

日本評価研究

The Japanese Journal of Evaluation Studies

Vol. 2, No. 1, June 2002

研究論文

James W. Altschuld

The Preparation of Professional Evaluators:
Past Tense and Future Perfect

三好 皓一

プログラム・セオリー・マトリックスの活用について
- 評価における分析の有用性を高めるために -

研究ノート

小野 達也

地方自治体の行政評価システムの課題と展望

本間 まり子

ジェンダー視点に立ったJICAプロジェクトの評価
- 東アフリカの事例から -

盛 信博 青津 暢

開発援助の戦略的評価に関する一考察

実践・調査報告

Charles Gene Lyle

Reflections on a Career in Evaluation

委員会活動

編集委員会

研修委員会

国際交流委員会

その他

日本評価学会第3回全国大会について
評価関連ウェブサイトリスト

日本評価学会

The Japan Evaluation Society

『日本評価研究』編集委員会

Editorial Board

編集委員長
Editor-in-chief

三好 皓一（国際協力事業団）
Koichi MIYOSHI

副委員長
Vice-Editors-in-chief

古川 俊一（筑波大学）
Shun'ichi FURUKAWA

村松 安子（東京女子大学）
Yasuko MURAMATSU

常任編集委員
Standing Editors

牟田 博光（東京工業大学）
Hiromitsu MUTA

山谷 清志（岩手県立大学）
Kiyoshi YAMAYA

編集委員
Editors

石原 俊彦（関西学院大学）
Toshihiko ISHIHARA

大沢 真理（東京大学）
Mari OSAWA

賀来 公寛（東洋大学）
Kimihiro KAKU

河野 善彦（国際協力銀行）
Yoshihiko KONO

佐々木 亮（国際開発センター）
Ryo SASAKI

田中 弥生（国際協力銀行）
Yayoi TANAKA

長尾 眞文（広島大学）
Masafumi NAGAO

西野 桂子（グローバル・リンク・マネージメント）
Keiko NISHINO

松岡 俊二（広島大学）
Shunji MATSUOKA

森 茂子（日本大学）
Sigeko MORI

事務局
Office

〒135-0047 東京都江東区富岡2-9-11 京福ビル
財団法人国際開発センター 企画広報室内
日本評価学会事務局
TEL：03-3630-6994，FAX：03-3630-8120
E-mail:suzuki.n@idcj.or.jp

日本評価研究

第2巻 第1号 2002年6月

目 次

研究論文

James W. Altschuld

The Preparation of Professional Evaluators: Past Tense and Future Perfect..... 1

三好 皓一

プログラム・セオリー・マトリックスの活用について

評価における分析の有用性を高めるために11

研究ノート

小野 達也

地方自治体の行政評価システムの課題と展望.....29

本間 まり子

ジェンダー視点に立ったJICAプロジェクトの評価 東アフリカの事例から39

盛 信博 青津 暢

開発援助の戦略的評価に関する一考察.....55

実践・調査報告

Charles Gene Lyle

Reflections on a Career in Evaluation.....63

委員会活動

編集委員会.....69

研修委員会.....71

国際交流委員会.....72

その他

日本評価学会第3回全国大会について.....73

評価関連ウェブサイトリスト.....74

日本評価学会規約.....79

日本評価学会誌刊行規定.....84

『日本評価研究』投稿規定.....86

日本評価学会

The Japan Evaluation Society

【研究論文：依頼原稿】

The Preparation of Professional Evaluators: Past Tense and Future Perfect¹

James W. Altschuld
The Ohio State University
jaltschu@galileo.enc.org

Abstract

Although methodology (statistics and measurement) is an important component of evaluation, it is only one part of a much more complex enterprise. By inference, educational programs that only focus on methodology are not considered to be evaluation training programs. With that as an initial premise, the following topics are described in this paper: a brief, historical perspective of evaluation training including how it is evolving; a summary of the main findings from the most recent (1994) study of the results of evaluation training in the United States, Canada, and Australia; the current, in-process worldwide study of evaluation training; and trends that are projected to emerge with regard to training in the broader context of a field undergoing rapid growth and globalization.

Keywords

training, evaluator, curricula, evaluation, evaluation training,
evaluator training, evaluation curricula

Introduction

This seemingly simple topic, the preparation of professional evaluators, is neither well understood in the United States nor throughout the world. It would appear to be straightforward - here's what prospective evaluators need to know and to be able to do and here's the training that should enable them to perform as evaluators. This premise is easy to grasp and should be equally easy to implement. This is, however, not reality.

A brief anecdote demonstrates one aspect of the complexity of the issue. About 8 years ago, I attended a panel presentation at the annual conference of the American Evaluation Association (AEA). The panel consisted of recent or soon-to-be completers of graduate training programs in evaluation who were just starting their careers in evaluation. They were offering their perspectives on graduate education, the quality of the training they had received, how well they thought they

were prepared for conducting evaluations, and what they perceived they would be doing in the future. One of the presenters briefly described his program of study (measurement and statistics were highlighted extensively and exclusively with no other types of courses mentioned). After the panel presentations, I asked several questions of this individual (in a non-threatening manner) that went something like this:

Why would you consider these courses to be evaluation training?

Why would you not just call this a methodology curriculum?

What is evaluation curriculum, what is unique about it, and what are its distinguishing characteristics?

His honest reply was "I hadn't thought of curriculum in that way and I guess there is nothing in what I had that would make it stand out as an evaluation curriculum or program."

Like the background of this student, it would never be argued that methodology isn't a necessary part of evaluation training. Evaluators need a diversity of methods to engage in their craft. At The Ohio State University, all of my students must take qualitative, quantitative, and measurement courses in their graduate programs. Conversely, if that was the extent of their training then as the above individual answered they also would not have experienced an evaluation program. Methodology by itself is insufficient to constitute an evaluation program.

Intuitively from the sense that we have of ourselves and of our work, we know that evaluation is different from research and the study of methodology for the sake of methodology alone. It is this sense that motivates us to become members of organizations that are separate, stand-alone societies devoted solely to evaluation. While some research associations have evaluation divisions, the rapid rise of AEA, the development of state and local affiliates of AEA in the United States, the emergence of national societies (The Japanese Evaluation Society is an illustration) and international regional associations collectively attest to the premise that evaluation is neither research nor methodology.

Even though the skills acquired from study in these areas are vital to what we do, they cannot and they do not fully represent evaluation. As evaluators we have to:

- deal with political situations (who will be affected by the evaluation and in what ways will they be affected);
- consider and include multiple stakeholder groups in evaluation planning and implementation;
- take into account the context, the milieu in which programs and projects are implemented;
- uncover and work with logic models that depict the intricacies of programs and projects;
- understand a variety of specialized methods such as needs assessment, process monitoring, and cost analysis procedures;
- incorporate multiple methods into evaluation designs; and
- learn how to report results so that they are attended to and lead to positive change and improvement.

These and other subjects are not traditionally taught in research and methodology course sequences. They are what Patton (1990) has referred to as the 'swag', the bag of skills of the evaluator. These are what we bring to bear on our work, they are the assets and distinguishing features of what evaluators do. They demonstrate or at least are suggestive of the proposition that evaluation is different from research and that evaluation represents a different perception of how to think about the world and problems confronting organizations and agencies.

Evaluation is a set of procedures and perspectives applied in practical, real-world situations to assist decision-makers and decision-making audiences to judge the merit and worth of programs, plans, and processes. Worthen (1994) viewed it as a 'near field' or sub-discipline. He identified 9 characteristics of a profession and found that evaluation measured up quite well on most of them. Similarly, Scriven (1994) described evaluation as a 'trans discipline', i.e., a field like statistics that is taught in many areas and tends to cut across them. In other words, to these prominent evaluators, evaluation is either a separate area of study or will become so in the short term.

Returning to the original focus of this paper, what do we actually know about the training of evaluators? How do past and current training efforts fit with developing views of what the field is and will become? Was the graduate student wrong in his description of his program? Was his program of study too deficient in regard to creating and implanting the broader dimensions of a field beyond the narrow focus of methodology? Is there agreement on what evaluation is, especially in relation to the training of evaluators? Can or should there be agreement as to what would be a standard set of content and instructional approaches for training evaluators?

To look at some of these concerns the ensuing text is organized into the following sections: a synopsis of the recent history of training; a brief discussion of studies of evaluation training with an emphasis on the most recent evaluation training directory study (Altschuld, Engle, Cullen, Kim, & Macce, 1994); a description of a new study that is in process at this time (Engle & Altschuld); and a brief examination of a few issues that influence the training of evaluators.

The Recent History of Evaluation Training

The history of evaluation, depending on your disposition, can be attributed to either the ancient Chinese (Worthen, Sanders, & Fitzpatrick, 1997) or to the Jews. (See the Hebrew Bible and the Book of Daniel). In the 20th Century, one legislative act in the United States propelled the whole field into a fast forward mode. In 1965 the federal government enacted the Elementary and Secondary Education Act which pumped billions of dollars into education particularly targeted toward programs for disadvantaged students. As indicated in Table 1, an adage of evaluation is that when public dollars increase so will the demands for accountability. With few individuals trained as evaluators, a vacuum was created that became the impetus for the development of a field and training programs for practitioners.

Clearly, ESEA was a major factor in shaping modern program evaluation. Through the mid to late 1970s numerous programs were catalogued as teaching evaluation (primarily educational evaluation, but not limited solely to education). On the other hand, the programs were heavily rooted in methodology and generally had only a single course that had evaluation specific content (Davis, 1986). The methodological focus of a program as depicted by the graduate student in the anecdote was accurate. Whereas perhaps a 100 or more programs had been detected in the 1970s, the number had dwindled to 46 in the United States by the mid-1980s. They were included in the 1986 Directory of Evaluation Training Programs (May, Fleischer, Scheirer, & Cox, 1986). For earlier directory studies the reader is referred to Gephart and Potter (1976) and Conner, Clay, and Hill, 1980.

By 1992-1993, when the next directory study was completed, a smaller set ($n = 37$) of university- based evaluation programs had survived in the U.S. It should be noted that this study utilized a restrictive, but undoubtedly a more realistic definition of what constituted an evaluation

Table 1 The Four Phases of Evaluation Training (1965-2001)

Time Period	Nature of the Period	Comments
1965 - Late 1970s	Funding of ESEA led to the development of many evaluation training efforts in response to the increased demand for evaluators in order to meet the needs of the legislation.	Training programs were concentrated in, but not limited, to Colleges of Education. Education is the single largest industry in the United States so what occurs in it will also affect other fields (psychology, public policy and administration, etc.).
Late 1970s - Through the late 1980s	Federal funding(ESEA) diminished but money from other sources kept evaluation alive. Training programs as ascertained in late 1970s in accord with a rather broad definition began to consolidate into a smaller number of programs by the mid to late 1980s.	State and local sources of funds as well as support from Business and Industry provided a baseline for evaluation activity. As anticipated, the flurry of evaluation training engendered by ESEA waned with the reduction in funds. By 1986, 46 training programs were identified. (See the training directory studies cited in the text.)
Mid to late 1980s - Mid 1990s	Consolidation of training programs continues. Based upon more stringent criteria for what constitutes a program, 49 programs were identified for an updated directory. Eleven of them, however, were in Canada and Australia with one other located in a large U.S. agency.	Programs in the U.S. were fewer in number, but a subset of them were of considerable size. More programs were in other countries and the most extensive one was notably at a federal agency not in a university setting.
Mid 1990s - Present	Change in evaluation training will become apparent with the emergence of the web and Internet (EVALTALK), specialized training institutes, the globalization of evaluation, and training offered under the aegis of many evaluation groups and societies.	Any study of evaluation training will have to take an international perspective if it is to meaningfully characterize the preparation of evaluators. Globalization and almost instantaneous communication are radically and rapidly altering the evaluation training picture.

program. A program was defined as having two more courses focusing on evaluation topics such as evaluation models and theories, seminars on research into evaluation, specialized evaluation methods (needs assessment, cost-benefit analysis, etc.), reviews of completed evaluations, evaluation planning, the use of mixed methods, and so forth. Even with this restriction, the methodological emphasis of the 'evaluation programs' was still very much apparent. Methodology was dominant in the surveys that were completed and returned for the study.

The Results of the 1992-1993 Study

In 1992, Altschuld and Engle were appointed by AEA to co-direct a study of evaluation training in the U.S. The study was subsequently expanded to include Canada and Australia. By various means, the staff of the project were able to solicit the nomination of 93 potential training programs in universities essentially at the graduate level. Undergraduate training which had been observed in some of the earlier directory studies was no longer in existence.

Before summarizing highlights of the study, it is important to state that there are many single evaluation courses taught in American universities and elsewhere. Although they serve a valuable role in orienting individuals to evaluation (Morris, 1994), they are singular entities and cannot be thought of as being evaluation programs designed for the training of professional evaluators. Their purposes and foci are different and one course would not a program make!

As explained in Table 1, responses to the mail survey used in the study revealed that 49 programs were in existence (two university programs were missed - one in the U.S. and one in

Canada). Thirty-eight of the entries in the final directory that was generated from the survey results, came from the U.S. This number represented an 18% (15% drop if the missed program is taken into consideration) in programs in the country (Altschuld, Engle, et al. 1994). Interestingly and surprisingly, slightly more than half of all the identified programs had the word, evaluation, in their titles and, therefore, could be easily located by individuals seeking formal training. Prospective trainees would have had to search more diligently for the other. Twenty-two percent of the programs were outside of the United States.

Most of the training efforts were quite small in nature and were often staffed by less than one full-time equivalent (FTE) faculty member. When conducting the survey, it was observed that, in several instances, when a critical individual i.e., the mainstay of the program, retired or left the educational institution, the program simply disappeared. Thus, even years ago, the fragile nature of some evaluation training endeavors was a concern. But we should not lose sight of the fact that there were strong programs (approximately 15 -18 in number) that were thriving and educating a steady stream of evaluators who entered the profession and continued their membership in evaluation organizations.

With regard to evaluation specific content, the situation had improved and progress was being made, but it was far from ideal. Most programs had an introductory evaluation theory and models type of course and 40% did have a noticeable amount of other evaluation-specific courses. Yet when all the courses offered were examined, methodology still predominated. About forty percent of the courses described by survey respondents represented traditional (statistics and measurement) methodology areas. Needs assessment and cost-benefit analysis, methods which could be viewed as more closely connected to the work of evaluators comprised about 2% of the total of all courses listed in survey responses. As an overall judgment one could say that the results were positive and promising, and conversely that there still was a critical need for more explicit evaluation curricula.

The Current Directory Study (2001-2002)

Engle and Altschuld are co-directing the current, AEA-sponsored directory study which has expanded in scope from the previous ones. This expansion is necessitated by the dramatic changes taking place in the field during the past 10 years. They are given below.

First, there are now many more evaluation associations besides AEA both in the United States and elsewhere. AEA has seen a major increase in the number of formally affiliated state and local groups. Training opportunities are now more abundant and, at times, more available locally due to the involvement of multiple organizations across the nation.

Second, other types of training via intensive educational experiences such as those provided by Western Michigan University and the Evaluator's Institute have been developed and well received by the evaluation community. Little of this kind of training was in existence or known of at the time of the most recent, previous directory study (1992-1993). These efforts have been popular and have attracted outside financial support (The National Science Foundation, NSF). Because of their nature and intensity they are a form of training not easily duplicated at most universities due to constrained financial support and the limitation of having, at best, a small number of appropriate staff.

Third, and most noteworthy, evaluation has truly become a global enterprise. In the

United States, at one time in the distant past, we had an almost xenophobic perception of our role and importance to the field. We seemed to feel that evaluation was centered in the U.S. That time has thankfully passed and evaluation is now found in many areas of the world. We are ecstatic about this fact, enlightened by it, appreciative of our international colleagues, and welcome their involvement in the family of evaluators. Indeed, the 2000 conference of AEA held in Honolulu, Hawaii was remarkable for the number of international attendees and their contributions to the success of the conference activities and agenda.

There are societies across country boundaries and newly formed societies within countries (e.g., Asia and Japan, respectively). This is an exciting new development and it means that any study of evaluation training must account for opportunities throughout the world in universities as well coming through the auspices of professional groups.

Fourth, instant communication across the world due to the Internet and the web has already begun to change understandings of evaluation and evaluation training. This writer has been invited abroad on three separate occasions in the last two years. In each instance, almost all communication in relation to what he would be teaching or presenting has occurred via e-mail and almost all arrangements were made via the computer and its communication capabilities. Handouts and PowerPoint presentations were sent as attachments so that the carrying of needless papers through crowded airports was avoided.

EVALTALK has changed the way we do business. Evaluators freely and openly exchange ideas on the listserv and frequently ask questions about evaluation problems they are encountering. The listserv is a form of education.

Even though training is not now taking place in a formal fashion via distance education (utilizing the power of our communication tools), it is not only conceivable, but likely that it will emerge in the not too distant future. This will become a consideration for subsequent investigations of evaluation training.

Fifth, if one looks at the AEA-sponsored training sessions at the annual conference, it is obvious that the role of training for the association, aside from the professional discussion of ideas and the presentation of papers, is very large occupying as much as 30-40% of the meeting time. The national role in training has become and will remain quite large.

Sixth, an issue that has moved to the forefront of evaluation training, particularly in the United States, is that of the number of individuals from minority groups attracted to and working in the field of evaluation. Many programs in education and in social service, as examples, are aimed toward disadvantaged students or disadvantaged groups in urban areas. Is the profession attractive enough to entice students of color to work in it and to participate in evaluations of these programs? What kinds of training opportunities are available to these students in major universities and additionally in institutions that primarily serve minority populations? How do we train evaluators in general to be culturally and ethnically sensitive?

The current directory study, obviously could not proceed along the lines of earlier efforts. Too much has changed. The study will utilize three distinct questionnaires which will be distributed through and completed over the Internet. (The study is being funded by NSF. As indicated before, it will be co-directed by Engle and Altschuld and, in turn, implemented by one of NSF's contracted service providers.)

One questionnaire will go to a list of worldwide, university-based programs (including those with large numbers of minority students in the U. S.) that have been nominated for potential inclusion in an electronic directory of same. Who are the students?, is there a specialized content emphasis in the program?, what are the courses?, etc. are typical of the questions posed to the universities. The directory generated from the responses will be accessible via the AEA website. At this early point in the study it is apparent that numerous international training programs will be part of the new directory.

The second questionnaire will be carried out in a similar manner across the world, but it would be oriented toward training offered: under the guidance of professional organizations; through specialized regularly available institutes; and through governmental agencies. These types of programs would also be listed on the website, but under a different heading than university-based training. The number of such programs will be substantial and the directory will be the first time that this type of information has been systematically catalogued for evaluators.

The third questionnaire is more in the way of a research study. Its intent is to identify courses at institutions that historically serve minority populations. The key question is are students of color being exposed to the field of evaluation and its conceptual framework. This study will help AEA and other groups to form policies that will hopefully broaden participation in evaluation.

What are Issues that Should be Considered in the Training of Evaluators?

The subtitle of this article mentions the past tense and the future perfect. The past has been covered in previous sections. What of the future perfect? It's the tense that deals with what will have taken place or maybe it's more of a subjunctive idea of what should have or should be taking place in evaluation training?

The definitions of evaluation and their implications are important for the training of evaluators. A full examination of them is beyond the scope of this discussion. Let a simple observation suffice. Dimensions of evaluation such as the: nature of decision-making, values of stakeholders and their involvement, types of programs to be evaluated, timing of the evaluation (when in the life of a program it occurs), methodological possibilities for evaluating programs and projects, politics of evaluation, and multiple types of evaluation activities (theory development, needs assessment, accountability systems, and so forth) are complex. That complexity explains why it is so hard for evaluators to agree about the content of and the instructional strategies to be used in an evaluation training curriculum.

We all have our independent views and perspectives on what models to use, about what evaluators do, and about the kinds of knowledge they should have. At times, these views are tempered by the idiosyncracies of the specific subject matter fields (education, health, mental health, social service, government, business and industry) in which we work and in which we conduct evaluations.

One area in which we have not yet attained the status of a profession (Worthen, 1994; Altschuld, 1999a; Altschuld; 1999b) is in regard to the certification of evaluators. A logical inference from not controlling entry into the field in this manner is that **we have not and are reluctant to specify the basic content, knowledge, and skills that evaluators should have.** This is the **most prominent and persistent issue** that we face in the training of evaluators and one that we must eventually address.

