact of measurement difficulty is strong and negative (-0.42) in Korea. Reliability of the self-assessment committees (0.33) and evaluation results (0.38) leads to trust in the Korean system. Like use, the usefulness of the evaluation results is the strongest and most positive variable in connection with the credibility accorded to the system. If a respondent were to increase his or her leadership score by one point, his or her ordered log-odds of being in a higher credibility category would increase by 0.97 if the other variables in the model are held constant.

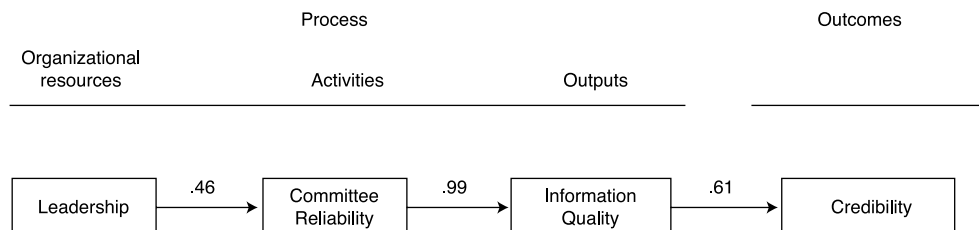
We conduct the same analysis as that for the path analysis of use to find the influential relationships between variables. Based on the results shown in table 7, leadership, committee reliability, information quality, and institutional credibility variables are selected. As shown in figure 2, these factors are aligned according to the logical model with leadership as an organizational resource factor, committee reliability as an evaluation activities factor, information quality as an output, and credibility as outcome

Table 7. Factors Affecting Credibility

		Variable	Credibility
evaluator characteristics		managing agency's attention	.07(.12)
user characteristics	position grade (basis: grade 6)	grade 5	-.20(.22)
		grade 4 or higher	-.20(.31)
	employment period	employment period	-.01(.01)
	work type (basis: performance management)	program management	-.09(.26)
organizational/cultural characteristics	leadership of organization head	leadership of organization head	.35(.14)*
	department level (basis: agencies)	vice ministry - committee	.12(.29)
		ministry	.22(.22)
evaluation process characteristics	self-assessment activities	measurement difficulty in SAMP	-.42(.12)*
		trust in self-assessment committee	.33(.17)*
	information quality	validity of evaluation results	.38(.18)*
		usefulness of evaluation results	.97(.17)*
/cut1			-12.03(1.08)
/cut2			-8.41(0.94)
/cut3			-5.78(0.88)
/cut4			-3.94(0.86)
/cut5			-1.53(0.88)
N			374
LR χ^2			190.85*
Pseudo R ²			.18

Note: significance levels are based on two-tailed or chi-square tests: ** p < 0.01, * p < 0.05.

Figure 2. Path Diagram of Credibility



Note: arrows (→) indicate direct effect.

variable.² In figure 2, the model fit for path analysis is fine (adjusted goodness of fit index = 0.990, comparative fit index = 0.990, normed fit index = 0.968, and root mean square residual = 0.015). The path diagram indicates the linear direction of influential relationships between variables: reliability of leadership and the self-assessment committee leads to trust in the quality of the information, which, in turn, leads to credibility in the system. In other words, the more government workers positively recognize leadership, committee reliability, and information quality, the more they trust the performance management system.

The columns of each variable in table 8 indicate the total effect. Leadership has a direct and total effect of 0.46 on committee reliability, a total effect of 0.46 on information quality, and a total effect of 0.28 on institutional credibility.

Table 8. Standardized Total Effects (Institutional Credibility)

	Leadership	Committee Reliability	Information Quality
committee reliability	.46*		
information quality	.46*	.99*	
institutional credibility	.28*	.61*	.61*

** p < 0.01, * p < 0.05

Sustainability of the System: Institutional Redesign in Korea?

In the United States, it has been pointed out that assessment results and information about federal government projects have not been utilized properly in the budget review process (Frisco & Stalebrink, 2008; Gilmour & Lewis, 2006; Radin, 2006). Federal managers responded positively to the system’s effectiveness in collecting evidence related to project efficiency during the assessment process but negatively to how resources were allocated in order to assess project efficiency (Newcomer, 2007).

