Book Reviews

Program Evaluation and Performance Measurement: An Introduction to Practice, James C. McDavid & Laura R. L. Hawthorn, SAGE Publications, Inc (2006), 496 pages, ISBN: 9781412906685

This book deals with the theories and practices of evaluation. It extensively covers useful information on concepts and methodologies that are actually demanded of in carrying out evaluation and performance measurement of programs in public sector, and it is expected to play the role of textbook or guideline for the readers. For these reasons, this review examines contents of the book without stepping out of the boundaries of marring the essence of the contents.

This book, compose of total of 12 Chapters, begins with comprehensive consolidation of concepts and issues that will be the requisites in understanding the theme to be dealt with later, and continues on with discussions on individual themes.: key concepts and issues in program evaluation and performance measurement, understanding and applying program logic models, research designs for program evaluation, measurement in program evaluation, applying qualitative evaluation methods, assessing the need for programs, concepts and issues in economic evaluation, performance measurement as an approach to evaluation, design and implementation of performance measurement systems, using and sustaining performance measurement systems, program evaluation and management: joining theory and practice, and the nature and practice of professional judgment in program evaluation.

Although conclusions are not deduced comprehensively, it was possible to verify the directionality asserted by the authors in the process of their deployment. The following is the summary of this:

- Selective application of quantitative and qualitative analysis methods
- Maintenance of understanding and balance on the political practices of an organization
- Integration of performance measurement and evaluation practice
- Development of sound professional judgment

Logic Model categorizes and explains the causal linkages of program. This is essential in program evaluation.

The program is in open-systems, that is, it is in interactive relationship with its environment. Logic model visually represents[1] the program in such open-systems. It illustrates how the resources of the program has been converted into activities and intended results. It is possible to categorize the program and identify it with the external environment by disclosing the causal linkages of its construct through this. Categories presented here are as follows: inputs, components, implementation objectives, and outcomes.

Logic model is essential in program evaluation. Logic model that relies on the qualitative research method is an important foundation in understanding whether effectiveness of program exists or how it is produced, and it is also very useful in developing performance measurement system for monitoring of program outputs and outcomes. However, one must bear in mind that the logic model cannot contain all the aspects of the program in detail.

Process of visual representation of program in logic model is iterative and combines with various other activities. In such process, the relationship with the

---

1) From such meaning, it has been disclosed that utilization of flow charts is also effective, which is in complementary relationship with logic model.

organizational goal is an important issue. Since the organization ecologically pursues behavioral goals – for example, putting priority on the survival of the organization – it may not coincide with normative goals, and, eventually, they could be conflict between the program goals and behavioral goals. In addition, there is time-limitation due to environmental changes in the program. For these reasons, efforts to structuralize the performance measurement system are being emphasized.

Experimental design is main tool in research design and the linkage with program logic model is important.

The key to evaluation is judging whether the program was effective. Therefore, research design must be able to examine the linkage with outcomes related to program.

There is need to pay attention to statistical validity, internal validity, construct validity and external validity that Cook & Campbell presented. These eventually results in the issue of variable and generalization. Research design must focus on constructing and completing evaluation that can be trusted, and experimental design is seen as a main tool - randomized experimental design and quasi-experimental design.

Examination method that can exclude other factors that can impart influence on causal linkages at program evaluation must be chosen. Research design plays important role in segregating such causal linkages, and the linkage with program logic model is required. However, research design that examines all the causal linkages within the logic model requires enormous resources and has limitations on control. These limitations mean that other observations and professional judgment other than the evidences can be considered together. There is need for evaluator to definitively assess particularly the logic and conditions of experimental design.

Focus on verification of measurement validity in the process of measurement.

Measurement is defined as the process of translating constructs into procedures for data collection.

Again, the constructs are manifested as the levels of measurement by being translated into variables via the measurement procedures – at the levels of nominal, ordinal, and interval/ratio. Here, the measurement procedures act very importantly because variables rely on them and because they affect the reliability of variables.

Validity of measurement is important in deciding whether the program evaluation and performance measurement system are appropriate. The validity is required in measurement method for program constructs, in selection of variables to measure constructs, and in relationship with other variables.