The Canadian Evaluation Society (Love, 1994), has tried to deal with the issue by proposing a credentialing system that would require evaluators to have exposure to four areas of content (general understandings of evaluation, needs assessment, monitoring program performance, and evaluating results). The idea is to credential (guarantee exposure to the areas), but not to certify (guarantee a certain level of skill in each area). Whether we agree or disagree with the Canadians, it is clear that we will have to attend to what it means to be an evaluator and what training for evaluation would entail. We must expand the debate and discussion of this issue.

The **second issue pertains to resources for evaluation training**, especially at the university level. The absolute number of university programs in the new directory will probably be larger than in the 1994 one, but primarily due to its international scope (programs from many countries will now be included). Will programs with a small FTE-base of support be able to continue if a key staff member leaves or if there is a negative change in administrative support. If we could better establish what a solid, fundamental program looks like and what resources are needed for implementation and continuance, then we would have a stronger base for promoting evaluation training.

The **third and last issue relates to the balance of training between universities and the emergence of other training opportunities**. What is the appropriate role for university-based training? Should university-based programs serve as the entry mechanism for the field? Would and should some aspects of training at the university level be for quite different purposes (to conduct research on evaluation and to teach evaluation) than other types? Would such purposes be evident in other training venues?

What role should the other training opportunities play in expanding general and specialized evaluation understandings and skills? What is the articulation between these two types (university and professional development) of training? What about the continuing educational needs of evaluators for maintaining skills? How will continuing education for evaluators be handled especially if continuing education credits are to be given and universally recognized? Who will cover the costs of continuing education? Could or should alternate forms of training evaluators eventually replace current ones which still reside within universities?

Concluding Remark

Many subtle undercurrents swirl around how we do train and should train evaluators. Hopefully, the reader is now disposed to raise even more questions about such training and to enter into the dialogue necessary for developing the programs of the future. If that outcome has been achieved, then this paper will have enhanced the debate about the preparation of professional evaluators and it will have met its intended purpose.

Notes

1 Partially based on *Evaluation Curricula and the Preparation of Professional Evaluators*, a presentation by James W. Altschuld to The Japanese Evaluation Society on September 9, 2001, Tokyo.

References

Altschuld, J. W. (2001, September 9). *Evaluation curricula and the preparation of professional evaluators*. A presentation to The Japanese Evaluation Society, Tokyo.

- Altschuld, J. W., Engle, M., Cullen, C., Kim, I., & Macce, B. R. (1994). The 1994 directory of evaluation training programs. In J. W. Altschuld and M. Engle (Eds.), *The Preparation of Professional Evaluators: Issues, Perspectives, and Programs*. New Directions for Program Evaluation, 62, 71-94.
- Altschuld, J. W. (1999a). The certification of evaluators: Highlights from a report to the Board of Directors of the American Evaluation Association. *American Journal of Evaluation*, 20(3), 481-493.
- Altschuld, J. W. (1999b). The case for a voluntary system for credentialing evaluators. *American Journal of Evaluation*, 20(3), 507-517.
- Connor, R. F., Clay, T., & Hill, P. (1980). *Directory of evaluation training*. Washington, D.C.: Pintail Press.
- Davis, B. G. (1986). Overview of teaching evaluation across the disciplines. In B. G. Davis (Ed.), *Teaching of Evaluation Across the Disciplines*. New Directions for Program Evaluation, 29, 5-14.
- Gephart, W. J., & Potter, W. J. (1976). *Evaluation training catalog*. Bloomington, IN: Phi Delta Kappa.
- Love, A. J. (1994). Should evaluators be certified? In J. W. Altschuld and M. Engle (Eds.), *The Preparation of Professional Evaluators: Issues, Perspectives, and Programs*. New Directions for Program Evaluation, 62, 29-40.
- Morris, M. (1994). The role of single evaluation courses in evaluation training. In J. W. Altschuld and M. Engle (Eds.), *The Preparation of Professional Evaluators: Issues, Perspectives, and Programs*. New Directions for Program Evaluation, 62, 51-60.
- May, R. M., Fleischer, M., Scheirer, C. J., & Cox, G. B. (1986). Directory of evaluation training programs. In B. G. Davis (Ed.), *Teaching of Evaluation Across the Disciplines*. New Directions for Program Evaluation, 29, 71-98.
- Patton, M. Q. (1990). *The challenge of being a profession*. *Evaluation Practice*, 11, 45-51.
- Scriven, M. (1994). *The final synthesis*. *Evaluation Practice*, 15, 367-382.
- Worthen, B. R. (1994). Is evaluation a mature enough profession that warrants the preparation of evaluation professionals. In J. W. Altschuld and M. Engle (Eds.), *The Preparation of Professional Evaluators: Issues, Perspectives, and Programs*. New Directions for Program Evaluation, 62, 3-16.
- Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines*. White Plains, NY: Addison, Wesley, Longman, Inc.

(2001.9.9受理)

【研究論文】

プログラム・セオリー・マトリックスの活用について 評価における分析の有用性を高めるために

三好 皓一
国際協力事業団

Miyoshi.Koichi@jica.go.jp

要 約

本論は、評価分析の有用性を高めるためにプログラム・セオリー・マトリックス（PTM）の概念と評価における活用について考察する。PTMは、プロジェクト内容と評価5項目等の評価分析結果を同一のマトリックス内に表示し分析枠組みを提示することによって評価の分析を深化させる。具体的には、開発援助の各々の事業を統合して一連の政策や施策の下に位置付け評価を行うための、また、個々の評価結果を知識として蓄積し一般化・概念化し、より有効な教訓や提言を導き出すための、有用な手法としてPTMの活用を検討する。

キーワード

プロジェクト・デザイン・マトリックス（PDM）
プログラム・セオリー・マトリックス（PTM） 政策評価、プログラム評価、メタ評価

1. はじめに

開発援助の評価においては、ロジカル・フレーム・ワークの一環であるプロジェクト・デザイン・マトリックス（PDM）が広く使われている。PDMは、プロジェクトの「投入（Input）」、「活動（Activities）」、「成果（Output）」、「プロジェクト目標（Project Purpose）」、「上位目標（Overall Goal）」へと連なる一連の因果関係、すなわちプログラム・セオリーを明確に提示する。PDMによる評価は、評価用のPDMの作成、評価グリッドの作成、実績・実施プロセスの確認、評価5項目による価値判断と効果発現・阻害要因

の分析からなる。評価は、DACの評価5項目である 妥当性（relevance） 効率性（efficiency） 有効性（effectiveness） インパクト（impact） 自立発展性（sustainability）などを評価の設問として、PDMを基に評価グリッドを作成し分析することになる。PDMは、歴史的には、開発協力における個別事業に対するプログラム・セオリー評価を促進することに寄与してきた¹。

しかし、今日、評価では、個別の事業のみならず、政策（ポリシー）や施策（プログラム）に対する評価が強く求められている。2001年の省庁の再編に伴う政策評価の導入と、この政策評価の実施を実体面で裏付ける「行政

機関が行う政策の評価に関する法律（行政評価法）」が2002年4月より施行される。また、国際協力の世界では、開発協力が開発途上国の開発に効率的・効果的に行われているかが一層強く求められており、1996年5月に経済開発協力機構（OECD）の開発援助委員会（DAC）上級会合で採択された「21世紀に向けて：開発協力を通じた貢献」（通称「DAC新開発戦略」）は、先進国の開発努力の成果を評価するために、開発途上国における開発指標を選択し目標数値を定めている。この戦略の基で、援助国はもとより、開発途上国においては、一層の開発政策や開発戦略の明確化とその成果の達成が求められるようになってきている。最近の貧困削減戦略書（PRSP）やセクター・プログラムの展開は、このような背景を反映したものであり、評価は、個々の事業を超えて、政策、施策の評価が重要視されるようになってきている（三好2001a）。

政策や施策の評価は、ある特定分野や開発課題のもとに計画された、共通の目標を持つ複数の事業（プロジェクト）を包括的に取り上げて評価するものである。プログラム・レベルの評価は、基本的には政策、施策、事業よりなる一連の政策体系の評価であるが、一般に政策、施策、事業は混在しており、これらを何らかの操作可能な（明確な、具体的な目的と手段の体系によって他と区別された）仕事のかたまりとして捉えられるようになってはじめて意味のある評価が可能になる（古川2001）。特に、日本の開発援助は形態（スキーム）別に実施がなされており、この政策、施策、事業の混在が著しい。このため、政策（ポリシー）評価や施策（プログラム）評価を行うためには、各々の事業を統合して一連の政策や施策の下に位置付け評価することが必要であるが、従来このための明確な方法論を欠いていた。

また、今日では評価結果の活用も大きく変わってきている。評価結果の活用は、初期には手段としての活用、すなわち評価の結果を使って目に見える形で直接的な利害関係者である実施担当部局等の事業実施関係者に働き

かける行動として捉えられた。どのように教訓を利用したか、どのように提言に従ったか、がここでは問題にされてきた。しかし、このような評価の活用についての狭い見方は現在では急速に改められてきている。今日では、評価対象事業での直接的な明白な行動の変化を期待するだけでなく、評価結果に対する利害関係者の理解を深めることで事業に関わる課題や問題を認識し、同じような事業に対して効果的な対策や対応が取れるようにすることが期待されている。そのためには、評価の教訓や提言の一般化、概念化が強く求められている（三好2001a）。

では、どのようにこのような動きに対応していくのか。筆者は、個別事業を対象としたPDMによる評価ではこれに対応することは難しいと考えている。本論は、このようなPDMによる評価手法の欠点を補い、評価における分析の有用性を高めるために、プログラム・セオリー・マトリックス（PTM）の概念と評価における活用の有効性について議論する。PTMは、個別のプロジェクト評価の分析で有効であることに加え、個々の評価結果を総合的に分析するに際し、PDMでは提供できない、より実践的な教訓・提言を提示するための分析枠組みを提供する。PTMを使つての個別事業に対する評価については先行研究は多い（Funnell 2000）。本論では、PTMの活用について、個別事業に加え、複数事業に焦点を当て、その活用の方向性について考察する。具体的には、開発援助の各々の事業を統合して一連の政策や施策の下に位置付け評価を行うための、また、個々の評価結果を知識として蓄積し一般化・概念化し、より有効な教訓や提言を導き出すための、有用な手法としてPTMの活用を検討する。プログラム評価とメタ評価についての手法について考察するとともに、業績評価の指標の適正化、ベンチマーキングの確保、内部評価・自己評価の強化について、その有効性について示唆を得る。

他方、PDMでの評価では、評価グリッドは、PDMとは別の独立した表として作成さ

れる。評価分析結果は、PDMのマトリックス内に同時に掲載されることはなく、PDMは独立した単独の基本的な表として示される。このため、評価をPDMとを密接に関連付けられず、各分析の深度には不均衡が生じ、適切なデータや情報が作成されない可能性を高くしている。結果として、評価5項目、因果関係を含む一連の評価が適切に実施されない可能性が高くなる。プログラム・セオリー評価の基本を理解していない人が、PDMを基に評価を行う場合には、投入・活動、成果に焦点が当てられ情報収集が行われ、プロジェクト目標や上位目標がおろそかになる、評価時点での現状とあるべき姿との比較が、主に投入・活動、成果を中心に行われる、投入・活動、成果、プロジェクト目標、上位目標と評価5項目との関連を適切に分析し得ない、結果として因果関係を明確化し得ないことが多い(Funnell 2000、三好・源 2001)。本論では、この点についても、PTMの概念を整理するために考察する。

なお、本論では開発援助におけるPTMの活用を考察するが、基本的には政策、施策、事業からなる一連の政策体系の評価を対象としており、ここでの考察結果は日本国内の事業の評価に適用することができるものである。PTMの活用は、行政の政策、施策、事業を整理し、評価実施可能性を高めるために有効であり、特に、地方自治体と中央官庁との事業が混在する事例等その適用範囲は広い。

本論の構成は、次節でPDMによる評価を、3節でPDMの制約とPTMの基本的概念を整理する。第4節では、プログラム評価への適用を、第5節でメタ評価における適用を考察し、評価分析におけるPTMの有効性を例示する。これらの考察を踏まえ、第6節で結論を整理し今後の方向性を提示する。

2. PDMによる評価

(1) PDMとプログラム・セオリー

プロジェクト・デザイン・マトリックス(PDM)は、プロジェクトの構成要素である「投入」、「活動」、「成果」、「プロジェクト目標」、「上位目標」や、プロジェクトを取り巻く「外部条件(Important Assumptions)」の論理的な相関関係、すなわちプログラム・セオリーをマトリックスとして示したものである。

PDMでは、プロジェクト活動と外部条件が満たされた時にプロジェクトの成果が、成果と外部条件が満たされた時にプロジェクト目標が、そしてプロジェクト目標と外部条件が満たされた時に上位目標が達成されるという論理構成が成り立つ。プロジェクトは、投入を効果的に用いてプロジェクトが実施する活動により成果を得、成果がPDMの論理構成を通じてプロジェクト目標、上位目標が達成されることになる²。

このようなPDMの論理性は、投入、活動、成果、プロジェクト目標、上位目標へと連なる概念的連鎖を提示する。この概念的連鎖は、プロジェクトの理論的根拠、すなわち構成要素間の因果関係をプログラム・セオリーとして規定することになる。

(2) PDMによる評価

開発協力におけるプロジェクトでは、広くPDMを基に評価が実施される。国際協力事業団(JICA)に例をとれば、終了時評価、または、事後評価を実施するに当たっては、まず評価対象プロジェクトのPDMの存在を確認することから評価が開始されている。PDMが存在しない時には、評価に先立ち既存資料などを基にPDMが作成されることから、PDMが評価の基本とされていることが解かる。

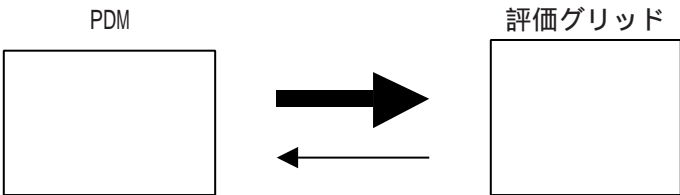
PDMによる評価では、PDMの論理性を基に、PDMの指標を用いてプロジェクトの各

表1 PDMの段階別目標とDAC評価5項目

評価項目	上位目標	プロジェクト目標	成果	投入／活動
妥当性	*****			
効果	*****			
有効性		*****		
効率性			*****	
自立発展性	*****			

（出所）『JICA事業評価ガイドライン』、ノルウェー外務省（1996）『開発援助の評価 評価のためのハンドブック』等を参考に作成。

図1 PDMと評価グリッド



（注）矢印の方向は上方の流れを、矢印の太さは情報の量を示す。
（出所）筆者作成。

レベルの目標達成度を検証し、その後に、DACの評価5項目である妥当性、効率性、有効性、インパクト、自立発展性により評価を実施することになる。PDMの段階別目標とDACの評価5項目の関係を示すと表1のようになり、この5項目を基本に評価を試みることになる。プログラム・セオリーを示すPDMの記載事項を基に評価の設問を設定し、一般的に、評価グリッドにより作業が行われ、評価グリッドに基づく調査結果を使い、評価5項目についての調査結果をまとめることになる。その後、横断的評価として因果関係を特定することになる³。

3．PDMの制約とPTM

(1) PDMによる評価の制約

ここでは、PDMの評価における制約を見ておきたい。PDMに基づく評価作業はPDM

と評価グリッドを主体とし行われるが、ここでの問題は、PDMは基本的な資料として評価分析とは別に存在し、PDMから導き出された設問に基づく各々の分析があたかもPDMとは独立したものであるかのように行われることにある（図1）。PDMの情報は、評価グリッド作成時にPDMから評価グリッドに流れるが、逆は少なく、評価は評価グリッドに基づき行われ、実際にはPDMと評価グリッド間の継続的な検証・考察が行われないことになる。それによって各分析の深度には不均衡が生じ、適切なデータや情報が作成されず、結果として評価5項目、因果関係を含む一連の評価が適切に実施されない可能性が高くなる。プログラム・セオリー評価の基本を理解していない人には、PDMを基に評価を実施するには次のような難しさが存在することになる。多くの評価で、投入・活動、成果に焦点が当てられ情報収集が行われ、プロジェクト目標や上位目標がおろそかになること、評価時点での現状とあるべき姿との比較が、

主に投入・活動、成果を中心に行われること、投入・活動、成果、プロジェクト目標、上位目標と評価5項目とも関連を適切に行い得ないこと、結果として因果関係を明確化し得ないことになる（Funnell 2000、三好・源2001）。

他方、このような単独プロジェクトにおける評価活動の分散化は、複数の評価結果を活用して総合的な分析への可能性をも狭めている。PDMの単票としての融通性のなさは、評価5項目と結合して分析を進展させる柔軟性がないばかりでなく、複数のPDMを結び付けて分析することの可能性をも排除している。結果として評価者の柔軟な思考を制限することになっている。

（2）プログラム・セオリー・マトリックス（PTM）

プログラム・セオリー・マトリックス（PTM）⁴は、基本的には上記表1の各項目を並列に並べ、PDMと評価5項目を同一のマトリックス内で検討分析するための分析道具である。図示すれば、表2のとおりになる。これは、形式的にはPDMを横に倒し、評価5項目による評価とを結び付けたものである。実態的には、プログラム・セオリーに基づく評価分析を同一枠組みで実施する仕組みを提供する。それぞれの欄に、各項目の関連を明記しておけば、より適切な分析に役立つ（網掛けの数字は例示）。分析に必要があれば、列や行の欄を増やすことも可能であり、また、例えば、欄が増えたとしてもコンピュータのソフトウェアの発達した今日では分析においても問題は生じない。

PTMは、従来からプログラム・セオリー

評価の実施に使用されており、特に前節で述べたPDMにおける制約を克服するのに有効とされる。PTMは、評価において、プロジェクトの構成要素である投入・活動、成果、プロジェクト目標、上位目標に対し均衡のとれた扱いを担保することになる、簡単に確保できるデータや情報に偏りがちな情報収集作業を、何が重要な情報かという思考へ変えるとともに、情報収集手法を明確化する、PTM自体に、評価項目や比較作業を内在させる、事業の意図した目標に対してプロジェクトと外部の要因の効果を明確化することによって、プロジェクトの因果関係を説明する組織的な基盤を提供することを可能にする（Funnell 2000）。

4．プログラム評価の試み

（1）プログラムとプロジェクト

表3は、PTMによる評価分析をプログラムに適用するための枠組みを提供するために、プロジェクトとプログラムの関係を示したものである⁵。個々のプロジェクトのPTMを統合して、プログラム評価手法の確立、メタ評価の実施促進、業績評価の指標の適正化、ベンチマーキングの確保、内部評価・自己評価の強化等が可能と考えるが（後述）するためには、基本となるプロジェクトとプログラムの関係についてPTM概念を使って整理しておくことが必要である。

プログラム・レベルの評価では、政策、施策、事業を一連の政策体系として何らかの操作可能な（明確な、具体的な目的と手段の体

表2 プログラム・セオリー・マトリックス（PTM）

上位目標	プロジェクト目標	成果	活動	投入	効率性 = /	有効性	インパクト	自立発展性	妥当性

（出所）筆者作成。

表3 プロジェクトとプログラムの関係

	上位目標 (Long-term Outcome)	プロジェクト目標 (Mid-term Outcome)	成果 (Output)	活動 (Activities)	投入 (Input)
ポリシー (政策)	P	MTOI/MTOII			
プログラム (施策) I	P=LTOI	MTOI	PA/PB		
プロジェクト (事業) A	LTOI=PA	MTOI=PA	PA	PA	PA
プロジェクト (事業) B	LTOI=PB	MTOI=PB	PB	PB	PB
プログラム (施策) II	P=LOTH	MTOII	PC/PD		
プロジェクト (事業) C	LTOII=PC	MTOII=PC	PC	PC	PC
プロジェクト (事業) D	LTOII=PD	MTOII=PD	PD	PD	PD

(注) = は、左右の事項が一致することを示す。

(出所) 筆者作成。

系によって他と区別された)事業のかたまりとして捉えることによって、はじめて意味のある評価が可能になる。日本の開発援助は形態別に実施がなされており、政策、施策、事業が混在しており、各々の事業を統合して一連の政策や施策の下に位置付け評価するための枠組みが必要である。