The GAO (2008) also found that, despite legislative and administration efforts, the overall use of performance data did not change between 1997 and 2007 and that even for a survey item asking about the use of such data in “adopting new programs or changing work process,” the percentage decreased significantly. The Obama administration criticized the GPRA and PART for the failure of government agencies to make use of performance data and stressed the importance of such usage in improving outcomes

2. Although measurement difficulty is a significant variable for credibility in an ordered logit regression model, it is no longer a significant factor in path analysis in figure 2.

(OMB, 2011). In time, the Obama administration discontinued the use of the PART assessments, and Congress established the GPRAMA, which requires agencies to provide performance data on a quarterly basis and to use those data or else risk closer oversight by the OMB and Congress (Kamensky, 2011).³ The OMB used the PART for the last time in 2008.

In Korea, we know that the conceptual use of data is relatively weak, and faith in the performance management system is not strong. In addition, 2009 survey results (J.-H. Park & Lee, 2009) show that trust in the self-assessment committee was 67% and that the usefulness of data was judged to be 62%, which was measured by whether the data provided valuable information for finding the causes of problem and for improving programs. In comparison with our survey results (61% and 65%, respectively), there seems to be no difference between 2008 and 2013. To promote the use of information data in Korea, and especially to enhance instrumental and conceptual use, we advise that leadership commitment is demonstrated, trust in evaluation committees is built, and information is made to be useful. In any event, Korea, like the United States, seems to be faced with a choice of either gradual improvement or overall reform.

The government-supported Korean Institute of Public Administration conducted a survey among civil servants in charge of government evaluations and specialists in the field of policy evaluation that revealed that the GPRAMA is the preferred system among foreign systems (Hwang, 2013). In addition, the institute suggested abolishing the SAMP and SABP (KPART) and improving institutions focused on performance information use (Lee, 2013). To date, Korea has used a catch-up strategy for economic development (Guillén, 2001) and has adopted the same strategy to establish or improve institutions, in particular, preferred institutional or mimetic isomorphism to minimize future uncertainty and risk (Joo & Halx, 2012). Therefore, mimetic isomorphism might be a key predictor of institutional change in the Korean performance management system. However, Korea has been a deliberate pursuer of performance management and thus, will pay close attention to the how the new U.S. system fares.⁴

3. The problem of use in performance management is the same in the U.K. as in the U.S. (James, 2004). The Public Service Agreement system introduced in the U.K. was an exemplary performance management system. However, according to the National Audit Office, approximately 50% of the surveyed high-ranking officials pointed out serious problems in the performance assessment process and 75% stated that the lack of cooperation between departments was very serious (James, 2004). According to a survey of parliament and committee staff, committees' attention to results was low (Johnson & Talbot, 2008). Consequently, the Public Service Agreement system was replaced with business plans in 2010.

4. While the PART of the United States was instituted in 2002, the SABP (KPART) of Korea

According to the Government Accounting Office (2014), agencies' reported use of performance information generally did not improve between 2007 and 2013, which suggests that Korea will wait before making changes to its the system.

CONCLUSION

The purposes of performance management systems vary, but the most important one is to make use of performance measures. Recently, however, as shown by the changes in the conventional performance management system used by the United States, such systems have been relatively unsuccessful in this respect. In this study, we have examined the use of performance information in and the trust extended to the Korean performance management system, which is very similar to the U.S. system, and draw the following conclusions.

First, the use of performance measures is passive rather than instrumental and conceptual; moreover, the perceptions of government employees are negative. These findings suggest that the performance management system in Korea is not being well used.

Second, leadership, measurement difficulty, reliability of the self-assessment committee, and performance data quality are important to the level of trust accorded to the Korean system. Furthermore, leadership, committee reliability, information quality, and credibility of the performance management system are correlated by a linear causal relationship. Leadership is a strong and positive predictor for the use and credibility of performance information in Korea. In addition, the usefulness of data has a strong impact on use and credibility, and the measurement difficulty variable might be an influential factor for performance management.