Much care must be exercised in selection and usage of data in measurement. When the management of outputs is the main concern, it is not easy to directly collect outcome related data, and proxy measurement that employs outputs as proxies also could be problematical in that it entails assumptions. In addition, there is a tendency to rely on already existing data in performance measurement system. Therefore, there is a need to focus on verification of validity of the measurement data and utilization of surveys as a key means of measurement.

Qualitative method based on constructive paradigm is an effective means of describing and conveying the results.

Qualitative method, unlike quantitative method, pursues things other than positivist approach method. The underlying base for this is the so called constructive paradigm that emphasizes the sense-making of human beings. Although debates on paradigm is continuing, methodological pluralism could be the current solution.

Although both methods is being used complementarily in evaluation regardless of whether which one is used ahead of the other, their differences are clearly distinguished. Qualitative method that pursues the things that are naturalistic is focused on seeking the answer to how social experiences are created and accepted as being meaningful. Qualitative method within the limitation of data utilization provides realizable and effective method in describing and conveying results.

Evaluator can construct conceptual frameworks that can lead evaluation. Credibility and generalizability of findings are the key tasks in executing this. Comparative measurement with other data and verification of coincidence through feedback from other provider are the means of enhancing the credibility and generalization.

How is qualitative method linked with performance measurement? Although the performance measurement has the tendency of relying on quantitative method, format in which stories on the performances is additionally verified is employed. Using cases and qualitative evidence can render findings more credible and useful. But precautions are required in order to be methodologically defensible as well as being persuasive.

In needs assessment for program, benchmarking through comparison is emphasized.

Needs assessment is one of core areas in the public sector. It can be carried out in diverse range of domains in the performance management cycle, and the needs here can be variable in reflecting our values and as support for particular programs and policies. That is, reflecting of needs onto program or policy presents the problem of having to accompany political choice. The concept of needs in our society has been changing and there are differences in the concepts of public needs. Such differences may arise from the insufficiency in the resources needed in describing problems.

Comparison concept is reflected onto all needs assessments. This signifies the comparison for measurement of scope and types of needs, and benchmarking is emphasized as the standard for measurement. The subjects of benchmarking are diverse. All the theories, models, frameworks, ethical values, service providers, and current and prospective clients can be the subjects.

Because the needs assessment, which is completed through various stages, can be contentious, format that can be methodologically defensible must be chosen. Survey is used frequently in order to obtain information from the current and prospective clients. Design and implementation of survey requires caution

in extracting issues. Bias in the responses is assessed as a significant problem. This can be mitigated through careful survey design and triangulation of survey results with other sources.

Economic evaluation requires caution on the validity of the assumptions and its need will continually increase.

Economic evaluation is based on the principles of welfare economics that consider the benefits and costs from societal perspectives. In general, cost-effectiveness analysis and cost-utility analysis are executed rather than cost-benefit analysis because it is difficult to convert the outcomes in the public sector into monetary values, and it can be obtained by focusing on finding the answer to how to accomplish the chosen outcomes.

These analytical methods have powerful appeal to the evaluators and policy decision makers because the most rational choice can be sought in providing information necessary in allocation of resources and decision-making particularly through decision-making on profitability (bottom line) by comparison of alternatives. There are limitations to be overcome here. Evaluation of the costs and practical benefits of a program is a difficult task. Since these in general cannot be evaluated directly, thereby frequently requiring assumptions and the validity problem of the assumptions may be presented.

In economic evaluation, the issue at hand is whether it is possible to predict the actual outcomes of program and policy alternatives being compared. In order for the outcomes of program to be measured or computed easily, certain program technologies[2] are necessary, and precautions are required in understanding the actual causal linkages and in validly measuring the inputs and outcomes.

Along with growing costs and demands for services, proving the expenditure in the public sector is being emphasized. Accordingly, the need for economic evaluation is also increasing. However, the quality of the studies thus far is evaluated to be insufficient.

---

2) It is defined as means-ends relationships that are used in programs to accomplish the program objectives.