表3では、上位目標、プロジェクト目標、成果、活動、投入が、政策(ポリシー)、施策(プログラム)、事業(プロジェクト)において、どのように位置付けられるか示している。例えば、プログラムとプロジェクトの上位目標は、ポリシーとして示される政策目標と一致することになる。また、ポリシーは、プログラムIとIIの実施によって実現される。プログラムIIは、プロジェクトAとBの実施によって実現される。プログラムIIも同様にプロジェクトCとDの実施によって実現される。それぞれのプロジェクトはそれぞれの投入とそれを使っての活動によって成果が達成されて、プロジェクト目標、上位目標が達成されることになる。

プロジェクトでは、投入、活動、成果は操

作可能であるが、プロジェクト目標と上位目標の達成は、それぞれの関係を規定する外部条件によって定められることになり、プロジェクト関係者の力の及ぶものではない。しかし、プログラムIIは、プロジェクトAとBのプライオリティ付けや予算配分、プロジェクトを規定する法令策定などにより成果と上位目標の関連に影響を与えることができる。また、ポリシーでは、その策定にあたっては、政策目標の選択として上位目標の確定、プログラムのプライオリティ付けなどによりプロジェクト目標と上位目標の関連に影響を与えることができる。表3では、網掛け部分が操作可能なものを示している。

ここで特筆すべきことは、個々のプロジェクトの評価では、プロジェクトを実施することによって上位目標やプロジェクト目標の達成については分析できるが、そのあり方について分析はできない。しかし、個別プロジェクトのPTMを統合することによって、政策やプログラムのあり方についての分析の可能性を生み出す。表4は、表3のうち操作可能な網掛け部分を抜き出して作成したプログラム

表4 プログラム・デザイン・マトリックス

上位目標(Long-term Outcome)	プロジェクト目標(Mid-term Outcome)	成果 (Output)	活動 (Activities)	投入 (Input)
P	MTOI	PA	PA	PA
		PB	PB	PB
	MTOII	PC	PC	PC
		PD	PD	PD

(出所) 筆者作成。

・レベルのPDMである。ここでは、便宜上、プログラム・デザイン・マトリックスと呼んでおく。このような表を作成することによって、政策、施策、事業の手段と目的を持った一つの仕事のかたまりとして一連の政策体系を捉えることが可能となり、プログラム評価を可能にする。このプログラム・デザイン・マトリックスは、プログラムのロジックである政策体系・施策体系のプログラム・セオリーを提供することになり、これを基にPTMを作成し評価分析を行うことが可能になる。また、一連のプログラム・デザイン・マトリックスの作成過程を通じ、プロジェクト間の比較（プロジェクトのPTMの比較など）により、政策立案の適否、施策の整合性など多くの分析が可能になるとともに、プログラム評価の評価可能性についても検討が可能である。

(2) フィリピン人口家族計画評価

フィリピンの母子保健・家族計画（リプロダクティブヘルス）評価は、1992年4月から2000年2月の間に実施された日本の援助案件を対象として実施している。これらの援助は、共通の目的の下に、プロジェクト方式技術協力、現地国内研修、青年海外協力隊（個別派遣とグループ派遣）、人口家族計画特別機材、開発福祉支援事業、無償資金協力、草の根無償と、多種多様なスキームで、緩やかに連携の下に実施されている。この評価では、これらの案件を包括的プログラムと見なし、PTM

による評価を試みている。しかし、実際の援助事業では、それぞれのスキームによる援助活動については明確にPDMが作成されている訳ではなく、この評価でPTMを作成することにより各々の援助活動を統合して一連の政策、施策、事業として認識することを可能にした。このことにより、今後の評価の可能性を大きくするとともに、事業運営の明確化・改善に役立つ情報を提供している。

手法としては、全案件の概略をマトリックスで整理している。具体的には、プロジェクト内容（上位目標、プロジェクト目標、成果、活動、投入）を横に並列に並べ、これを「プログラムアプローチ・ロジック・モデル（PLM）1」してとりまとめている（一部を表5として示す）。次に各々のプロジェクト目標や成果のレベルを調整し（実際には、各項目におけるバラツキが大きい）、同じような成果と活動をまとめて、「プログラムアプローチ・ロジック・モデル（PLM）2」を作成している（一部を表6として示す）。ここで使ったPLMは、PTMの変形であり、プロジェクト毎のPDMでは表せない、それぞれのプロジェクトの活動を統合した援助活動を明確に提示しており、これをもとに個々のプロジェクトのPDMを使つての評価では対応することが難しい分析を可能にした⁶。また、作成過程において、各スキーム間の比較を行うことになり、個々の事業毎の評価を行うことでは特定できない一連の事業の実態を明らかにすることを可能にした。

表5 プログラムアップローチ・ロジックモデル (PLM1)

[illegible]

(出所) 国際協力事業団企画・評価部評価監理室 (2002a)

表6 プログラムアップローチ・ロジックモデル (PLM2)

Overall Goal	Program Purpose	Outputs	Outcomes	Name of Project	Activities	Inputs	備考
リージョン画におけるリレーション画における、RHが向上する。基礎的健康改善が向上する。	1. RH分野における、サービスデリバリーのための基盤が整備される。	1. タララ州における家庭サービス、母子保健サービス提供体制の改善。 2. 母子保健サービスの向上。 3. 地方自治体スタッフの保健プロダクトIIへの導入が向上する。 4. 人口・家族計画及び母子保健サービスの向上を促進する。	1. 家族計画・母子保健サービスの提供と供給体制プロジェクト基盤整備 2. 母子保健サービスの向上 3. 施設・医療資機材及びIEC資機材の整備。 4. 人口家族特別機材供与 (基礎診断機器、統計資料作成用機器、車両、TBA補助機材、血圧計、診療台、医薬品、バイク、油圧用器具、検査室)	1. 家族計画・母子保健サービスの提供と供給体制プロジェクト基盤整備 2. 母子保健サービスの向上 3. 施設・医療資機材及びIEC資機材の整備。 4. 人口家族特別機材供与 (基礎診断機器、統計資料作成用機器、車両、TBA補助機材、血圧計、診療台、医薬品、バイク、油圧用器具、検査室)	1. 家族計画・母子保健サービスの提供と供給体制プロジェクト基盤整備 2. 母子保健サービスの向上 3. 施設・医療資機材及びIEC資機材の整備。 4. 人口家族特別機材供与 (基礎診断機器、統計資料作成用機器、車両、TBA補助機材、血圧計、診療台、医薬品、バイク、油圧用器具、検査室)	1. 家族計画・母子保健サービスの提供と供給体制プロジェクト基盤整備 2. 母子保健サービスの向上 3. 施設・医療資機材及びIEC資機材の整備。 4. 人口家族特別機材供与 (基礎診断機器、統計資料作成用機器、車両、TBA補助機材、血圧計、診療台、医薬品、バイク、油圧用器具、検査室)	1. 家族計画・母子保健サービスの提供と供給体制プロジェクト基盤整備 2. 母子保健サービスの向上 3. 施設・医療資機材及びIEC資機材の整備。 4. 人口家族特別機材供与 (基礎診断機器、統計資料作成用機器、車両、TBA補助機材、血圧計、診療台、医薬品、バイク、油圧用器具、検査室)
現在の子供数 Pilot: 2.5人 Control: 3.2人							
望む子供の数 Pilot: 3.4人 Control: 3.7人							
難民実行率 Pilot: 69.8% Control: 66.7% *非妊婦中の139人(Pilot), 27人(Control)							
難民方法 Pilot: 多岐 (Injection:24%, Pili:22%) Control: Pili:55%							

(出所) 国際協力事業団企画・評価部評価監理室 (2002a)

この評価調査では、このPLM2を基に、計画達成度を把握し、5項目評価を行っている。このPLM2をさらに発展させ、指標、指標の入手手段、外部条件、評価5項目等の欄を設けて分析することも可能である。この調査では、一般的なりプロダクティブヘルスの指標のうち、既存資料とアンケート調査で入手できたものを記載し、分析を試みている。しかし、この評価でのPTMを使っでの分析では、試行錯誤を繰り返してのPLM2を作ることまでに多くの労力が費やされており、PLM2を十分に活用しての評価分析にまでは至っていない。

5. メタ評価PTMの活用

(1) メタ評価とPTM

メタ評価とは、評価手段、成果目標、前提条件等を特定したり、それらの相互関係を確認する指標を明示したりし、一連の評価結果を総合的に分析するものである。メタ評価は、過去に実施した評価の結果を知識として確実に蓄積し、評価結果に関心ある人々に伝える役割を果たし得る。

一般に、評価の教訓や提言は、評価対象となった事業の実施者等の直接事業にかかわっている利害関係者には直接的に関連するものであり、それなりの影響を与える。しかし、評価の教訓や提言は通常一般化、概念化されておらず、評価に興味がある者でも、その教訓や提言を実際の事業の実施や、また、政策決定に活用していくにはそれなりの知識を有

していることが求められることになるので、その活用には制約が出てくる。評価結果の活用は、特定の教訓や提言を理解する力に左右されることにもなる（Kirkhart 2000、三好2001a）。

メタ評価は、評価の教訓や提言をより概念化することにより、評価の影響を大きくし、永続性を持たせるものとする。一つ一つの評価結果では、活用の源泉も限られるが、メタ評価を実施することによって、より広い範囲から評価結果を概念化・一般化でき、事業にかかわる課題や問題の認識を深めることや、特定の考え方や事業等を進めることができる。表7は、メタ評価におけるPTMの活用を示している。A、B、Cは、プロジェクト（またはプログラム）を表しており、それぞれのPTMを並べることにより、プロジェクト（またはプログラム）間の比較検討を行うことができる。

日本では、メタ評価の例はほとんどないが、本論で例示する南南協力評価ではPTMを活用しメタ評価の手法を一部取り入れている。この評価調査では、既存の終了時評価報告書を含む評価結果をベースにPTMを作成しメタ評価を行っている。

(2) 南南協力評価におけるPTMの活用

この評価調査は、南南協力の概念と現状の比較とともにシンガポールとタイをケース・スタディとして南南協力の視点から評価を行い、今後のJICAの南南協力支援の効果的な戦略と実施上の留意点などについて提言を行うことを目的としている。この評価調査では、

表7 PTMを活用したメタ評価

	上位目標	プロジェクト目標	成果	活動	投入	効率性	有効性	インパクト	自立発展性	妥当性
A										
B										
C										

（出所）筆者作成。

表8 南南協力評価におけるPTM

Code	Course title	Implementation organization	Evaluator	Survey year	Post technical cooperation	Certification examination	Purpose	Output	Efficiency	Effective
S1	Computer Software Technology	Japan-Singapore Institute of Software Technology (JSIST)	AFC	1998	Established by PTTC	x	To provide a basic knowledge of the system life cycle, and the tools and techniques used during the System Analysis and Design stages.	Identify the stages in a systems development life cycle. Understand the importance and purpose of systems analysis/design. Use the various tools and techniques of system analysis. Describe the activities and procedures in the design of files, codes and other man-machine interface for systems design. Apply the techniques of structured design to design systems from requirements specifications, producing designs which are well structured.	The operation of activities at JSIST is conducted efficiently and orderly system for the management and operation of the course has been established. The amount of money paid to JSIST has been sufficient for the proper operation and management of the course. A majority of ex-participants feel that the degree of goal achievement has been significant.	80% of ex-participants feel the course to be useful and re their work and organisation.
S2	Effective Management of Port Operation	Port of Singapore Authority (PSA)	AFC	2000	JICA training	x	Participants are able to conduct efficiency management of port operation	Acquire skills on planning and organization, and use of resources in order to ensure a high and efficient level of service to port users.	The course is considered as being very efficient - very good results achieved in relation to the efforts and resources put in. The cost and utilization of resources has always been within the budget and plans. The officers of implementing organization consider the results attained from the course in relation to resource allocation to be very efficient. The participants agreed that the course had enhanced their knowledge in the field of port operations management.	The course had achieved of enhancing the capability development of the port effective management operations. 79% of ex-participants more than 50% of acquire knowledge would be applied daily work, and almost all participants applied the their work. All of supervisors of the course has ability of participants efficiency of port management.
							Participants are able to conduct comprehensive and high level consulting works, to manage consulting projects and guide other consultants.	Acquire more practical and applicable consulting techniques. Improve their management ability for consulting projects.	The course is considered as being very efficient - very good results achieved in relation to the efforts and resources put in. The cost and utilization of resources has always been within the budget and plans.	The course had achieved of enhancing the capability development of the port productivity management improvement.

Impact	Impact	Relevance	Sustainability	Course customization	Lessons learned and recommendations
Understand the relevant to	25% of ex-participants was able to disseminate the knowledge learnt to others in the organisation. There is a large degree of exchanging of views among participants.	Since the establishment of the course, there has been no change in the course needs. The curriculum is sufficient for fulfilling the course needs and there is no need for review of the curriculum. 100% of ex-participants organisations want to continue sending participants.	100% of ex-participants organisations want to continue sending participants.	The curriculum is considered by JSIST to be appropriate to fulfil the objectives of the course, therefore, no curriculum was newly adopted or abolished.	The course should be extended for another 5 years. The course organisers should select participants more stringently (in terms of both knowledge level and English skill). The curriculum should be modified to include more theoretical and practical sessions, and the course organisers may need to evaluate the duration of the course.
And its purpose ability participants in the field of port	Nearly all ex-participants had shared the acquired knowledge with others. Nearly all ex-participants had institutionalized the procedures and practices taught into an instruction manual or written a report for the benefit of others. The course had led some participants to initiate a number of projects in their countries.	All the supervisors of ex-participants indicated that the opportunities for practicing the concepts learnt during the course, at work in their countries were very good. All ex-participants felt that the course was somewhat relevant and useful to their work.	The sustainability of any programme or training depends on Policy Support Measures, Appropriate Technology/Concepts, Socio-cultural Aspects, Institutional and Management Capacity, and Economic and Financial Aspects.	No information	The organisers should continue to select participants that are holding managerial positions in their respective ports, and to require that participants be proficient in the English language. In order to maintain contact with the participants after the course, the organisers should indicate in the General Brochure the expectation for the participants to update their contact address on a regular basis. It is recommended to extend the course for another 5 years.
Ex-participants had increased the > improve the environment.	92% of ex-participants had shared the acquired knowledge with others. 50% of ex-participants had institutionalized the procedures and practices taught into an instruction manual or written a report for the benefit of others.	92% of the supervisors of ex-participants felt that the course had somewhat improved the ability of the participants to manage consultancy projects and to guide other consultants.	The sustainability of any programme or training depends on Policy Support Measures, Appropriate Technology/Concepts, Socio-cultural Aspects, Institutional and Management Capacity, and Economic and Financial Aspects.	No information	The organisers should preferably select participants from national productivity organisations, and to require that participants be proficient in the English language. In order to maintain contact with the participants after the course, the organisers should indicate in the General Brochure the expectation for the participants to update their contact address on a regular basis.

(出所) 国際協力事業団企画・評価部評価監理室 (2002b)

PTMは、シンガポールとタイの第三国研修の終了時評価結果を使いメタ評価を行うために使用している。

この評価対象の各第三国研修では、全体20件のうちPDMを採用しているものは5件、評価5項目で評価を行っているもの7件であった。このような状況を考慮し、評価者の判断で評価報告書の記載事項を分類しPTMを作成している。項目として、プロジェクト内容であるプロジェクト目標と成果、評価5項目である効率性、目標達成度、効果、妥当性、自立発展性、更に研修の設立に大きな影響があると考えられるコース編成、教訓と提言を対象としてマトリックスを作成している（一部を表8として示す）。また、この評価の中で実施されたインタビュー結果も評価報告書の分析と同じ項目に南南協力の項目を加えて分類し、マトリックスを作成している。

この評価では、個々の評価結果によって作成したマトリックスに基づき、研修間の比較を行うことによって、シンガポールとタイにおいて実施された第三国研修の特性を分析している。しかし、PDMを使って実施した評価の数は少なく、実際にメタ評価を行うには十分な情報の確保は難しかった。但し、この分析により評価のあり方について考察を行うことができた。メタ評価は、評価結果を同一の観点から分析するものであり、この分析を通して評価方法についても考察することができ、評価の方法論の議論に役立った。

6. 結論と今後の方向性

(1) 結論

評価は、評価が活用され、実際のプロジェクトやプログラムの実施に影響を与えてこそ意味がある。しかし、個々の評価が実際の活動に影響を与えることは難しいことも指摘されている（Patton 1997、三好 2001a）。また、評価の議論では、個々の評価の方法論に焦点が与えられており、評価を使い、いかに影響

を与えていくかという視点が希薄である（三好 2001a）。このような状況の中で、PTMは、個々の評価分析とともに、評価結果を知識として蓄積し、より有効な教訓や提言を導き出すことを可能にするものであり、評価の有用な手法となり得る。

本論では、PTMの活用を政策、施策の評価への適用の可能性とその方向性について考察した。ここでは、本論の考察の結果として、PTMを活用する有効性を整理しておく。

PTMを使うことによって、PDMと評価5項目を同一のマトリックス内で分析することを可能にする。これにより、評価の設定である、評価時点の現状、あるべき状況との比較、そして因果関係を、相互に関連づけて分析できる。

個々のプロジェクトを、PTMを使い総合したり統合したりすることにより、プログラム評価の基盤を構築し得る。複数のプロジェクトの、例えば、同一分野の日本のプロジェクトをとりまとめた上での評価を実施することが可能になるし、他の援助国のプロジェクトを同じように扱えば、他の国のプロジェクトを含めて評価が可能となる。また、総合したり統合したりすることにより、個々のプロジェクトの上位目標、プロジェクト目標、成果、活動、投入についてプロジェクト間の差異を比較検討することが可能となり、特定の政策体系の下にプロジェクトの発掘・形成、計画策定、運営管理に有用な情報を提供しえることになる。

同一分野や同一課題に属する個々のプロジェクトの内容と評価をPTMとして併記することによって、同一分野、同一課題でのプロジェクトの特性・傾向などを分析・評価することができる。また、ベスト・プラクティスの特定化も可能であり、特定プロジェクトの指標化などが容易になる。また、これらの分析により、社会の知識の継続的な蓄積手段として評価を認識することを促進する。

このような分析を基に、上位目標の選定、プロジェクト目標のプライオリティ付けなどにより具体的な政策提言、政策決定が可能に

なる。

(2) PTMの活用促進

PTMはわが国の評価では新しい評価概念である。まだ実施事例も数少なく、今後PTMの適用を推進していくためには、第一義的には、個別のプロジェクト評価において、PTMを適用しその手法を改善していくことが必要である。しかし、筆者は、より施策・施策指向の評価を目指し、PTMを複数のプロジェクトに適用し、個々のプロジェクトでは導き出せない評価分析を行っていくことが重要と考える。そのためには、本論で議論したように、PTMを活用したプログラム評価手法の確立とメタ評価の実施促進を重視すべきである。

プログラム評価手法の確立

フィリピンのテーマ別評価ではPTMを使ってプロジェクトを取り纏めて「プログラムアプローチ・ロジック・モデル (PLM) 2」を作成し評価を行っている。この評価で作成されたPLM2は、従来不明確であった、プログラムとプロジェクトにおける、上位目標、プロジェクト目標、成果、活動、投入の関係を明確にし、プロジェクトをプログラムの部分として評価分析する枠組みを提供した。

このような試みは、上位目標、プロジェクト目標、成果、活動、投入の関係が明確にでき、単独のプロジェクトだけでなく、政策や施策や評価の基盤を作れる。このような分析基盤の構築は、プログラム評価手法の確立のための端緒を開くものであり、政策評価や施策評価の有用性を高めるゆえに、積極的に促進し実績を積み上げていくとともに、それらの実績を通してプログラム評価手法を確立していくことが重要と考える。

メタ評価の実施促進

メタ評価は、評価の教訓や提言をより概念化することにより、評価の影響を大きくし、持続性を持たせるものとする。一つ一つの評

価結果では、活用の源泉も限られるが、メタ評価を実施することによって、より広い範囲から評価結果を概念化・一般化でき、事業にかかわる課題や問題の認識を深めることができる。また、メタ評価の結果は、特定の考え方や事業等を支持したり、支援したりし、政治的な議論に使うことによって、政策決定や事業の推進に大きな影響を与え得る。

しかし、メタ評価の実績は日本ではまだ少ない。JICAでは、2001年度から、過去の評価案件を対象にメタ評価を取り入れた総合分析を実施する。最初の試みとして、過去4年間に実施された保健 / 医療分野の評価78件と教育分野の評価39件を対象に行うことにしている。今後は個別事業の評価の実施とともにメタ評価を行う体制を作り上げ、評価結果の概念化・一般化を行い、影響を与える工夫を行っていくことが重要である。