In light of these results, it seems likely that the Korean system undergo change. In particular, external factors, such as reform of the U.S. system, might act as key drivers for change in the near future.

was adopted in 2008 after much research by Korean government supported institutes (i.e., Korean Institute of Public Administration, Korea Institute of Public Finance) about the effects of PART.

REFERENCES

- Alkin, M. C. 1985. *A guide for evaluation decision-makers*. Beverly Hills, CA: Sage.
- Alkin, M. C., & Taut, S. 2002. Unbundling evaluation use. *Studies in Educational Evaluation*, 29(1): 1-12.
- Askim, J., Johnsen, Å., & Christophersen, K.-A. 2008. Factors behind organizational learning from benchmarking: Experiences from Norwegian municipal benchmarking networks. *Journal of Public Administration Research and Theory*, 18(2): 297-320.
- De Kool, D. 2012. The utilization of performance information: The Dutch network ecological monitoring. *Environmental Policy and Governance*, 22(1): 1-13.
- Dull, M. 2008. Results-model reform leadership: Questions of credible commitment. *Journal of Public Administration Research and Theory*, 19(2): 255-284.
- Forss, K., Rebien, C. C., & Carlsson, J. 2002. Process use of evaluations: Types of use that precede lessons learned and feedback. *Evaluation*, 8(1): 29-45.
- Frisco, V., & Stalebrink, O. J. 2008. Congressional use of the Program Assessment Rating Tool. *Public Budgeting and Finance*, 28(2): 1-19.
- Government Accountability Office (GAO). 2008. *Government performance: Lessons learned for the next administration on using performance information to improve results*. Washington, DC: Government Printing Office.
- Government Accountability Office (GAO). 2009. *Government performance: Strategies for building a results-oriented and collaborative culture in the federal government*. Washington, DC: Government Printing Office.
- Government Accountability Office (GAO). 2014. *Managing for results: Agencies' trends in the use of performance information to make decisions*. Washington, DC: Government Printing Office.
- Gilmour, J. B. 2007. Implementing OMB's Program Assessment Rating Tool (PART): Meeting the challenges of integrating budget and performance. *OECD Journal on Budgeting*, 7(1): 1-40.
- Gilmour, J. B., & Lewis, D. E. 2006. Assessing performance budgeting at OMB: The influence of politics, performance and program size. *Journal of Public Administration Research and Theory*, 16(2): 169-186.
- Government Innovators Network. 2005. Program Assessment Rating Tool. Retrieved on January 1, 2015, from <http://www.innovations.harvard.edu>.
- Guillén, M. F. 2001. *The limits of convergence: Globalization and organizational change in Argentina, South Korea, and Spain*. Princeton, NJ: Princeton University Press.
- Hatry, H. P. 2008. The many faces of use. In W. van Dooren & S. van de Walle (eds.),

Performance information in the public sector: How it is used (pp.). Houndmills, UK: Palgrave Macmillan.