Performance measurement requires approach from the viewpoint of new public management.

Historically, the performance measurement has been connected with the financial accountability. As the interest in the efficiency of public sector increased, the importance of performance measurement has also been receiving highlights. Succession of numerous cases - Programmed Planed Budget System(PPBS), Management By Objective(MBO), Zero-Based Budgeting(ZBB) – are a part of these procedures and is based on the cognizance that it is desirable to put focus on the results in performance measurement system.

In the 70's and the 80's, people became aware of the importance of balance of budget due to enormous financial losses and inflation they have experienced, and people, as tax-payers changed their viewpoints on the resources inputs in the public sector[3]. Such situations brought about changes in the viewpoints and realities on the management of public sector, which were expressed through emphasizing of downsizing, reduction of red tapes, efficiency and effectiveness.

Such concept of New Public Management exists as the governing concept in designing, execution and operation of government programs and services, and reflects the key metaphors of the public sector for performance measurement. That is, business-like practices are emphasized for the government, organizations are recognized from the perspectives of open-systems in which they biologically interact with the surrounding environment and assessed as machines consisted of complex systems.

Performance measurement being accomplished on such foundation exists as a part of performance management cycle along with program evaluation. These are mutually complementary in acquiring and analyzing intended information in order to reduce the uncertainty in program and policy decisions.

Key factors in elevating the possibility of success of performance measurement system – sustained leadership, communications, clear expectations for the system, sufficient resources, logic model and

measurement process.

With generalization of performance measurement on public sector, the sustainability of performance measurement system has become an important issue. The key factors in elevating the possibility of success in designing and implementing the performance measurement is presented below.

1) Sustained leadership. In particular, in the case of performance measurement at the government level, leadership is required at the 2 levels, namely high level officials of the departments who were appointed and the officials who were elected as the end users. 2) Communications. This increases the common understanding and participation on the performance measurement process. Diverse range of benefits of the multi-channel such as top-down, bottom-up and horizontal sharing of information must be sought after. 3) Clear expectations for the system. In general, performance measurement aims to improve program through provision of information on managers. Therefore, there is a need to be open and honest about the purposes in order for stakeholders and problems to be dealt with properly. 4) Sufficient resources. Under the situation of limitations on resources, performance measurement resources of other programs must be utilized assertively. 5) Logic model. This identifies the key programs and organizational constructs. Its process is important in choice of the constructs and performance measurement. 6) Measurement process. Valid measures that can entice the confidence of stakeholders must be produced.

Fulfillment of the presented criteria does not guarantee the success of definitive performance measurement. Performance measurement is also a craft and has considerable room for creativity and professional judgment.

Consider rational/technical factors as well as political/cultural factors, and reduce the possibility of gaming on performance measurement.

Although the performance measurement model is based on rational and technical processes, in reality, considerations for all political and cultural factors within the organization are required.

Performance measurement has always been at

---

3) Institutionally, the Proposition 13 in California in 1978 is considered as its starting point in the United States.

the center of the process of conversion into result-based performance management - The Government Performance and Results Act (GPRA), Program Assessment Rating Tool (PART), etc. As the value of performance measurement for budgetary and resource allocation is increasing, improving accountability for the program has become as important as the efficiency and effectiveness. However, there are several problems and issues in securing this. Among these, the complicated accountability and gaming on the performance measurement are situated at the center.

Under the democratic culture, there are two types of accountability model, namely, the hierarchical accountability and accountability of result. However, such accountability must consider the organic functions with other public organizations and private sector organizations as partners in accordance with complexity of accountability relationships, that is, dispersion of the accountability.

Performance is on target and gaming occurs when there are effects linked with this. Gaming that exists in the organizational life imparts influence on the validity of performance information and the behaviors of organizational participants (individual influence including managers and government official). Efforts are required to reduce such gaming. Multiple measures, each with its own validity and reliability profile can be means of reducing focal point on the gaming behaviors.

Cooperative combination with evaluation and management. Establish learning organization and develop professional judgment.