(3) 今後の方向性

本論では、プログラム評価の手法確立とメタ評価に焦点を当ててPTMの活用の有効性を考察した。しかし、PTM手法の評価における活用については、今後に検討すべきことは多い。本論での考察結果を踏まえれば、PTMを有効に活用することによって評価活動の有用性が高められる。ここでは最後にPTMの活用について今後の可能性として、業績評価の指標の適正化、ベンチマーキングの確保、内部評価・自己評価の強化における有効性を提示しておくこととする。

業績評価の指標の適正化

今日ODA事業のアカウンタビリティと評価結果のフィードバックがますます重視されてきており、ODA事業の裨益効果をより具体的に定量的に評価することが求められてきている。そのためには、事業の達成すべき目標、成果、活動を計画段階から評価概念を導入し、その妥当性、達成度、効率性、効果、自立発展性を定量的な指標を設定し評価して行くことが必要になる。また、事業の実施の

進捗に合わせ指標をモニタリングするとともに中間評価を行い、外部条件の変化があればこれを計画に反映させ適切に修正する、事業の終了時には目標、成果を指標で評価する、その結果として事後評価ではこれらの評価をレビューしより有効な評価結果を得ることが必要である。

このような事前評価から事後評価までの一貫した評価システムは、成果を重視した評価（Result-based Management）の中核を占めるものであり、ニュー・パブリック・マネジメントの一翼を担うものである。しかし、日本では、概念自体新しいものであり、日本に適した完成されたモデルがある訳ではない。2000年よりJICA、JBICでは事前評価を実施し事前評価表の公表を行っているが、ここでの問題は如何に適切な目標とその指標を選択していくかということである。メタ評価では、プロジェクトのPTM間の目標やその指標について比較分析される。また、メタ評価の分析の課程では、各評価の実施方法についても、情報を明確化するために検討される。これらの検討は、事前評価の実施方法、目標とその指標化に対して多くの貴重な情報を提供することになると考える。

ベンチマーキングの確保

ベンチマーキングとは、最も優れたもの、すなわちベスト・プラクティスと比較し、学習し、事業を改善していく手法である。ベスト・プラクティスとの比較分析により、事業の持つアドバンテージとディス・アドバンテージを再発見し得るとともに事業プロセス自体を革新しながら変革することが可能となり、組織的な学習効果がもたらされる。ベンチマーキングは、従来の事業手法からは生み出し得ない革新的な手法を導き出すことも可能であり、事業の手法のみならず、意識の変革にまで及び得る⁷。

開発事業においても、最近では、世銀をはじめ多くの援助機関でベスト・プラクティスの活用が強調されている。しかし、必ずしも適切な手法を提示しているわけではない。この

ようなベスト・プラクティスの活用には、メタ評価の手法が有効である。PTMを活用したメタ評価は、このようなベンチマーキングのために、プロジェクト間の比較検討により、多くのプロジェクトやプログラムの事例を提供し得る。また、プロジェクトやプログラムを比較することにより、ベスト・プラクティスを提供し得る。また、計画中、実施中のプログラムやプロジェクトに対して、適切な比較対象を提供し得る。この点を踏まえれば、メタ評価をベンチマーキングの情報基盤として活用していくことは、開発協力において事業の改善のために有効な手段であり、メタ評価とベンチマーキングを関連づけて展開していくことを重視すべきである。

内部評価・自己評価の強化

内部評価・自己評価は外部評価と比べて客観性に欠けるとの指摘がある。しかし、一方で内部評価結果はマネジメント改善に繋がる確率が高い、という報告もある⁸。内部評価は、マネージメント・ツールとして有効なものであり、今後とも評価の主要部分を占めるものである。それゆえに、如何に内部評価の質を向上させていくかは大きな課題である。マネージメント・ツールとして実施者の主体性を重視するとともに、評価の信頼性を強化していくことが必要である。いわゆる「お手盛り」評価にならないためには、マネジメントに関わる関係者が「より良い事業を行うための評価」という認識を共有し、協働して評価に取り組むことが必要になる⁹。

PTMを活用したメタ評価は、内部評価者に対して、事業の比較対照を提供することができる。前述した、事前評価から事後評価までの一貫した評価体制は、その主要部分を事業実施部署による内部評価・自己評価に基本を置く体制である。行政評価法で定められている事業評価、業績評価は、内部評価に基本をおいている。これらの体制が有効に機能するためには、自己満足にならない、チェック機能の内在する制度として構築される必要がある。メタ評価の有用性については前述した

が、他方、内部評価・自己評価では、評価結果を常に他と比較することが重要である。メタ評価を使つての目標、その指標の適正化やベスト・プラクティスとの比較やベンチマーキングの検討などを取り入れていくことを心掛けるべきである。

内部評価にあたっては、不適切な評価は、全ての知る権利のある対象者に事業の価値を根拠のある評価として提示し得ないことになり、結果として正当な判断をさまたげることになる。また、評価対象事業が、不適切な評価の基に事業を正当化され継続的に実施されることになり、将来に亘って問題を大きくする。このような点を踏まえて評価体制を整備し、評価を実施していくことが重要である。

付記

本論の内容は筆者の個人の見解であり、JICAの統一見解を示すものではない。

注記

- 1 PDMによる評価については、国際協力事業団企画・評価部評価監理室（2002）『JICA事業評価ガイドライン』及び三好・源（2001）を参照されたい。なお、「JICAガイドライン」は2002年に「実践的評価手法：JICA事業評価ガイドライン」として国際協力出版会より発行。
- 2 国際協力事業団企画・評価部評価監理室（2001）『JICA事業評価ガイドライン』
- 3 国際協力事業団企画・評価部評価監理室（2001）『JICA事業評価ガイドライン』を参照されたい。
- 4 PTMは、プログラム・ロジック・マトリックスとも呼ばれる。
- 5 ポリシー、プログラム、プロジェクトと政策、施策、事業は、それぞれ相対する概念として考えられている。しかし、実態面では、プロジェクトは、政策、施策、事業を含むものとして提示されるのが一般的である。これは、プロジェクトのPDMにおいても、（アウトカムと規定される）政策である上

位目標と施策であるプロジェクト目標が含まれており、実践上の用語使用を複雑にしている。本論では、政策、施策、事業を含んでプロジェクトを扱う場合に「プロジェクト」と「プログラム」の用語を、事業のみを扱う場合に「事業」の用語を便宜上使うこととした。

- 6 作成方法の詳細については、評価報告書掲載の作成マニュアルを参照されたい。
- 7 ヒブラー（1999）は、プロセスに着目し、ベスト・プラクティスを同一分野の事業に限らず広く他分野にも求めることを推奨している。
- 8 Sonnichsen（2000）p.83、三好・田中（2001）を参照されたい。
- 9 Stufflebeam（2001）は、広報の内意を受けた調査（Public Relations-Inspired Studies）と政治的に規制された調査（Politically Controlled Studies）を挙げ、これらを、問題のある評価アプローチとして擬似評価（Pseudo-evaluation）として位置付けている。

参考文献

- 上山信一（1997）『行政評価の時代 経営と顧客の視点から』、NTT出版
- 援助評価検討部会・評価研究作業委員会（2000）『「ODA評価体制」の改善に関する報告書』、外務省経済協力局評価室
- 国際協力事業団（2000）『ODAセミナー～より良いODA評価に向けて 報告書』
（<http://www.jica.go.jp/evaluation/etc/2001oda.html>）
- 国際協力事業団企画・評価部評価監理室（2001a）『評価結果のフィードバック』
（<http://www.jica.go.jp/evaluation/etc/2001feedback.html>）
- 国際協力事業団企画・評価部評価監理室『JICA事業評価ガイドライン』
- 国際協力事業団企画・評価部評価監理室（2002a）『特定テーマ評価 フィリピン人口・健康セクター・USAID連携： Part1

- 母子保健・家族計画（リプロダクティブ・ヘルス）分野』（印刷中）
- 国際協力事業団企画・評価部評価監理室（2002b）『特定テーマ評価 南南協力支援』（作成中）
- 国際協力事業団国際協力総合研修所（2001）『参加型評価基礎研究 - 国際協力と参加型評価』
- ノルウェー外務省（1996）国際協力事業団企画部評価監理課訳、『開発援助の評価 評価のためのハンドブック』、国際協力事業団企画部評価監理課
- 古川俊一（2001）『地方自治体の行政評価』『日本評価研究』1（1） 25-37
- ロバート、ヒブラー、トマス・ケリー、チャールズ・ケッテマン（高遠裕子訳）（1999）『ベスト・プラクティス 成功企業に学ぶ顧客戦略』、ティビーエス・ブリタニカ
- 源由理子（2001）『PCM手法と開発援助プロジェクトの評価 マネジメント支援としての評価』、PCM読本編集委員会編『PCM手法の理論と活用』、（財）国際開発高等教育機構
- 三好皓一（2001a）『評価のフィードバック-事業の利害関係者への評価の影響について』『国際開発研究』9（2） 71-86
- 三好皓一（2001b）『成果を重視した評価へ』、三好皓一・高千穂安長共編『国際協力の最前線 グローバル・ホットイシュー』、玉川大学出版部、59-75
- 三好皓一（2001c）『援助協調枠組みの変化』三好皓一・高千穂安長共編『国際協力の最前線 グローバル・ホットイシュー』、玉川大学出版部、42-58
- 三好皓一・田中弥生（2001）『参加型評価の将来性 参加型評価の概念と実践についての一考察』『日本評価研究』1（1） 63-79
- 三好皓一・源由理子（2001）『国際協力事業団「評価ガイドライン」の理論的枠組み ガイドライン作成のプロセスにおける一考察』『日本評価研究』1（2） 89-100
- Caracelli, Vallerie (2000). Evaluation Use at the Threshold of the Twenty-First Century. In Caracelli, Vallerie J. & Preskill, Hallie (Eds.), *New Direction for Evaluation*. American Evaluation Association, Safrancisco, CA: Jossey-Bass.
- Cousins, J.B. & Earl, L. M. (1995). *Participatory Evaluation in Education: Studies in Evaluation Use and Organizational Learning*. London : Falmer Press.
- Fetterman D. M., Kaftarian, A. J. & Wandersman, A. (eds.) (1996). *Empowerment Evaluation: Knowledge and Tools for Self-assessment and Accountability*. Thousand Oaks, CA : Sage Publications.
- Funnell, Sue C. (2000). Developing and Using a Program Theory Matrix for Program Evaluation and Performance Monitoring. In Rogers, Patrica J., Hacs, Timothy A., Petrosino, Anthony & Huebner, Tracy A. (Eds). *Program Theory in Evaluation: Challenges and Opportunities, New Direction for Evaluation*. American Evaluation Association, Safrancisco, CA: Jossey-Bass.
- House, Ernest R., Howe, Kenneth (2000). Deliberative Democratic Evaluation. In Greene, Jennifer C. & Henry, Gary T. (Eds.), *Evaluation as a Democratic Process: Promoting Deliberation, Dialogue, and Inclusion, New Direction for Evaluation*. American Evaluation Association, Safrancisco, CA: Jossey-Bass.
- Jackson, Edward T. & Kassam, Yusuf et al. (1998). *Knowledge Shared: Participatory Evaluation in Development Cooperation*. West Hartford, Connecticut: Kumarian Press.
- Kirkhart, Karen E (2000). Reconceptualizing Evaluation Use: An Integrated Theory of Influence. In Caracelli, Vallerie J. & Preskill, Hallie (Eds.), *New Direction for Evaluation*. American Evaluation Association, Safrancisco, CA: Jossey-Bass.
- OECD-DAC (2001). *Evaluation Feedback for Effective Learning and Accountability*, OECD.
- OECD-DAC (1998). *Review of The DAC*

- Principles for Evaluation of Development Assistance.*
- Patton, Michael Q. (1997). *Utilization-Focused Evaluation*, The New Century Text, 3rd ed., Thousand Oaks, CA: Sage Publications.
(邦訳『実用重視の事業評価入門』(2001) 大森彌監修、山本泰、長尾眞文編、清水弘文書房)
- Rogers, Patricia J., Hacsí, Timothy A., Petrosino, Anthony & Huebner, Tracy A. eds., (2000). *Program Theory in Evaluation: Challenges and Opportunities, New Direction for Evaluation*. American Evaluation Association, San Francisco, CA: Jossey-Bass.
- Rossi, Peter H., Freeman, Howard E., and Lipsey, Mark W. (1999). *Evaluation: a systematic approach*, 6th ed., Thousand Oaks, CA: Sage Publications.
- Sonnichsen, Richard C. (2000). *High Impact Internal Evaluation: a practitioner's guide to evaluating and consulting inside organizations*. Thousand Oaks, CA: Sage Publications.
- Stuffle, Daniel L. (2001). *Evaluation Models*. American Evaluation Association, San Francisco, CA: Jossey-Bass.
- United States General Accounting Office (1991). *Designing Evaluations*. Washington D.C.: U.S. General Accounting Office.
- Weiss, Carol H. (1998). *Evaluation: Second Edition*. Prentice-Hall, Inc., Upper Saddle River: New Jersey.
- (2001.11.13受理)

Utilization of Program Theory Matrix For Improving The Usefulness of Evaluation Analysis

Koichi Miyoshi

Japan International Cooperation Agency

Miyoshi.Koichi@jica.go.jp

Abstract

This paper aims at discussing the concept and utilization of program theory matrix (PTM) in evaluation in order to improve the usefulness of evaluation analysis. PTM is effective for evaluation of projects and compensates for the disadvantage of evaluations using the project design matrix (PDM). Furthermore, PTM provides a framework for analysis to synthesize a number of evaluation results and to create useful lessons and recommendations.

Many studies were conducted on PTM utilization in evaluation of an individual project. This paper mainly focuses on examining the possibility utilizing PTM for a number of projects. To put it concretely, the paper seeks a methodology of policy and program evaluation by synthesizing project evaluation results by utilizing PTM. The paper also seeks a way to generalize and conceptualize number of evaluation results by the meta-evaluation method. The paper examines two evaluation studies as case studies. These evaluations, namely the Thematic Evaluation on Population and Health Sector in the Philippines under JICA/USAID Collaboration: Part1 and Evaluation on Support for South-South Cooperation, provide some implications for utilization of PTM. In addition, the paper illustrates the effectiveness of PTM for results-based management, benchmarking and internal/self-evaluation.

Keywords

Project Design Matrix (PDM), Program Theory Matrix (PTM),
Policy Evaluation, Program Evaluation, Meta-evaluation

【研究ノート】

地方自治体の行政評価システムの課題と展望

小野 達也

富士通総研

onotatsu@fri.fujitsu.com

要 約

1990年代後半以降、全国の多くの地方自治体が行政評価の仕組みを導入する過程において多種多様な問題に直面している。これらの問題は「評価ツール（道具立て）」「導入戦略（段取り）」「実践活動（現場）」の3分類と、「広範」「徹底」「外部」という業績測定の発展方向の3分類という2次元の枠組みで整理・配置できる。この配置からは、さらに マネジメント・サイクル、外部との関係、基礎的ツール、評価システム自体のマネジメント、職員の意識・能力という5グループの問題群が抽出される。

行政評価システムが真価を発揮するための条件としては、（1）行政マネジメントの改革のためには各グループの問題について、「統治システム」「行政システム」「組織」「作業プロセス」「人」という改革のレベルに応じて取り組む必要があること、（2）行政活動がいかなる成果を挙げたかを、国民・市民が理解・納得できるように説明する責任であるアカウンタビリティを遂行すること、（3）自治体の外において 評価ツールの開発・提供・流通、改革レベルに応じて外部とリンクしたマネジメント・サイクルの実現、及び 国との関係における制約の解消が求められること、を挙げることができる。

キーワード

行政評価、業績測定、地方自治体、行政経営

1. はじめに

近年多くの地方自治体が行政評価に取り組むようになるにつれて、様々な問題点が自治体の内外から指摘され、議論されるようになってきている。本稿は、自治体が現実の局面で遭遇する諸々の問題を整理分析し、行政評価システムが真価を発揮するための条件という観点から今後を展望することを目的とする。

2. 地方自治体の行政評価

最近の行政評価の導入状況に関する調査によれば、都道府県・政令指定都市ではほとんどの団体が行政評価に取り組んでおり、市区町村でも取り組みが急増している¹。例えば総務省の調査（総務省自治行政局2001a）によれば平成13年7月末現在、46都道府県、全12指定都市において「導入済み」「試行中」「検討中」の何れかであり、「導入済み」とした

団体も都道府県では37団体（79％）、指定都市でも7団体（58％）に達している。前年との比較では、導入済みとした自治体が都道府県で対前年比で54％増、政令指定都市で75％増となっている。また市区町村では、「検討中」以上のところが全体の56％（全国3,235団体のうち1,809団体）、「導入済み」は5％（150団体）とまだ少数ながら対前年（77団体）比で95％増である。「導入済み」と「試行中」（140団体、前年94団体）をあわせると、対前年比で70％増となっている。

自治体で急速に取り組みが広がりつつある背景には、年々厳しさを増す財政状況の他、行政改革大綱（財政再建が主目的）の策定、総合計画の策定・進行管理手法の見直し、市民本位・住民満足への志向、職員の意識改革の必要性、市民からの要求などがあるといえる（小野・田淵2001, p. 4）。一般に行政評価の目的はマネジメントとアカウンタビリティに大別され（Aristigueta 1999, p. 149）またマネジメント改善の目的は、パフォーマンス（計画、予算、サービス提供など、しばしばこれが最重要とされる）モチベーション及び（自治体を構成する主体間の）コミュニケーションそれぞれの改善に分類することもできる（Greiner 1996, P.19）。日本の自治体及び国の行政評価において掲げられている導入目的も同様であるといえる。

自治体における行政評価の取り組みには内容・呼称とも様々なものがあるが、現在主流をなしている取り組みは次の、互いに共通部分のある4種類に分類することができる（小野2001, pp. 4-5）²。

- ・事務事業評価システム
- ・政策・施策レベルの評価システム
- ・総合計画の進行管理システム
- ・ベンチマーク指標によるベンチマーキング・システム

なお一般に、行政評価public sector evaluationをプログラム評価program evaluationと業績測定performance measurementの和集合と捉えれば、上に挙げたような自治体の行政評価はそのほとんどが「継続中または終了後の行政

活動の実績を主たる評価対象とし、場合により将来の予測や目標設定を含む」タイプであることから、業績測定あるいは業績測定を中心とする業績評価performance evaluationに相当する（古川・北大路2001, pp. 29-32；古川2001, pp. 30-31；古川2000上, pp. 62-67；Hatry 1999, p. 8；Hatry et al. 1981；小野2001, p. 2；島田・三菱総合研究所1999, pp. 41-42）といえる。なお、業績測定に相当しない行政評価としては、一部政策の評価をプログラム評価として行っている場合、古川（2001）が指摘するように欧米では（狭義）政策分析policy analysisとされることが多い大規模プロジェクトの費用便益分析などによる事前評価などがある。

このように行政評価に着手する自治体が増え、また導入した自治体も試行錯誤を続ける中で、数値指標の難しさを筆頭に、自治体の担当職員を中心とした悩みの訴えがアンケートなどを通じて紹介され、議論されるようになっていく（総務省自治行政局2001b；上山他編2000；小野2001, pp. 57-58）。

いま、地方自治体の行政評価は、一定の成果を挙げたとされる事例が増える一方で、表面的な導入に留まったり、困難に遭遇して停滞する事例も増えている。この結果、行政評価システムの効能とともに、その難しさが認識されつつあるといえる³。さらに地方自治体の現場において、行政評価の限界をやや短絡的に論ずるケースも見受けられる。

3．行政評価導入上の問題

地方自治体の行政評価システムが遭遇する様々な問題と、行政評価システムが成功するための条件については、最近のいくつかの文献がそれぞれの観点から列挙して検討を行っている（古川・北大路2001；上山2002；総務省自治行政局2001b；小野2001；小野・田中2001など）。本節と次節のねらいは、行政評価導入上の様々な問題を一定の概念枠組に沿って配置・構造化し、もって行政評価シ

システムが真価を発揮するための条件を導くことにある。

(1) 整理の枠組み

地方自治体が行政評価の導入に取り組む場合に遭遇する問題は、取り組む主体が都道府県、市区町村の何れであるかを問わず、また行政評価システムがどのような類型（事務事業評価システム、政策・施策レベルの評価システム、総合計画の進行管理システム、ベンチマーク指標によるベンチマーキング・システムなど）であるかを問わず、日本の現状では共通する論点が多いため、特に区別せずに論ずるのが一般的である。