- Huberman, M. 1990. Linkage between researchers and practitioners: A qualitative study. *American Educational Research Journal*, 27(2): 363-391.
- Hughes, A. 2006. Testimony to the senate homeland security and government affairs subcommittee on federal financial management, government information, and international security. *OMB Watch*, June 13.
- Hwang, H.-S. 2013. *Recommendations for improving the Korean public service evaluation system from a comparative perspective*. Seoul, Korea: Korean Institute of Public Administration.
- James, O. 2004. The UK core executive's use of public service agreements as a tool of governance. *Public Administration*, 82(2): 397-419.
- Johansson, T., & Siverbo, S. 2009. Explaining the utilization of relative performance evaluation in local government: A multi-theoretical study using data from Sweden. *Financial Accountability and Management*, 25(2): 197-224.
- Johnson, C., & Talbot, C. 2008. UK parliamentary scrutiny of public service agreements: A challenge too far? In W. van Dooren & S. van de Walle (eds.), *Performance information in the public sector: How it is used* (pp. 140-156). Houndmills, UK: Palgrave Macmillan.
- Joo, Y. H., & Halx, M. D. 2012. The power of institutional isomorphism: An analysis of the institutionalization of performance-based pay systems in Korean national universities. *Asia Pacific Education Review*, 13(2): 281-297.
- Julnes, P. de L. 2008. Performance measurement beyond instrumental use. In W. van Dooren & S. van de Walle (eds.), *Performance information in the public sector: How it is used* (pp. 58-71). Houndmills, UK: Palgrave Macmillan.
- Julnes, P. de L., & Holzer, M. 2001. Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Public Administration Review*, 61(6): 693-708.
- Kamensky, J. M. 2011. The Obama performance approach: A mid-term snapshot. *Public Performance and Management Review*, 35(1): 132-147.
- Kirkhart, K. E. 2000. Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation*, 88: 5-23.
- Kong, D.-S., Bang, M.-K., & Yoon, K.-U. 2012. Influential factors in the evaluation results of finance programs. *Journal of Korean Policy Analysis and Evaluation*, 17(4): 27-56.
- Laegreid, P., Roness, P. G., & Rubecksen, K. 2008. Performance information and performance steering: Integrated system or loose coupling. In W. van Dooren &

- S. van de Walle (eds.), *Performance information in the public sector: How it is used* (pp. 125-141). Houndmills, UK: Palgrave Macmillan.
- Lee, H.-S. 2013. *Improving the development and operation of performance management plans*. Seoul, Korea: Korean Institute of Public Administration.
- Melkers, J., & Willoughby, K. 2005. Models of performance-measurement use in local governments: Understanding budgeting, communication, and lasting effects. *Public Administration Review*, 65(2): 180-190.
- Moynihan, D. P. 2008. *The dynamics of performance management: Constructing information and reform*. Washington, DC: Georgetown University Press.
- Moynihan, D. P., & Lavertu, S. 2012. Does involvement in performance management routines encourage performance information use? Evaluating GPRA and PART. *Public Administration Review*, 72(4): 592-602.
- Moynihan, D. P., & Pandey, S. K. 2010. The big question for performance management: Why do managers use performance information? *Journal of Public Administration Research and Theory*, 20(4): 849-866.
- Moynihan, D. P., Pandey, S. K., & Wright, B. E. 2011. Setting the table: How transformational leadership fosters performance information use. *Journal of Public Administration Research and Theory*, 22(1): 143-164.
- Newcomer, K. 2007. How does program performance assessment affect program management in the federal government? *Public Performance and Management Review*, 30(3): 332-350.
- Office for Government Policy Coordination (OGPC). 2008. *Performance management and evaluation in public service*. Seoul: Korea.
- Office of Management and Budget (OMB). 2011. *The president's budget for fiscal year 2012: Analytical perspectives*. Washington, DC: Government Printing Office.
- Park, J.-H., & Lee, K.-H. 2009. *A study of the content of performance information and the relevance of performance measures*. Seoul, Korea: Korean Institute of Public Administration.
- Park, M. J., & Choi, S. R. 2010. Analysis of the impact of the program assessment rating tool on program budget reduction. *International Review of Public Administration*, 14(3): 25-35.
- Park, N.-W. 2006. *Performance management in the Korean public sector*. Seoul, Korea: Center for Performance Evaluation and Management.
- Radin, B. A. 2006. *Challenging the performance movement: Accountability, complexity, and democratic value*. Washington, DC: Georgetown University Press.
- Rossi, P. H., & Freeman, H. E. 1985. *Evaluation: A systematic approach*. London:

Sage.

- Stalebrink, O., & Frisco, V. 2011. PART in retrospect: An examination of legislators' attitudes toward PART. *Public Budgeting and Finance*, 31(2): 1-21.
- Taylor, J. 2009. Strengthening the link between performance measurement and decision making. *Public Administration*, 87(4): 853-871.
- Van Dooren, W., Bouckaert, G., & Halligan, J. 2010. *Performance management in the public sector*. Oxford, UK: Routledge.