The purposes of evaluation impart influence on relationships between evaluators and managers. Formative evaluation focused on providing program-improvement information is recognized as being friendly to the managers, and enables win-win strategy between the two. On the other hand, the summative evaluation is quite different from above as it in general determines the future of programs. Therefore, ability to evaluate with considerations for both the rational/technical perspectives and political/cultural perspectives is important.

It is believed that the best evaluation can be made through cooperative combination with management rather than considering evaluation as an activity against management. Having evaluation as an external function can be an alternative in solving political problems from the perspective of the organization. In spite of this, the internal evaluation is a dominant execution method in public sector. However, the securing of reliability must be presumed.

Here, learning organization based on self-evaluation is emphasized. As the key methods, double-loop learning, that is, learning how to learn, and empowerment evaluation are presented. The former signifies behavioral correction to attain objectives and learning to assess the appropriateness of them, while the latter signifies that managers improve practice and foster self-determinant through evaluating their own programs.

How should the relationship between the evaluators and managers be formed? Excessive approach to the managers can hinder the objectivity of the evaluation process. In particular, it is more so for the objectivity of evaluation findings and conclusions. The objectivity coincides with the repeatability through repetition. However, in reality, repetitive evaluation itself is not easy. Evaluation is a craft that combines methods and artistry, and professional judgment is important. Stakeholders must acquire confidence through securing of ethical value system, honesty, accuracy, fairness, impartiality, competence, high skill and credibility.

Professional judgment has integrated meaning in evaluation practice.

Professional judgment has integrated meaning in evaluation practice. It acts on overall aspects of evaluation ranging from evaluation questions, research design, conduct, information analysis and interpretation to conveyance of results. There is no methodology in evaluation without room for dispute. In addition, the practice of the evaluation is more diverse than normative approaches. Findings, conclusions and recommendations based on evidences and professional judgment can reduce the uncertainties in evaluation questions and relevant information. Such judgment is required not only of the evaluators but also of the managers.

Professional judgment can be categorized into technical judgment, procedural judgment, reflective judgment and deliberative judgment proposed by Fisher & Cole. Evaluator's values, beliefs and expectations create experiences by combining with shareable and practical knowledge, and such experiences form the professional judgment that influences the decision. Here, external environment is also a factor in determining the judgment.

To evaluate is to make judgment. It is possible to develop sound professional judgment through reflection on one's practice. This is based on evaluation competencies, that is, the professionalism, and implies core body of knowledge. However, the balancing is needed in application of theoretical knowledge and practical know-how. One's experience is essential in completely integrating the impersonal knowledge into professional knowledge by combining with tacit knowledge. Professional judgment substantially depends on developing and practicing evaluation craft. Training aimed at understanding their roles and stages of formation, and conscious development of such are needed on the basis of awareness of its importance.

In conclusion of the review

This book is meaningful in further progressing the theories and practices in the evaluation area, and contains notable assertions and grounds for argument. In particular, it is possible to see that the authors approached with interest in grafting the theories onto the realities, which appears to the reflection of knowledge experienced by the both authors in theory and practice.

Their assertions are based on numerous preceding researches and evaluation cases. Nonetheless, it can be seen that they are putting more weight on the practice than on the evaluation theories. Although they are presenting the importance of the issues that can arise in evaluation practices, in particular, political views and professional judgment, they leave the key judgment to the readers. In developing the contents, there were some confusion in dealing with the programs, policies and organizations. This could perhaps be understood as the limitations in comprehensively dealing with the government and public sector.

I feel that readers who are wishing to or is already working in the area of evaluation or performance measurement would be able to acquire indirect experiences through this book. I have no doubt that this book will be helpful in recognizing and establishment of judgment that can cope with political issues that one would confront ceaselessly throughout the process of evaluation from research design to selecting methodology, executing evaluation and deduction of results.

Under the reality in which discussions and debates on the reinforcement of accountability associated budget along with expansion of program, I hope that readers would be able to acquire more prudent interest and intellectual stimulation on evaluation and performance measurement.

Reviewed by Young-Soo Ryu
Center for Public Administration and Policy
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061-0002, USA
E-mail: soory@kistep.re.kr