この行政評価システム一般の導入に係る問題点を分析的に論ずるにあたっては、様々な観点からの整理が考えられる。まず、行政評価導入もひとつの事業であり、企画立案から実施（及び評価）にいたる時系列・論理的順序に従った整理が考えられる。行政評価システムの場合、これは「企画（または設計、検討など）段階」「試行段階」「実施（または運用）段階」と分けることが多い。しかし、自治体が行政評価システムを実際に導入する過程では試行錯誤が欠かせず、本格的に取り組む上では段階的に改良してゆくことが不可欠であるため、問題点をこのような観点から分類するのは必ずしも容易ではない。

自治体における行政評価導入上の様々な悩みを取り上げた総務省の報告（総務省自治行政局2001b）では、この企画、試行、実施という3段階（複数の段階にまたがる項目も多い）と、寄せられた悩みを分類した「評価方法」「検討・導入の進め方」「評価対象」「体制と公表」といったアド・ホックな区分をあわせ用いて整理している。

J. M. Greinerは、業績測定の本格的な普及・定着を阻む要因を、これまでの歴史を振り返って次の4つに分類している（Greiner 1996, pp.16-28）。即ち institutional obstacles：行政機関が評価というものを受容する過程での問題群、 pragmatic obstacles：行政機関トッ

プが評価の効果に対して、また職員が評価作業に対してとる、横並び意識や様子見を含むプラグマティズムによる態度、 technical obstacles：評価の概念問題に起因するものからデータ収集にいたる様々な技術的問題群、 financial obstacles：評価に必要な費用と労働時間の問題、の4つである。

本稿ではまず、このGreinerの4分類を参考にしながら、評価に限らず行政の現場に新たな手法を導入する場合に一般化できると考えられる「評価ツール（道具立て）」「導入戦略（段取り）」「実践活動（現場）」の3分類を設定した。評価ツールとは、評価指標や評価シートといった文字通りのツールのほか、目標設定やデータ収集などの手法なども含む広義の「道具立て」の集合であるとする。導入戦略は評価システムの導入目的や導入ステップ、制度や体制の整備など、評価システムの導入から活用に至るすべての段取りを含む。実践活動は、（戦略に基づいて）職員が実際に評価活動に携わる局面すべてである。

評価ツールに関する問題がGreinerのtechnical obstaclesに相当し、導入戦略上の問題がinstitutionalとfinancialに、実践活動における問題が概ねpragmaticに相当する。

さらに、もうひとつの軸となる分類として、「広範（extensive）」「徹底（intensive）」「外部（external）」という業績測定の発展方向による分類を設定する。これはG. Bouckaertが、欧州における業績測定発展の3次元としているものである（Bouckaert 1996）。

"extensive"とは、評価対象の範囲の拡大であり、試行の対象となった一部の事業から全事業へ、事業レベルの評価から政策・施策レベルへの拡張などに相当する。"intensive"とは、評価結果のより明確な活用に向かうことであり、当初の予算編成の補助資料や各種の参考資料といった位置付けから、様々なマネジメント（資源の配分、人事・組織管理など）に本格的に活用するようになる方向である。"external"は内部管理だけでなく外部（との関係）における活用、すなわち市民一般へのアカウンタビリティ遂行、議会や様々なステー

クホルダー、行政サービスの対象者などへの情報提供、コミュニケーション、連携を目指す方向である⁴。

今取り組まれている行政評価の手法はほぼすべて、諸外国の取り組みから直接・間接に学んだもの（少なくとも刺激を受けたもの）ということができる。そこで、その手法導入にあたっての問題点が内外を問わず公共部門一般にあてはまるのか、また日本固有の問題なのかという区別・認識も重要である⁵。

（２） 地方自治体の行政評価導入上の諸問題

上の「評価ツール（道具立て）」「導入戦略（段取り）」「実践活動（現場）」及び「広範（extensive）」「徹底（intensive）」「外部（extensive）」という2次元の分類枠組みを用いて、地方自治体の行政評価導入上の諸問題を整理すると表1のようになる⁶。表中、日本の行政評価に特徴的な問題という色彩の濃いものには を付した。

表中、導入戦略に係る問題の一部は広範、徹底、外部の3分類に共通のものである。また、評価ツールと実践活動に関しては、一部

表1 地方自治体の行政評価導入上の問題群

	広範 extensive	徹底 intensive	外部 external
評価ツール	(1)評価指標設定手順の不備 (2)評価シート設計上の不備（客観性の不足） (3)不用意な横断的比較（客観性の不足）		(12)住民からのデータ収集の軽視 (13)アカウンタビリティや情報提供、コミュニケーションにおけるＩＴの活用不足 (14)公表ツールの工夫不足
	(4)評価ツールにおけるＩＴの活用不足 (5)評価シートと評価以外の調書との重複 (6)統一的な事業体系（目的・手段体系）の不在、予算・計画の不对応など (7)評価対象の枠組みとしての総合計画の非戦略性	(8)二次評価の欠如（客観性の不足） (9)コスト（事業別コスト、人件費など）の把握困難 (10)不明確な目標値設定基準 (11)不明確な達成度評価の方法	
導入戦略	(15)中長期的な戦略の不在 (16)行政運営の円滑化、住民・職員の満足向上など「前向き」姿勢の不足 (17)庁内の体制・人材・予算の不足 (18)トップ（首長）の理解・支持の欠如 (19)評価システムの不十分な評価		(30)議会との連携不足 (31)監査との連携不足 (32)住民への働きかけの不足、協働環境の未成熟
	(20)評価関連データの整備・蓄積・活用の不足 (21)職員への説明・研修の不足 (22)評価作業と他事務との輻輳 (23)評価作業の支援体制の不足 (24)職員への過大な負荷（一次・二次評価作業、評価担当セクションの事務）	(25)行政評価導入目的の未確立 (26)評価結果を活用するサイクルを回すための仕組みの不在 (27)他の改革の取り組み（公会計改革、ＩＳＯ、他の行革の仕掛けなど）との競合、連携不足 (28)マネジメント担当部局間の連携不足 (29)評価結果の盲目的利用（数字の一人歩きなど）	
実践活動	(33)数量データの不用意な取り扱い (34)評価の必要性・意義の理解不足 (35)評価作業への抵抗、回避		(41)アカウンタビリティへの消極的姿勢、否定的姿勢 (42)顧客志向への消極的姿勢、否定的姿勢
	(36)アウトカム指標の設定困難 (37)評価能力の蓄積不足	(38)予算、計画関連作業における評価結果の活用不足 (39)評価に基づくマネジメント意識の欠如 (40)目標値の設定困難	

注： 印は、日本における特徴的な問題と思われるもの。

の問題が広範と徹底に共通する。これは、業績測定を活用モデルをサイバネティクス的に測定measurement、分析analysis、行動actionの3段階からなる制御系と捉えて各段階の問題点を考察するモデル（Smith 1996, pp.7-18）に即していえば、「広範」・「徹底」の2分類に共通する段階として「測定」のための評価ツールと「測定」の実践活動があり、それから「徹底」の分類において「分析」のための評価ツールと実践活動があることになる。

さて、表1の3行3列を眺めると、次の5本の問題群にグルーピングができることがわかる⁷。

マネジメント・サイクル

「徹底」の列に配置されている問題群。行政評価の結果を活用し、マネジメント・サイクルを回す上での問題。

外部との関係

「外部」の列に配置されている問題群。外部と関わるための発想やツール、実際の関係構築活動における問題。

基礎的ツール

「評価ツール」行と「広範」「徹底」列との交わりに配置されている問題群。一部と重なる。（行政内部に構築する）行政評価の各機能に関わるツールと手法に関する問題。

評価システム自体のマネジメント

「導入戦略」行に配置されている問題群。一部と重なる。広範、徹底、外部の何れの方向に発展する場合にも必要となる評価システム自体のマネジメント上の問題。

職員の意識・能力

「実践活動」行と「広範」「徹底」列との交わりに配置されている問題群。一部と重なる。（行政機関内部での）実践活動に関わる職員の意識と能力に関する問題。

4. 今後の展望 行政評価システムが真価を発揮するための条件

（1）行政マネジメント改革のために 改革レベルを意識した前進

多くの行政評価システムの導入目的は行政運営の改革・革新であるといえるが、前節で分類・整理した各グループの問題点を解決してゆくことによって、「広範（extensive）」「徹底（intensive）」「外部（extensive）」の各次元において行政評価システムは進化し、行政運営の革新に資することとなる。逆にいえば、これらの多岐にわたる課題がある程度解決されない限り、自治体の経営革新にまではいたらない。

さて、これら多種多様の問題群のどれから取り組むべきであろうか。問題のなかにはすぐにも取り組んで解決できるものもあれば、個々の自治体内部の努力では解決不可能なものも含まれる。ここでは公共システムのどのレベルの改革に必要なという観点から、問題を解決して行政評価システムが真価を発揮するための条件を整理する。公共システムの改革レベルはD. Osborneらに従い、上位から統治システムgoverning systems、行政システムadministrative systems、組織organizational level、作業プロセスwork processes、人people levelの5レベルとした（Osborne & Plastrik 1997、訳語は邦訳書による）⁸。

統治システムとは改革の対象として最も高いレベルであり、中央政府や地方政府の全体、教育制度（システム）や福祉制度（システム）そのものなどが相当する。行政システムとは、統治システムを構成する「予算・財政システム」「人事システム」などである。行政システムの改革なくして統治システムの改革はありえないが、その逆は真ではない。組織とは文字通り、ひとまとまりの施策・事業を所管する部署である。組織レベルでの改革は行政システムのレベルでの改革ほど大きな影響力はもたない。作業プロセスとは個々の事務・事業の執行プロセスであり、その限りでの行

政サイクルの改革どまりである。人とは公共部門の職員であり、すべての改革の前提となる。しかし意識改革が成し遂げられても、あるいは良い人材が揃っていても、それがプロセスや組織の改革をもたらずとは限らない。

この「統治システム」「行政システム」「組織」「作業プロセス」「人」という改革のレベル別に、前節の問題点を裏返しにした対策の例を示したのが表2である。表中 ～ の番号は前節において分類した問題群のグループである。

このように、行政評価システムが真価を発揮するための条件として、各グループの問題への対処は、それがどのレベルの改革に必要なのかを明確にして臨む必要がある。

(2) アカウンタビリティの遂行

行政評価システムの外部への発展において重要なアカウンタビリティを行政評価システ

ム導入の目的として掲げることも多い。行政のアカウンタビリティの概念は近年拡張され、今日では行政機関が税金の使途に限らず自己の遂行する政策・施策・事業についてその目的、方法とその成果を広く国民・市民に明らかにし、その理解を求める責任と解される（西尾1993）。したがって、単に評価結果など評価に関する情報を公開・公表するだけでは前節で掲げた問題への対処として十分ではない。アカウンタビリティとは、行政活動がいかなる成果を挙げたかを納税者に説明する責任、国民・市民が理解・納得できるように説明する責任を含む。

(3) 自治体の外における環境整備

行政評価システムが十分に機能するための必要条件として、個々の自治体の努力に加えて、次のような自治体の外における環境整備が挙げられる。

表2 行政マネジメント改革のための対策

改革のレベル	具体的な対策
人	意識改革（評価の必要性と意義、成果志向、顧客志向など）・・・研修などによる啓発 評価能力の向上（研修や評価ツールの整備） 外部の人材・機関の活用、人的交流
作業プロセス	評価作業における市民とのコミュニケーションの確立 行政評価システムの継続的改良（評価指標、評価シート、目標達成度評価など個々のツール） 評価作業におけるITの活用 評価システムの運用体制の整備（作業スケジュール調整、評価作業のサポートなど）
組織	行政評価システムの中期的導入戦略の策定 行政評価システムを核としたマネジメントの実行（組織内の各種マネジメント - 予算、計画への活用）
行政システム	他の改革の取り組みとの効果的な連携 評価結果に基づく行政の守備範囲の見直し 行政評価システムを核としたマネジメントの実行（庁内横断的マネジメント - 予算、計画、組織、人事への活用） 総合計画の有機的活用 統一的な事業体系の確立 評価と連携した会計制度の改革 行政評価システムの評価と継続的改良（作業負荷、分散処理などによる評価作業全体の効率） 他の管理的事務との輻輳の改善 国、都道府県（または市区町村）の評価との連携
統治システム	議会との連携 監査との連携 評価における市民との協働 首長のリーダーシップ発揮

注：斜体字は、自治体（首長部局）の外部との連携が必要なもの。

評価ツールの開発・提供・流通

重要性が認識されるとともにその難しさも明らかになった数量的な評価を行うためのツールなど、各種の評価に必要な道具立てを自治体が揃えるにあたっては、一定の公共財としてツールやツールを開発するノウハウが存在することが望ましい。各種の手法の開発とその提供・流通も必要である⁹。これらツールや手法の背景となる概念上の問題も多く残されている¹⁰。

改革レベルに応じたマネジメント・サイクルの実現

より上位の改革レベルにおいては、行政評価を軸としたマネジメント・サイクルが、首長部局の枠を越えて回らなければならない。すなわち、都道府県に対する市区町村（及びその逆）議会、監査委員、さらには市民にも有機的な連携が求められることとなる。

国との関係における制約の解消

行政システム以上のレベルで行政評価を活用した改革を実行しようという場合、その足かせとなるのが、税財源と権限の中央集権である。俗に3割自治といわれるような状況の改善や2000年の地方分権一括法を受けた実質的な権限移譲がまたれるところである。

謝辞

本論文の草稿に対して、筑波大学古川俊一教授及び査読者の方から数々の有益なご指摘を頂きました。深く感謝します。

注記

- 1 総務省の調査の他に最近のものでは総合研究開発機構（2002）や三菱総合研究所（2001）日経産業消費研究所（2001）などがあり、何れも自治体における行政評価の急速な広がりを報告している。
- 2 地方自治体の行政評価システムの代表的類型として、Furukawa and Hoshino（2001）は、事務事業評価システム、業務棚卸表、総合計画の進行管理、ベンチマーキングを

挙げている。

- 3 これは日本に限ったことではなく、例えばすでに20年、30年の歴史をもつ米国の地方政府の業績測定にしても、発展段階に応じて様々な問題が認識され、対策が講じられ続けていることが、本稿で引用したものを含む多数の文献で述べられている。
- 4 小野（1999）は、行政評価システムを全庁共通の中核的な道具とするための4条件として 適用範囲の広範性（行政活動全体の体系的な取り扱い） 評価の時系列的な連続性（事前・中間・事後評価の連続性） システムの汎用性（庁内横断的な構成と運用） システムの拡張性（評価システム外の改革との連携）を掲げたが、これらは概ね上の広範（ ） 徹底（ ） 外部（ ）に相当するといえる。
- 5 古川・北大路（2001）は、「評価システム構築にあたっての課題」を「公共部門に特有の課題」（企業との対比、内外を問わず）と「日本の行政に固有の課題」に分けて列挙している。またTanaka & Ono（2002）は米国との比較の観点から、日本の自治体の取り組みを評価したものである。
- 6 地方自治体が抱える様々な問題の具体的内容については、小野（2001）が表1に掲げた問題群よりも大括りの20課題にまとめて詳しく現状を紹介している。
- 7 5グループと表1中の諸問題との対応は次の通り。 マネジメント・サイクル 25～29, 38, 39； 外部との関係 12, 13, 14, 30, 31, 32, 41, 42； 基礎的ツール 1～11； 評価システム自体のマネジメント 15～25, 28, 30, 31, 32； 職員の意識・能力 33～37, 39, 40
- 8 欧州の行政改革研究において定評のあるPollitt & Bouckaert（2000）は、公共部門の改革の結果として得られたものを論ずるためには、「どのような」（誰にとっての、誰による、どのような目的での）結果なのかを抜きには語れないとして、結果をoperational results, results as process improvements, results as system improvement, results as

the realization of a visionという4つのレベルに分類しているが、これは上の5レベルと同様の考え方に立つものと考えることができる。

- 9 日本は米国と比べると、自治体間の情報流通や研究と実践の橋渡しの機会が不足している（小野・田中2001）。
- 10 欧州評価学会会長を務めるなどしたC. Pollittは、現在多くの国が取り組む業績測定が抱える問題として、conceptual, motivational, technicalの3種類があるとしている（2002年2月国際公会計学会関東部会における講演）。motivationalは前節の「導入戦略」「実践活動」に関わるものであり、technicalは「評価ツール」に関わるものである。一方conceptualな問題は、個々の行政機関を離れたレベルで取り組まなければならない。日本においても、行政の実状を踏まえた研究が必要となろう。

参考文献

- 上山信一（2002）『日本の行政評価 総括と展望』、第一法規
- 上山信一他編（2000）『実践・行政評価』、東京法令出版
- 小野達也（2001）『地方自治体の行政評価システムの課題と成功への条件』、「富士通総研研究レポート」No.117
- 小野達也（1999）「行政評価のかんどころ」、島田晴雄・三菱総合研究所政策研究部『行政評価』東洋経済新報社
- 小野達也、田中啓（2001）「行政評価の日米比較 行政経営の革新に向けて」、政策メッセ2001発表論文
- 小野達也、田渕雪子（2001）『行政評価ハンドブック』、東洋経済新報社
- 島田晴雄、三菱総合研究所政策研究部（1999）『行政評価』、東洋経済新報社
- 総合研究開発機構（2002）『行政運営の新しい手法に関する調査』
- 総務省自治行政局（2001a）『地方公共団体における行政評価の取組状況』
- 総務省自治行政局（2001b）『行政評価導入上の悩みと解決策』
- 西尾勝（1993）『行政学』有斐閣
- 日経産業消費研究所（2001）「全国調査670市・東京23区 本格化する民間管理手法導入」、『日経地域情報』No.368
- 古川俊一（2001）「地方自治体の行政評価」、『日本評価研究』1（1）：25-37
- 古川俊一（2000）「政策評価の概念・類型・課題（上）（下）」、『自治研究』76（2）、（4）
- 古川俊一、北大路信郷（2001）『公共部門評価の理論と実際』、日本加除出版
- 三菱総合研究所（2001）『地方自治体における行政評価への取り組みに関する実態調査』
- Aristigueta, M. P. (1999). *Managing for Results in State Government*, Westport, Conn: Quorum Books.
- Bouckaert, G. (1996). Measurement of Public Sector Performance: Some European Perspectives. In Halachmi, A. and Bouckaert, G. (Eds.) *Organizational Performance and Measurement in the Public Sector*, Westport, Conn: Quorum Books.
- Furukawa, S., & Hoshino, Y. (2001). Knowledge-Based Governance by Performance Measurement. *The Japanese Journal of Evaluation Studies*, 1（2）, 13-27.
- Greiner, J. M. (1996). Positioning Performance Measurement for the Twenty-First Century. In Halachmi, A. and Bouckaert, G. (Eds.) *Organizational Performance and Measurement in the Public Sector*, Westport, Conn: Quorum Books.
- Hatry, H. P. (1999). *Performance Measurement: Getting Results*. Washington, D.C.: The Urban Institute Press,
- Hatry, H. P.; Winnie, R. E.; and Fisk, D. M. (1981). *Practical Program Evaluation for State and Local Governments*. Washington, D.C.: The Urban Institute Press.
- Osborne, D.; and Plastrik, P. (1997). *Banishing Bureaucracy - the five strategies for reinventing government*. New York: Addison Wesley.

- (邦訳 (2001) 『脱官僚主義』 PHP)
- Pollitt, C.; and Bouckaert, G. (2000). *Public Management Reform - A Comparative Analysis*, Oxford: Oxford University Press.
- Smith, P. (1996). A Framework for Analysing the Measurement of Outcome. In Smith, P. (Ed.) *Measuring Outcome in the Public Sector*, London: Taylor & Francis.
- Tanaka, H. & Ono, T. (2002). *Performance Measurement Utilization in Japan: A Critical View*. A paper presented at the Annual National Conference of the American Society for Public Administration.
- (2002.3.29受理)

Performance Measurement in Local Government Problems and Prospects

Tatsuya Ono

Fujitsu Research Institute
onotatsu@fri.fujitsu.com

Abstract

The purpose of this paper is to put forward a framework for seeking a solution to the various problems in performance measurement in local governments.

After a brief review of the current status of their practices and accomplishments, it identifies the structure of problems to be solved by placing them in a two-dimensional table, where one dimension is categorized as "tools and techniques", "strategies" or "practices", and the other is categorized as "extensive", "intensive" or "external". Scrutinizing this table gives us a list of five groups of problems as follows: i) Management cycle; ii) External relation; iii) Basic methodology; iv) Management of performance measurement; and v) Attitude, capability and behavior of government officials.

The paper concludes with a suggestion of necessary conditions of successful practice to be met inside and outside local governments. First, each of five groups of problems should be tackled distinguishing the five levels of reform, that is, governing systems, administrative systems, organizational level, work processes, and people level. Secondly, accountability should be improved by focusing outcome. Lastly, there are three conditions outside local governments: i) development and dissemination of tools and methodologies; ii) collaboration with entities outside and citizens to incorporate management cycle; and iii) progress of decentralization as part of nationwide administrative reform.

Keywords

public sector evaluation, performance measurement, local government, public management

【研究ノート】

ジェンダー視点に立ったJICAプロジェクトの評価
東アフリカの事例から

本間 まり子

snmariko@hotmail.com

要 約

本研究ノートは、東アフリカにおいてJICAが実施しているプロジェクトにおけるジェンダー視点に立った活動を評価し、その特徴を分析することによって、ジェンダー・メインストリーミングの必要性を明らかにするものである。

プロジェクト活動は、その進行状況を「ステップ」に整理でき、そのステップごとに「ありうべき、ないしは必要とされるジェンダー視点に立った活動要素」が考えらる。本研究ノートは、それらの要素をまとめて「ジェンダー視点の活動9要件」として提示し、各プロジェクトで実際に行われている活動と対比することによって、ジェンダー視点からのプロジェクトの評価を試みる。次に、プロジェクトにおけるジェンダー視点に立った活動を促進するうえで、専門家やカウンターパート（ここではフォーカル・ポイントと呼ぶ）が重要な役割を果たしていることを論じる。最後にフォーカル・ポイントの活動とプロジェクトサイクルを構成するPDMとの関係を論じることで、プロジェクト本来の目標をより有効かつ効果的に達成するためには、ジェンダー・メインストリーミングが必要であると結論づける。

キーワード

ジェンダー視点、フォーカル・ポイント、プロジェクト・サイクル、
PDM、ジェンダー・メインストリーミング

1. はじめに

本稿は、国際協力事業団（以下JICA）が東アフリカで実施しているプロジェクトにおけるジェンダー視点に立った活動を評価し、ジェンダー・メインストリーミング¹の必要性について分析するものである。筆者は、2001年5月から6月に、JICAの専門家育成個人研修²の海外研修の一環として、JICAプロジェ

クトにおけるジェンダーを専門とする技術協力専門家の活動内容や環境を調査した。本稿は、その調査結果に基づいている。

以下では、第一に調査対象としたプロジェクト概要について紹介し、第二に、プロジェクトにおけるジェンダー視点に立った活動状況について分析する。第三に、それらの活動とフォーカル・ポイント³の関係について述べ、最後にそれらの現状分析を通じて、ジェ

ンダー・メインストーリーミングの重要性と方法について考察する。

2. 調査対象プロジェクトの概要

本稿は、東アフリカにおいてJICAが実施している、4件のプロジェクトを対象とした調査を基にしている。「タンザニア ソコイネ農業大学地域開発センター」、「エチオピア地下水開発・水供給訓練計画」、「ケニア 中等理科教育強化計画」、東アフリカ3ヶ国（ケニア、タンザニア、ウガンダ）における「アフリカ人造り拠点」の4件であり、いずれもプロジェクト方式技術協力⁴である。各プロジェクト概要については、Project Design Matrix（以下PDM）と呼ばれるプロジェクトに必要な、活動、投入、成果、目標、外部条件、指標等の諸要素と、それらの論理的な関係を示したプロジェクトの概要表をもとに、表1にまとめた⁵。

各プロジェクトの特徴は、三様に分類できる。協力の型による分類としては、「訓練、研修」型（3件）、「地域開発」型（2件）、「研究協力」型（2件）がある⁶。カウンターパート機関による分類としては、大学、もしくは同等の研究機関（3件）と、省庁（1件）がある。プロジェクトの年数による分類としては、1年目（1件）、3年目（1件）、4年目（2件）がある⁷。

3. JICAプロジェクトにおけるジェンダー視点に立った活動

ここでは、プロジェクトにおけるジェンダー視点に立った活動の現状について分析する。調査対象プロジェクトでは様々な活動が行われている。プロジェクトの活動はPDMに規定されるが、プロジェクトで実際に実施されている活動は、PDMの記載内容より具体的であり、多岐にわたっており、「目的」に向かった一連の流れが存在している。これらの

調査対象プロジェクトで実際に実施されている活動を、筆者が「ステップ」という形で順序立てたものが、図1から図4である⁸。

プロジェクト毎に作成されたこれらの図では、中央の欄には「ありうべき、ないしは必要とされるジェンダー視点に立った活動要素」を表示している。それらは、各ステップの活動に対応した、筆者が必要だと考えるジェンダー視点に立った活動要素である。右の欄には、それらの要素に対応した形で、プロジェクトで「実際に実施しているジェンダー視点の活動」を表示している。

これらの図を通じて、プロジェクトの活動がステップごとに実施されていること、ステップに対応して必要なジェンダー視点に立った活動要素が存在していることがわかる。同時に、実際にプロジェクトで行われているジェンダー視点に立った活動が、どのステップのどの要素に対応しているかが判断できる。また、どのステップにおいて、必要な活動が実施されていないか、どの要素における活動が今後なされるべきかということも一目で理解できる。

これらの図に表示されている「ありうべき、ないしは必要とされるジェンダー視点に立った活動要素」が、偏り無く必要な要件を満たしているかを判断するために作成されたものが、表2である。ここには、JICAの「WID配慮の手引書」（1993）による「WID配慮の5要件⁹」を参考に、プロジェクト実施段階において、筆者が必要だと考える「ジェンダー視点に立った活動9要件」が表記されている。この表は、これらの要件を各プロジェクト活動がどれだけ満たしているかという評価を、〇、△、×の三段階で試みた。△は、そのプロジェクトの活動ステップに対応した、「ありうべき、ないしは必要とされるジェンダー視点に立った活動要素」に含まれ、プロジェクトでその要素を充たす活動が、実際に実施されている要件であることを表している。△は、そのプロジェクトの活動ステップに対応し、ジェンダー視点に立った活動要素として含まれているが、プロジェクトでは現に実施

されていない要件であることを表している。
×は、そのプロジェクトの活動ステップでは、そのジェンダー視点に立った活動要件は充たされていない（必要とされないと認識される）ことを表している。

これらの図表から判断すると、調査対象プロジェクトでは、様々なジェンダー視点に立った活動要素が存在している。そして、それらは、現在プロジェクト準備段階にあるAプロジェクトを除いて、実際のプロジェクト活動を通じてある程度充たされていることが分かる。また、それらの活動要素は、筆者の考える、プロジェクトの実施段階に必要な「ジェンダー視点に立った活動9要件」を、ある程度充たしていることが確認出来た。

一方、図1～4と表2に浮き彫りにされた結果と、表1にあるPDMの内容を比較してみると、その間にはギャップが存在している。実際に、これだけ多く必要なジェンダー視点の活動要素が存在し、プロジェクト活動で現に実施されているにも関わらず、Dプロジェクトを除いて、PDMにジェンダー視点に立った活動に関する記述は見られない。

この状態は、JICA全体で見られる。調査対象となったプロジェクト方式技術協力を実施しているJICA社会開発協力部は、アフリカ地域において8件のプロジェクトを現在実施している。しかし、そのうちPDMにおいてジェンダーに関する記述があるのは1件のみである。JICAはプロジェクトを、ジェンダー視点から、女性がプロジェクトの主な対象者である「WID案件」と、女性が対象者の一部である「WID配慮案件」の2種類に分類していた（国際協力事業団 企画部1996）¹⁰。しかし、この分類方法では、どちらにも当てはまらないものは、ジェンダー視点の活動は関係ないと捉えられることが多かったように思われる。今回調査対象となった4つのプロジェクトにも、あまりジェンダーとは関係がないと思われている案件がある¹¹。

本項における分析結果によれば、ジェンダー視点に立った活動をする必要がないと、JICA内で考えられているプロジェクトでも、あり

うべき、ないしは必要とされるジェンダー視点に立った活動要素が存在し、プロジェクトの現場では、必要に応じてそれらの要素を充たす活動が実施されていることが多いのである。

4. ジェンダー視点に立った活動とフォーカル・ポイント

前項の分析により、プロジェクトの活動に対応した「ジェンダー視点に立った活動9要件」が考えられ、それらはPDMに記入されていない場合でも、必要に応じてプロジェクト活動を通じて充たされていることが分かった。しかしそれはまた、プロジェクト内で何らかの理由で明確に必要と認識されなければ、そのジェンダー視点に立った活動は実施されない危険性があることを示唆している。

プロジェクトの調査を通じて、ジェンダー視点に立った活動の実施には、それを促進するフォーカル・ポイントの存在が大きいということが判明した。表1からわかるように、調査対象の4件のプロジェクトでは、1件を除いて、TOR¹²に「ジェンダー」や「WID」という専門分野に特化して活動するよう記述されている専門家は配置されていない。しかし実際には、ジェンダー視点に立った活動を促進することを期待されている専門家の存在がみられた。これらの専門家を、ここではフォーカル・ポイントと呼ぶ。そして、彼らの活動をプロジェクト活動との関係から4種に分類し、分析する。すなわち、PDMに、ジェンダー視点について記述がされ、専門家やカウンターパートのTORとして明確化されているケース、必ずしもPDM等で明確化されていなかったが、プロジェクト・カウンターパートが、ジェンダー視点に立った活動のフォーカル・ポイントとなっているケース、必ずしもPDM等で明確化されていなかったが、日本人専門家がジェンダー視点に立った活動のフォーカル・ポイントとなっているケース、

ジェンダー視点に立った活動のフォーカル

- ・ポイントが不在のケース、である。

PDMにジェンダー視点について記述がされ、専門家やカウンターパートのTORとして明確化されているケース（D）

表1にある通り、Dプロジェクトでは、PDMにジェンダー視点について記述がなされており、その活動を担うべき日本人専門家が投入されている。すなわち「WID」分野の長期専門家である。このケースにあてはまるプロジェクトでは、TORにジェンダー視点に立った活動について記述されているこれらの日本人専門家がフォーカル・ポイントとなって、PDMに沿ったジェンダー視点に立った活動を実施している。また、Dプロジェクトでは、PDMに沿った形で、ジェンダー視点に立った活動を実施するカウンターパートも配置されており、フォーカル・ポイントとしての一翼を担っている。

必ずしもPDM等で明確化されていなかったが、プロジェクト・カウンターパートが、ジェンダー視点に立った活動のフォーカル・ポイントとなっているケース（B）

プロジェクト活動においては、その相手国のカウンターパートが大きな役割を占めており、彼（女）らがジェンダー視点に立った活動のフォーカル・ポイントになるケースもある。Bプロジェクトでは、数学分野のカウンターパートの女性が、ジェンダー視点の活動のフォーカル・ポイントとなっている。現在プロジェクトには数学、科学、物理、生物の4分野23人のカウンターパートがおり、彼女はその一人である。彼女のTORにジェンダー視点に立った活動に関する記述は特にはない。しかし、自身のプロジェクト活動を通じてジェンダー視点の必要性に気がつき、それがプロジェクト活動として実施されるよう、リーダーを始めとしたプロジェクト全体に働きかけた。

彼女が、ジェンダー視点の重要性について認識するに至った背景には、ケニアにおける

「理数科教育は女子には必要ない」という伝統的な考え方が存在する。またその影響で、女性の理数科教師の数が少ないことが、女子が理数科を選択するのを妨げる結果になっている。この状況を自ら克服して数学教師となったことが、彼女にジェンダー視点に立った活動をさせる要因となった。プロジェクト活動の他に、彼女は、女性教師が中心となって女子の教育推進活動を行うNGOにも、個人的に所属している。

このプロジェクトの特徴として、将来の持続性を視野に入れて、積極的な相手国カウンターパートの意見を登用していることが挙げられる。プロジェクト・リーダーによると、プロジェクトでは、日本人専門家と相手国カウンターパートを含めた全員でミーティングを重ね、PDMに沿ったプロジェクト活動の実施方法を検討している。こうした中、彼女はリーダーや自国のカウンターパート達を中心に、理数科教育を強化するうえで、女子に対する理数科教育の重要性の理解や、女子教員の増加の必要性を理解すること、教師間のジェンダー理解が必要であることの認識を深めた。そして、この彼女のフォーカル・ポイントとしての活動を受けた形で、プロジェクトは図2にあるようなジェンダー視点に立った活動を実施している。

これらの活動の具体的な効果として、プロジェクトでは、カウンターパートに女性を積極的に採用し、プロジェクトが実施している研修モジュールにジェンダー認識トレーニング採用することにより、理数科教師が、教育へのジェンダー配慮を行う必要があることが深く認識されるようになった。また、女子に対する理数科教育の促進が、女性理数科教師を今後増加させていき、女子に対する理数科教育が強化されることが期待されている。他に、ジェンダー理解促進トレーニングを受けた教師の間では、理数科教育に限らず、女子の教育へのアクセス一般に関する問題を認識するものも出てきており、今後、戦略的¹³なレベルから女子の教育を強化する活動が必要であると認識され始めている。

必ずしもPDM等で明確化されていなかったが、日本人専門家がジェンダー視点に立った活動のフォーカル・ポイントとなっているケース(C)

Cプロジェクトでは、社会経済分野で派遣されている長期専門家の女性が、ジェンダー視点に立った活動を中心的に行っている。彼女は自分の専門として、ジェンダーについて大学院で学んだ経緯がある。現在は、社会経済というジェンダーを含む広い専門性で、必要に応じて随時ジェンダー視点に立った活動を実施している。

例えば図3にあるように、プロジェクト対象者に関して現状把握のための調査を行う際に、男女別の調査を実施したり、プロジェクトの住民参加活動の項目選択に際してジェンダー視点を考慮し、女性のニーズに合った活動や、女性の参加しやすい活動形態を推進する、等である。住民参加の段階でも、積極的に女性の参加を促進する努力を行っている。

その効果として、プロジェクト活動への女性の積極的参加が実現した、プロジェクトの専門家やカウンターパートの中でジェンダー格差に対する認識が深まった、格差を是正するためにはジェンダー視点に立った活動が必要であるという認識が深まった、などが挙げられる。

ジェンダー視点に立った活動のフォーカル・ポイントが不在のケース(A)

Aプロジェクトでは、ジェンダー視点に立った活動のフォーカル・ポイントが不在である。しかし、ジェンダー視点に立った活動が全く行われていないわけではない。本調査中に、プロジェクト活動への理解促進と情報共有のためにワークショップが実施され、多くのカウンターパートが参加して、貧困対策に関する発表を行った。22件の発表のうち3件がジェンダーに関連するものであった。

図1にあるように、Aプロジェクトは現在準備フェーズであり、このワークショップの

他に、分析の対象となるような活動がまだ始まっていない。このプロジェクトの方針は、「プロジェクトに関係する機関が必要と認識する活動を実施する」というものであり、彼らの必要性の認識如何によっては、今後ジェンダー視点に立った活動が実施される可能性が大いにある。しかし、PDMへの記述もなく、フォーカル・ポイントも不在なことから、相手国カウンターパート達の認識が下がれば、ジェンダー視点に立った活動が全くなされないという危険性を含んでいる。

5. ジェンダー・メインストリーミングの重要性とその方法

これまでの分析を通じて、プロジェクトにおけるジェンダー視点に立った活動は、フォーカル・ポイントと呼ばれるプロジェクト関係者によって推進されているケースが多いことがわかった。彼(女)らの活動は必ずしもPDMに表現されていないにもかかわらず、プロジェクト目標を推進するにあたって、重要な役割を果たしている。しかしそれらの活動は、あらかじめPDMに表現されていなかったため、フォーカル・ポイントの個人的興味や問題意識、努力がなければ、実施されなかったかもしれない活動である。

これらの分析結果を踏まえ、本項では、これまで述べてきたジェンダー視点に立った活動を、プロジェクト・サイクルを通じて分析することで、持続可能なジェンダー視点の活動方法、すなわちジェンダー・メインストリーミングについて述べる。

JICAでは1994年度から、プロジェクト方式技術協力を中心として、プロジェクト・サイクル・マネジメント(PCM)手法を導入し、プロジェクト毎にPDMが作成されることになった。PCM手法とは、開発援助のプロジェクトの計画立案・実施・評価という一連のサイクルを、PDMと呼ばれるプロジェクト概要表を用いて運営管理する手法である。またこのサイクルは、プロジェクトの計画、

実施、評価のそれぞれの段階が一連の過程の中で位置付けられ、相互に関連していることを示している。

前項までの分析結果をこれにあてはめて考察すると、「計画段階」で作成されたPDMには記述がなくとも、「実施段階」で、プロジェクトをより成功裏に進める為に、ジェンダー視点が活動として明確化され、実施されているプロジェクトがある。つまり、計画段階ではジェンダー視点が必ずしも考慮されていたわけではないという意味で、上述のA、B、Cのようなプロジェクトがあることがわかる。そうしたプロジェクトでは、ジェンダー視点に立った活動があらかじめPDMに記述されておらず、フォーカル・ポイントの個人的興味や努力如何によっては、実施されなかった可能性についてはすでに述べた。

こうした状況を回避するために、筆者が、必要と考える「ジェンダー視点に立った活動の要件」を、プロジェクトサイクルに対応させ、図式化したものが図5である。見られるように、「ジェンダー視点に立った活動の要件」がプロジェクトの「実施段階」で効果的に充たされるよう、プロジェクト「計画段階」で「ジェンダー視点に立った活動の要件」と、プロジェクト目標との関係が認識され、PDMの成果と活動にも記述され、それに見合った投入がなされることが望まれる。またこうした一連の活動は、「評価段階」において評価され、次のサイクルへの引き継がれることが望まれる。

これをBプロジェクトの例として見る。図1から4は、プロジェクトの実施段階における活動を表しているものであり、Bプロジェクトにおいても、図2からわかるように、「実施段階」ではジェンダー視点に立った活動が行われている。しかし、PDMに記述がないために、プロジェクト目標である「パイロット・ディストリクトにおいて、現職員研修（INSET）により中等教育レベルの理数科教育が強化される」（表1のB）ことと、ジェンダー視点に立った活動は、直接結びつけられていない。この目標を効果的に実現するには、

理数科教育を受ける半数の女子生徒が置かれている状況を認識し、それに対応した活動が必要になる。このケースでは、特に伝統的な考え方が、女子の理数科教育へのアクセス及び参加を妨げている。そのため、このジェンダー不平等な状況を考慮した活動なしに、プロジェクト目標が達成されることは難しい。

現在、ジェンダー視点に立った活動は、数学分野のカウンターパートが、フォーカル・ポイントとなって推進している状態であり、彼女の進退如何によっては、実施されなくなる可能性が十分にある。PDMに記述がないために、プロジェクトで実施されたジェンダー視点に立った活動が、目標達成との関係から大きな役割を果たしたという点が評価され損なう、という問題も発生しうる。同じ観点から、プロジェクト活動が特定のジェンダーにマイナスの効果を与え、プロジェクト全体の有効性と効率性を低下させたとしても、今のPDMをもとにした評価では、評価されないことになる。

これまで論じられてきたような、プロジェクト形成段階からジェンダー視点を導入し、プロジェクト・サイクル全体にジェンダー視点が浸透することを、ジェンダー・メインストリーミングと呼ぶ。ジェンダー・メインストリーミングの必要性和有効性は、モーザ（1996）によって説かれており、本調査結果は、それを実証するものである。

6. おわりに

ジェンダー視点に立った活動は、JICAで従来ジェンダー配慮が必要ないと思われていたようなプロジェクトでも、実際には配慮されていることが多いが、PDMに記述がないため、評価に反映され損ない、このことが有効な評価の実施そのものを阻んできたと考えられる。ジェンダー視点に立った活動がプロジェクト形成段階から認識され、プロジェクトのすべてのサイクルに組み込まれることにより、適切な評価要素・指標が開発されるこ

とになり、より質の高い評価を行うためには必要かつ不可欠である。ひいては、プロジェクト目標の効果的な実現につながると考えられる。

(本論文の内容はあくまで著者の個人的意見であり、国際協力事業団の正式な見解を反映するものではない。)

注記

- 1 ジェンダー主流化の意。「ジェンダーと開発」の取り組みを開発の重点課題とし、ジェンダー平等を進めるための包括的戦略である、ジェンダー視点を企画立案段階から組み込んでいくことを指す。
- 2 国際協力事業団が実施している、技術協力専門家育成のための研修制度。国内海外あわせ一年以内の研修期間に、個々の研修員の必要に応じて研修内容を決定、実施する。
- 3 ジェンダー・フォーカル・ポイントともいう。ジェンダー推進のためプロジェクトや組織内外の機関との接点となり、ジェンダーについて他の職員に情報を提供したり、アドバイスを与える役割を担っている職員(この場合プロジェクト専門家やカウンターパート)を指す。
- 4 JICAの行っている技術協力の形態の一つ。「研修員の受入れ」「専門家派遣」「機材供与」の3つの形態を有機的に組合せ、1つの案件として数年間にわたって計画的に実施する方式。
- 5 JICAでは、すべてのプロジェクト方式技術協力においてこのPDMを作成している。
- 6 いくつかのプロジェクトは重複して数えられている。
- 7 ここにある数値は、本研修の実施された、2001年5月から6月時点のものである。
- 8 この活動ステップは、プロジェクトにおける活動内容紹介資料や研修を通じて収集した情報をもとに、著者が個人的に作成したものである。
- 9 ここでは、プロジェクトの計画から実施、評価までの3段階において必要なWID配慮要件を5つ提示している。すなわち「住民男女の現状分析」「女性参加促進方策」「女性の参加」「女性からの意見聴取」「WID専門性の活用」の5つである。
- 10 今後は、JICAの環境・女性課は、プロジェクトを、「WID案件(女性を主たる参加者、受益者とする案件)」、「ジェンダー平等案件(ジェンダー平等を上位目標、プロジェクト目標とする案件)」、「ジェンダー関連案件(ジェンダー視点に立った案件)」の3種類に、ジェンダー視点との関係から、分別しようとしている(2002予定)。
- 11 各プロジェクトの本部担当者とのインタビューでは、4件のうち2件のプロジェクト担当者が、「プロジェクトはジェンダーにはあまり関係の無い案件である」と回答した。
- 12 Terms of Referenceの略。JICAによる開発調査事業などに関して、通常、協力要請国から要請書とともに送付されてくる書類で、要請側が求める調査の内容などが記されている。このほかに広く業務内容や、要請内容を指すものとして用いられる(国際開発ジャーナル社 1999)。
- 13 Moser(1996)によると、ジェンダー・ニーズには、実際のものと戦略的なものの2種類が存在する。前者は「性別分業における女性の位置づけから生じるニーズ」であり、後者は、「男性に従属している女性の現状を把握する分析から出されたニーズ」である。この文章における戦略的ニーズとは、「男女間の理数科教育へのアクセスの不平等の原因そのものに対する働きかけ」に対する措置や活動を意味している。

参考文献

外務省 経済協力局編(2001)『我が国の政府 開発援助 2000 上巻』国際協力推進協会

キャロライン・モーザ(久保田賢一、久保田真弓訳)(1996)『ジェンダー・開発・NGO』新評論

国際開発高等教育機構編(2001)『PCM 開発援助のためのプロジェクト・サイクル・マネジメント』

国際開発ジャーナル社編(1999)『国際協力用語集【第2版】』

国際協力事業団 国際協力総合研修所編(2001)『国際協力と参加型評価』

国際協力事業団編(2002予定)『課題別指針

「ジェンダー主流化・WID」』

国際協力事業団編(1993)『WID配慮の手引書』

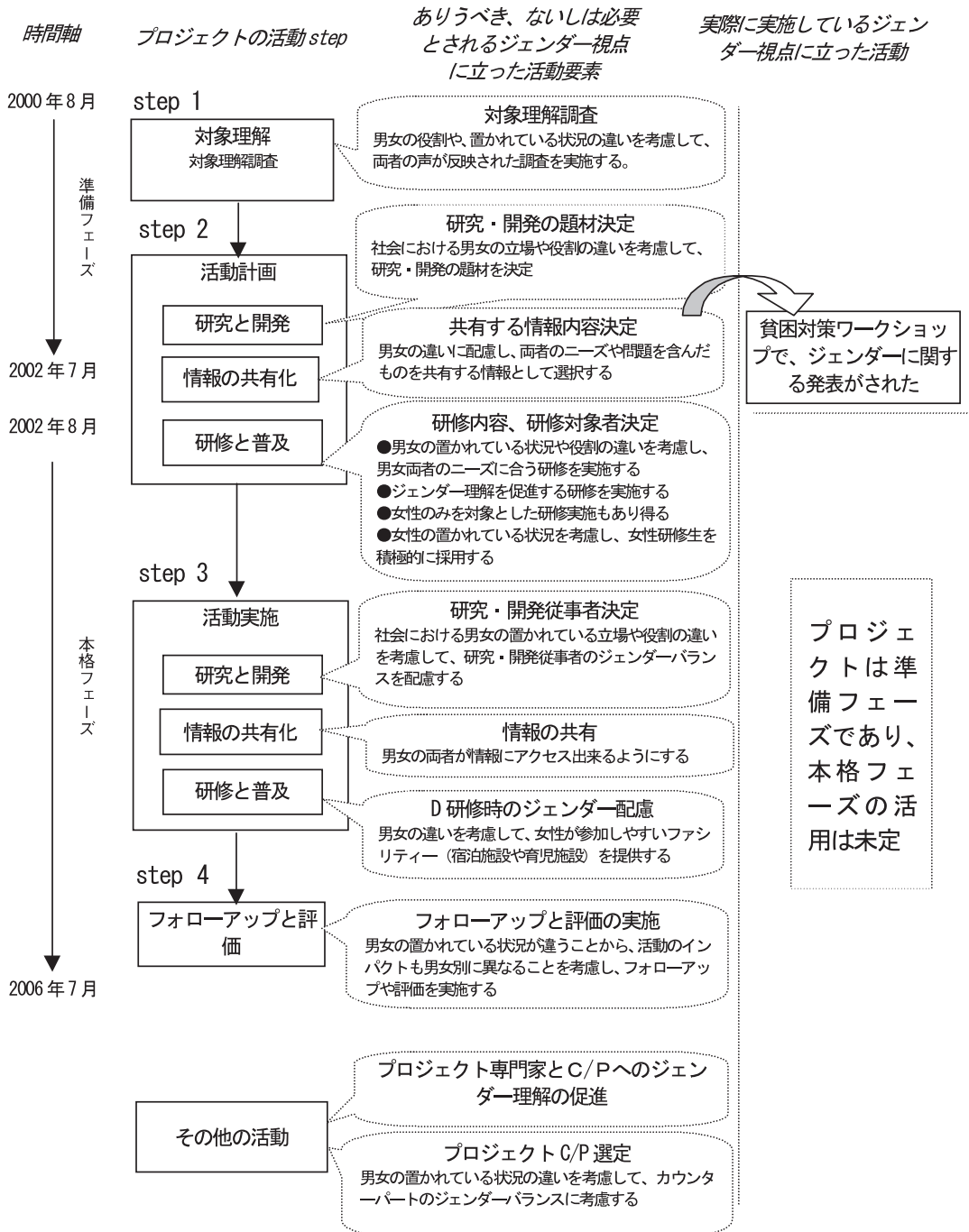
国際協力事業団 企画部編(1996)『「プロジェクト方式技術協力におけるWID配慮実施の現状と課題」調査 報告書』

国際協力事業団 農業開発調査部編(2000)『農林業協力のためのWID/ジェンダーハンドブック』

(2002.4.23受理)

図1 プロジェクトA

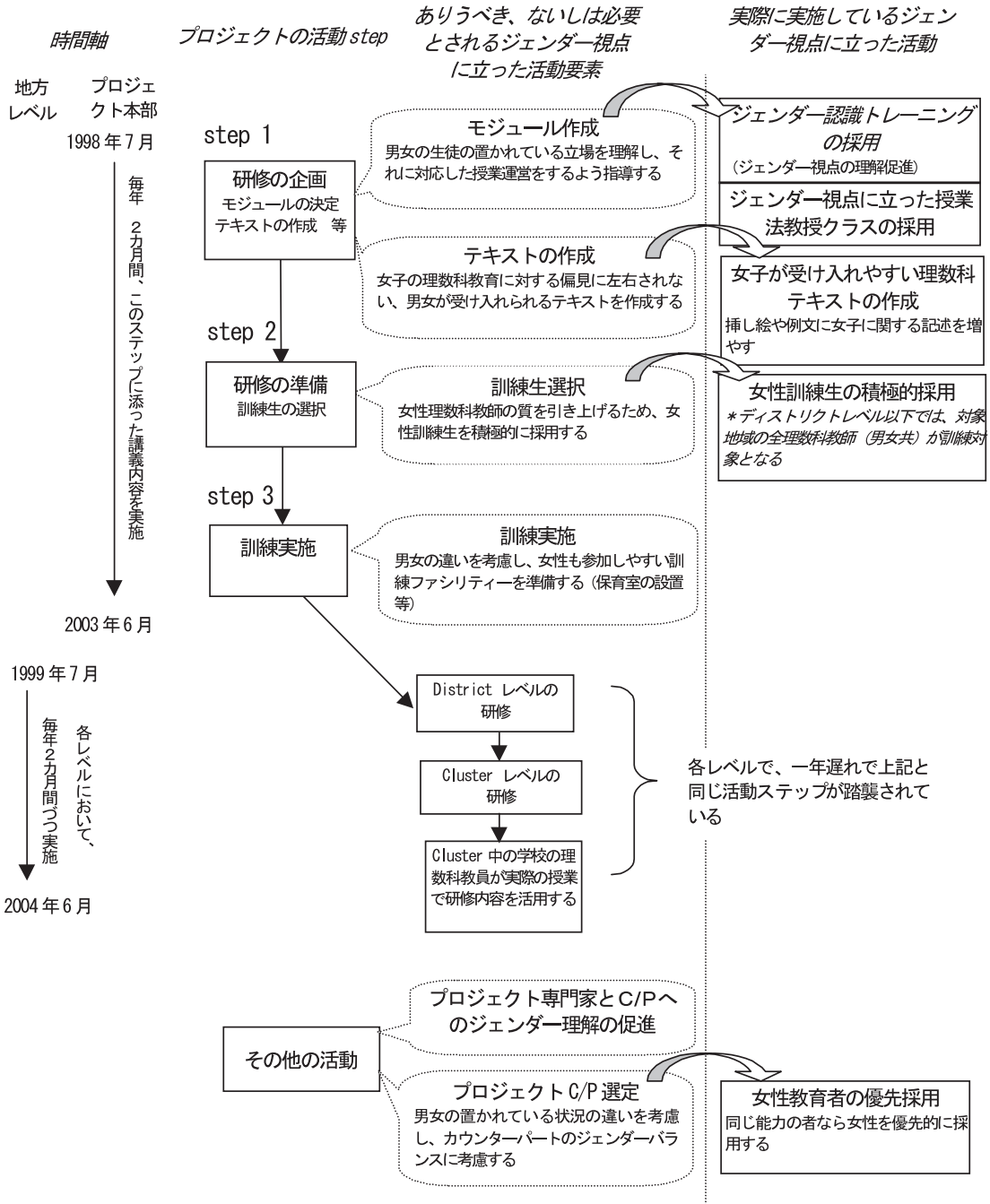
東アフリカ3カ国・アフリカ人造り拠点（準備フェーズ）（2000年8月～2002年7月2年目）



（出所）筆者作成（2001）

図2 プロジェクトB

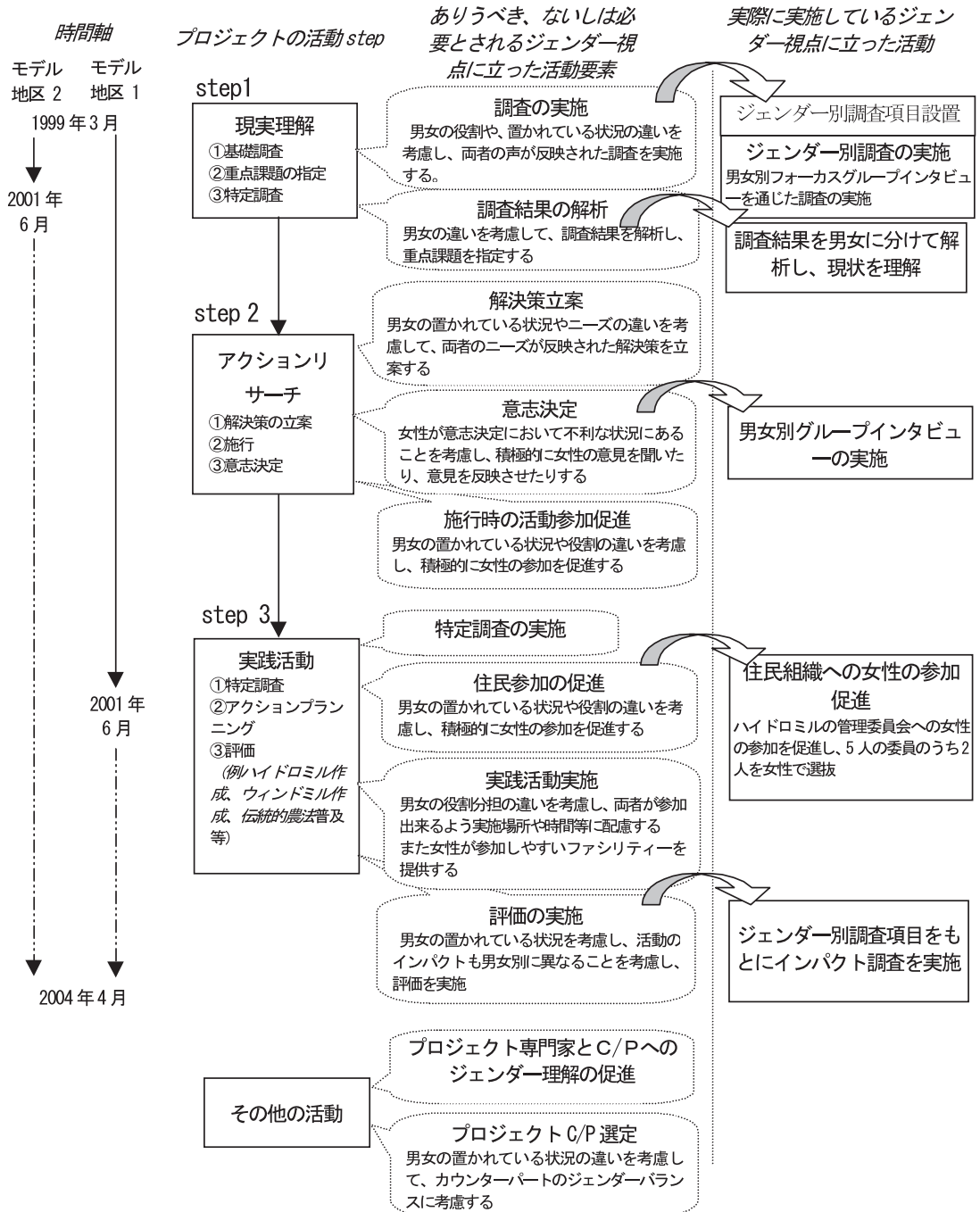
ケニア・中等理数科教育強化計画 (1998年7月～2003年6月 4年目)



(出所) 筆者作成 (2001)

図3 プロジェクトC

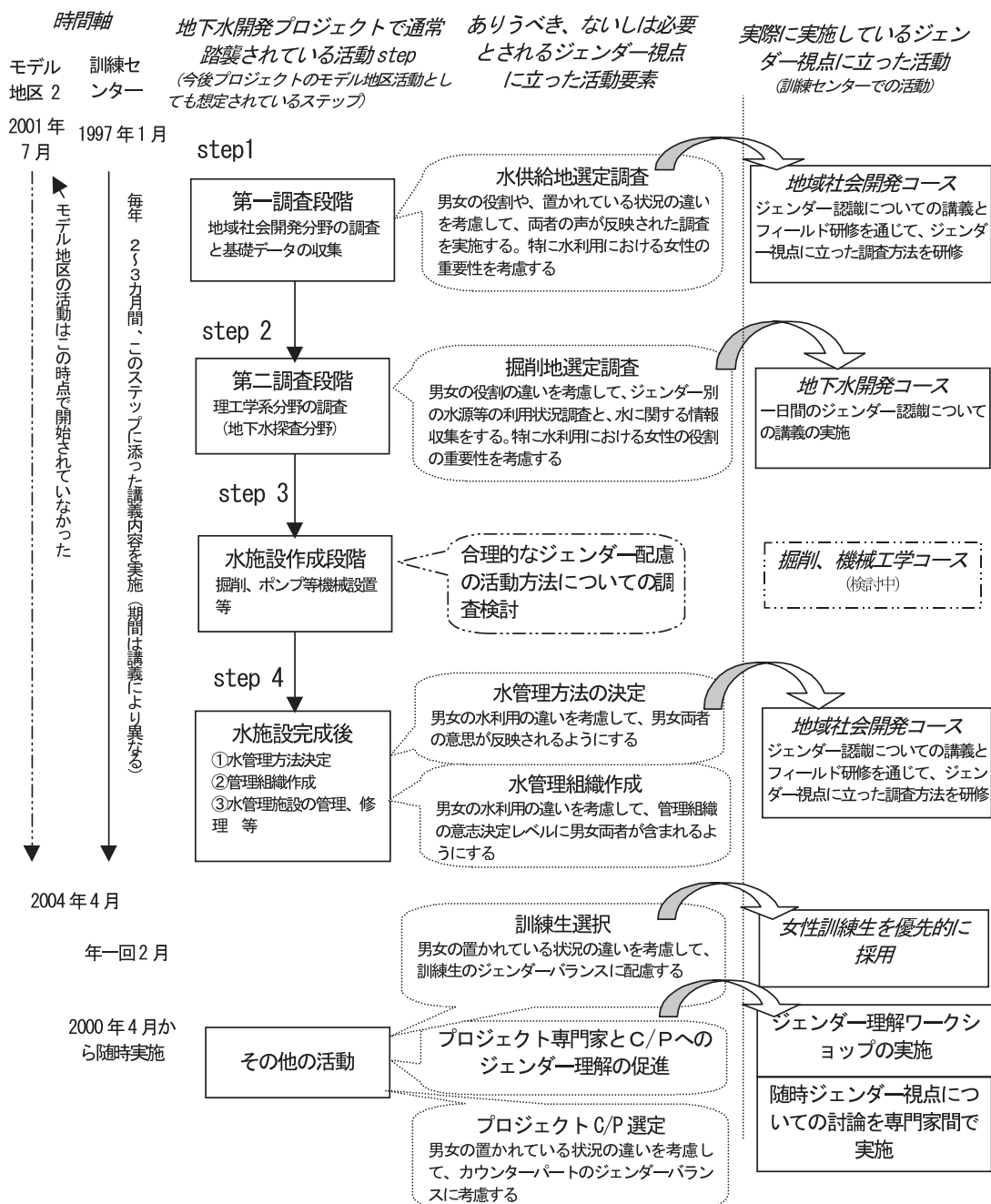
タンザニア・ソコイネ農業大学地域開発センター（1999年3月～2004年4月3年目）



(出所) 筆者作成 (2001)

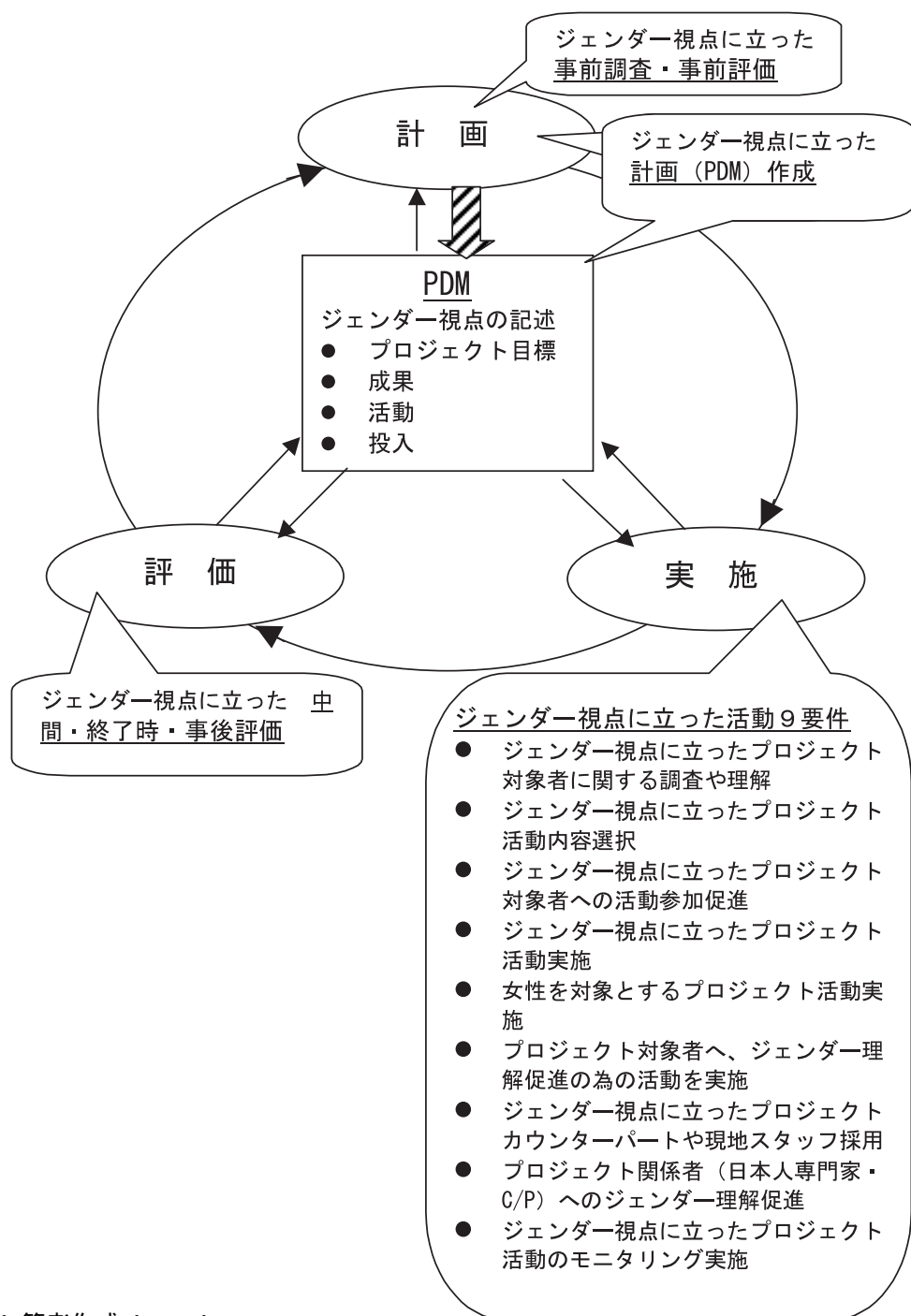
図4 プロジェクトD

エチオピア・地下水開発・水供給訓練計画 (1997年1月～2003年1月 4年目)



(出所) 筆者作成 (2001)

図5 プロジェクトサイクルにおけるジェンダー視点要件



（出所）筆者作成（2001）
国際開発高等教育機構（2001）p.4より

表1 調査対象プロジェクトの概要

	A アフリカ人造り拠点	B ケニア 中等理数科教育強化計画	C タンザニア ソコイネ 農業大学地域開発センター	D エチオピア地下水開発・水供給訓練計画
上位目標	アフリカにおける貧困削減、社会経済開発に資する人材育成	理数科目についてのケニア青少年の能力が向上する	対象地域の生活水準が向上する 持続可能な地域開発手法(SUA Method)がセンターや他の組織を通じて地域にも適用される	適正な水管理技術の研修を通じ、地下水開発及び、組織作りが強化され、十分かつ安全な水が供給される
プロジェクト目標	3機能（共同研究、研修普及、情報ネットワーク）を有するアフリカ人造り拠点の準備	パイロット・ディストリクトにおいて、現職員研修(INSET)により中等教育レベルの理数科教育が強化される	SUA Methodが2カ所の対象地域での活動を通じて構築される	地下水開発・水供給計画に依る人材を育成し、その過程において、ジェンダーと開発を重視する
成果	AICAD(アフリカ人造り拠点)の設置、他の参加大学との連携、共同研究開発機能の計画、開始、研修普及機能の計画、開始、情報ネットワーク機能の計画、開始	KSTC(プロジェクト本部)でパイロット・ディストリクトでの理数科分野でのキートレーナーのための養成研修システムの確立、INSETシステムの確立、リソースセンターとしてのKSTCとディストリクトセンターの役割強化	センターの設置、モデル地区の現状把握、住民によって地域の開発課題が明らかになる、内外の開発関連事例の調査、地域への支援計画の形成、パイロット・プロジェクトの実施、モニタリング・評価の実施	センターにおける、エン지니어、テクニシャン、普及指導員の技術訓練 実験的普及活動を行うモデル地区の確立
活動	AICADの設立、共同研究、研修普及、情報ネットワークに関し、ニーズの調査、それぞれの事業実施要領、計画の作成、その実施	カリキュラム、マニュアル、教材、トレーナー、評価システムの整備 INSETの実施と、教育マネージメント研修の実施 ニュースレター等の発行や、情報交換システムの整備、コンテストの実施	SUA(ソコイネ農業大学地域開発)センターに研究室(社会経済・環境保全・資源管理)および情報ユニットを置き、地域開発の諸問題についての調査研究、実践の計画、実施、モニタリング、評価への支援の実施	訓練センター センターの設立、カリキュラム開発、設備、機材の充実、インストラクターの能力の向上、訓練コースの実施 モデル地区 フィールドトレーニング、ケーススタディ、モニタリング評価の実施とその結果のセンターへのフィードバック
投入 (長期専門家)	チーフアドバイザー、業務調整員、共同研究、研修普及、情報ネットワーク	チーフアドバイザー、業務調整員、理数科教育(生物、科学、物理、数学)【2地区 30人の協力隊員派遣】	チーフアドバイザー、業務調整員、社会経済、資源管理、環境保全	チーフアドバイザー、業務調整員、掘削技術、WID、機械工学、地下水探査、地域開発

(出所) 各プロジェクトのPDMより(2001)

表2 ジェンダー視点の活動9要件

プロジェクト \ ジェンダー視点に立ったプロジェクト対象者に関する調査や理解	ジェンダー視点に立ったプロジェクト活動内容選択	ジェンダー視点に立ったプロジェクト対象者への活動参加促進	ジェンダー視点に立ったプロジェクト活動実施	女性を対象とするプロジェクト活動の実施	プロジェクト対象者へ、ジェンダー理解促進のための活動を実施	ジェンダー視点に立ったプロジェクトカウンターパートや現地スタッフ採用	プロジェクト関係者(日本人専門家・カウンターパート)へのジェンダー理解促進	ジェンダー視点に立ったプロジェクト活動のモニタリング実施
A								
B	x				x			x
C					x	x		
D	x				x			

：そのプロジェクトの活動ステップに対応した、ジェンダー視点に立った活動要件であり、実際プロジェクトでその要素を充たす活動が実施されている要件であることを表している。

：そのプロジェクトの活動ステップに対応し、ジェンダー視点に立った活動要素としては存在しているが、プロジェクトでは実際実施されていない要件であることを表している。

x：そのプロジェクトの活動ステップでは、そのジェンダー視点に立った活動要件は充たされない(必要とされないと認識される)ことを表している。

*Dプロジェクトは、訓練センターでの活動内容を使用

(出所) 筆者作成(2001)

Evaluation of JICA Projects from Socio-Gender Perspective: Case Studies in African Countries

Mariko Homma

snmariko@hotmail.com

Abstract

This paper evaluates JICA projects from a socio-gender perspective and studies the importance of gender mainstreaming in the projects.

Project activities can be categorized into several steps. In each step, there are gender-related elements. This paper provides nine necessary conditions for the gender aspect of project activities to be satisfied, and evaluates JICA projects based on them. JICA projects evaluated in this paper satisfy these conditions to a certain extent. Moreover, the gender related elements positively affect each project's outcome. However, satisfying the conditions depends on the effort of singular project members, who are called focal points.

JICA plans and conducts projects using the project cycle management method. In this method, projects have three cycles: planning, operation and evaluation, all through a project design matrix called PDM. This paper discusses about the strong relationship between PDM and the activities of the focal points.

In order to satisfy the nine conditions for the gender aspect, a conscious gender perspective must be objectives, this writer suggests that JICA implement gender mainstreaming in all projects.

Keywords

Gender Perspective, Focal Point, Project Cycle, Project Design Matrix, Gender Mainstreaming

【研究ノート】

開発援助の戦略的評価に関する一考察

盛 信博 青津 暢
コーエイ総合研究所

a1351@n-koei.co.jp a5648@n-koei.co.jp

要 約

限りある貴重な開発資源を総花的ではなく選択と集中、そして援助形態の組み合わせの妙を出すことがODA戦略の鍵となる。そのために必要なのがODAの戦略的評価である。

戦略を実行するための細かな事業計画が戦術であるが、従来のODA評価は戦術レベルの評価、つまり事業レベルの評価に偏りすぎている。それでは結果的に総花的な援助、つまり戦術ばかりで戦略なき援助となり、開発資源の効果的な活用とはならないであろう。

本稿では、戦略的ODA評価の事例研究として対バングラデシュODAのセクター／タイプ別マクロ経済的インパクト評価を試みる。開発援助の評価をマクロ経済指標へのインパクトに集約させるアイデアは、一定の戦略的インプリケーションを与えてくれるものである。

今後の課題は、まず大局的観点から被援助国の開発にとって効果的なODA分野を絞り、それから各重点分野における具体的なプロジェクトを複眼的観点から選択するという包括的かつ体系的な評価手法の開発とそれに基づく事例研究の積み重ねである。

キーワード

戦略的ODA評価、マクロ経済的インパクト評価、セクター／タイプ別ODA評価、
選択と集中、資源配分

1. ODA評価再考

「世界銀行の援助の有効性を評価するには、もはやプロジェクトのみに焦点を当ててはならず、部門と国のレベルでの世銀の活動全般の影響力を測らなければならない。これは新しい評価手段と様式を必要とする。」

World Bank (1998). *Assessing Aid* (小浜・富田訳 (2000)『有効な援助』東洋経済新報社)。
「我が国のODAは、個別プロジェクト支援

を中心とする援助形態から、特定地域、或いは分野の開発に向けて包括的な支援を提供し、援助効果を高める方向へ転換しつつある。」
援助評価検討部会・ODA評価研究会 (2001)「ODA評価研究会」報告書。

ODAをめぐる評価の理論の中では「プロジェクト」が援助活動の基礎単位と考えられており、また「プロジェクト」は目標が比較的単純で対象も把握しやすいためにその評価

も容易でODA評価の中でも理論化が最も進みそれなりの発展をみてきた。そのためODA評価とはすなわちプロジェクト評価であるといっても過言ではないようになっている。しかしながら、伝統的なプロジェクト評価は内部収益率や費用便益分析をはじめとする経済分析や財務分析などミクロの技術的な評価が多用される傾向にある。このようなプロジェクト評価は、個々の事業単位の評価で完結するというのが通例で個別化、断片化するデメリットがあると指摘されている。さらに、世銀では援助資金はファンジブル（流用可能）なので、援助の効果は援助プロジェクトそれだけの効果とは同じではなく個別の援助プロジェクトの収益率だけでは援助の本当の効果を明らかにすることはできないと明言している。世銀および我が国においても、従来の手法によるプロジェクト評価結果をいくら積み上げてもODA評価とはならないという認識から、これまでのプロジェクト単位のアウトプットに主眼を置く評価から上位レベルのマクロ的なインパクト評価へと主眼が移ってきている。

2. ODA評価の戦略的アプローチ

開発援助における戦略とは、「開発途上国におけるODA事業の長期的な基本構想」と定義されよう。戦略が実行可能であるためには、資源配分の裏付けが構想の中に入っていないなければならない。限りある貴重な開発資源を総花的ではなく選択と集中、そして援助形態の組み合わせの妙を出すことがODA戦略の鍵となる。そのために必要なのがODAの戦略的評価である。つまり、まず大局的観点から各ODA分野のプライオリティを明確にし、それから複眼的評価により各分野の個別案件を選択するというアプローチである。

戦略を実行するための細かな事業計画が戦術であるが、これまでのODA評価は戦術レベルの評価、つまり事業レベルの評価に偏りすぎている。それでは結果的に総花的な援助、

つまり戦術ばかりで戦略なき援助となり、開発資源の効果的な活用とはならないであろう。まず戦略があるべきであって、戦術はその次である。戦術の合算によって戦略が作成されるのではないのと同様に、ボトムアップ評価の合算による戦略評価はありえない。

本稿では、ODA戦略の策定に資するセクターレベルのマクロ経済的なインパクト評価を試みる。資源配分の決定に際してはマクロ経済のみならず、相手国の政治・社会的な要因も考慮すべきであろうが、以下において述べる開発援助の評価をマクロ経済指標へのインパクトに集約させるアイデアは、一定の戦略的インプリケーションを与えてくれるものである。

3. ODAのマクロ経済的インパクト評価手法

計量経済学的分析手法を用いて開発援助のセクター／タイプ別インパクトを定量的に示す。

分析に用いるデータおよび手法は以下の通りである。

(1) 分析データ

- ・セクター／タイプ別ODA 累積受取額（時系列データ；18カ年）
- ・セクター別GDP、輸出データ等（時系列データ；18カ年）

上記の各種基礎データを特定の指数に変換する¹。そして、これらの指数を対数変換し、グランジャー因果テストを行い、最後にその結果に基づいてパネル分析を行うというステップを踏む。

(2) グランジャー因果テスト

モデルを定式化する際の重要な問題の1つとして一方の変数が他方の変数と因果関係が

あるかどうかという問題がある。例えば、GDPとODAの関係を考えると、ODAの変化が原因でGDPが変化したのか（ODA → GDP）、それともGDPが変化したのが原因でODAが変化したのか（GDP → ODA）、または因果関係は無いのかなどの問題に答えるためにグランジャー因果という因果性の概念を導入した。先の例で言うと、もしODAに関する過去の情報がGDPの予測を改良するのに役立つならばODAはGDPのグランジャーの意味で原因になっていると言える。このことをテストするのがグランジャー因果テストであり、計量ソフト（TSP）を用いて検証する。同テストのモデル式は以下の通りである。

分析モデル：

$$\begin{aligned} Y_t &= a_0 + b_1 Y_{t-1} + c_1 X_{t-1} + \varepsilon_{1,t} \\ X_t &= d_0 + e_1 X_{t-1} + f_1 Y_{t-1} + \varepsilon_{2,t} \end{aligned}$$

（ここで、Y=GDPセクター等、X=ODAセクター／タイプ、 ε は確率的攪乱項である。）

パネル分析では以上のようにして検証した結果からモデルの推計F値が有意でモデルのF値が有意でないという結果、つまりXはYのグランジャーの意味で原因になっていないと統計学的に認められる変数を除いたものだけを用いてY（各GDPセクター等）を従属変数、X（各ODAセクター／タイプ）を独立変数としてモデル分析を行う。

（３） パネル分析

経済データの２つのタイプ（１．時系列＋２．横断面データ＝パネルデータ）を用いて分析対象の最適化行動を推定する手法がパネル分析である。本稿では、時系列データ＝１７カ年、横断面データ＝ODAセクター／タイプのデータを用いてどのセクター／タイプのODAが各種マクロ経済指標にインパクトを与えているかを検証する。

ところで、パネルデータから推定を行う際

の課題は推計モデルの選択である（固定効果モデルまたはランダム効果モデル）。推計においては各ODAセクター／タイプに固有な観測できない変数（たとえば、効率など）は個別効果として捉えられる。もし、援助効率が高く、マクロ経済にさらなるインパクトを与えているという傾向があれば個別効果と各説明変数間にプラスの相関が生じる。よって、もしこのような関係がある中でセクター別GDPをセクター別ODAに回帰させれば説明変数の係数にはプラスの偏りが生じる。そのため、個別効果が説明変数と無相関である場合と、相関がある場合とでは異なる推定量を用いて推計係数の偏りに対処しなければならない。前者の場合がランダム効果モデルであり、後者の場合が固定効果モデルである。どちらのモデルが適しているかは計量ソフトによる検定統計量によって判断でき、各モデルによって説明変数の係数も推計される。

４． ケース・スタディー

以上のような計量経済学的手法を用いて対バングラデシュのODA評価を行った結果を添付１に示している。添付表では推計モデル、推計係数そして推計係数の有意性を判断するt値などを示している。ちなみに、t値の大きさから統計的に有意と認められる推計係数を太字で示している（t値＞２）。また、従属変数（各GDPセクター、輸出）を上段に示し、説明変数（ODAタイプ、セクター）を左側側面に示している。先に述べたように、パネル分析に用いた説明変数はグランジャー因果テストで従属変数の変化の原因となっていると検証されたものだけを用いている。パネル分析結果によると、個別の因果テストでは関係が認められたものの、それらの説明変数の係数を同時推計したパネル分析では推計係数がプラスで有意な説明変数とそうではないものの、また符合がプラスのものとマイナスのものなど、各説明変数で異なる推計値が得られた。このことはある意味で合成の誤謬と言え

るだろう。つまり、個別セクターにおける個別プロジェクトのマクロ経済効果を評価してODA実効額を積算しても被援助国の援助吸収力、実施能力、管理能力などの制約条件により必ずしも望ましい経済効果へとは結びつかないということである。したがって、まずマクロ的視点からインパクトのあるODAセクターを特定し、それに基づき当該セクターの具体的なプロジェクトをデザインするというトップ・ダウンの評価アプローチが望まれる。

各説明変数の推計係数値の符号条件、有意性（t値）、弾力性（各係数値は弾力性を示している）をみると、ODAタイプでは一部を除き各タイプとも符号条件および係数の有意性を示している。また、資金協力と技術協力のマクロ経済へのインパクトは前者が後者よりも若干大きいと言えるが、共に効果的である。ODAのセクター別インパクト評価では、符号条件および有意性を満たす係数の数から、教育（ED：4つ）および社会保障、女性・青少年開発（SWY：3つ）分野への援助がバングラデシュのマクロ経済に累積的なインパクトがあると言える。その他のセクターでは農村・組織開発（RDI：2つ）、保健医療・家族計画（HPF：2つ）および通信（COM：2つ）分野への援助が比較的インパクトがあると言えよう。

添付2ではGDPセクターおよび輸出間の因果関係のテスト結果を示している。各列が因果関係の“原因”を示しており、グランジャー因果テストにより有意な関係が認められたものを“+”と表現している。このテスト結果によると、鉱業、農業、建設業などの成長と他のGDPセクターの成長との因果関係が広く認められる。つまり、これらのGDPセクターの成長が他のGDPセクターの成長を促す原因となっている可能性が高いと言える。また、この表の各行に着目すると、電気・ガス・水道セクターのポイントが最も多い。各行は因果関係の“結果”を示しており、あるセクターの成長が当該セクターの成長に寄与していることを示している。つまり、当該分野の付加

価値の増加は他の各GDPセクターの成長に大きく依存していると言える。

したがって、マクロ経済的観点から判断すると、ODAの重点セクターはEDおよびSWY、そしてRDI、HPF、COMセクターであり、効果的なODAタイプは資金協力と技術協力との連携を強化したタイプであると言える。また、これらの援助重点分野以外にも鉱業セクターの成長にインパクトが認められる水資源開発援助（WR）、農業セクターの成長にインパクトが認められる住宅・施設・上水道整備（PWS）に対する援助、そして建設セクターの成長にインパクトが認められる工業開発援助（INDなども考慮すべきであろう）。

5. 結語

本稿における主張は、限りある貴重な開発資源を総花的ではなく特定の分野へ集中させ、そのインパクトを他に波及させる戦略的な援助が必要だということである。そのためには、これまでのような戦術レベル（プロジェクト・ベース）の評価の積み上げではない新たな評価アプローチが望まれる。その一手法が、まず大局的観点から被援助国の開発にとって効果的なODA分野を絞り、それから各重点分野における具体的なプロジェクトを複眼的評価に基づき選択するという評価アプローチである。

上述したマクロ経済に対するインパクト評価はセクター／タイプ別ODAの事後評価であり、その評価結果は今後の開発資源投入の方向性を示す弱公準である。したがって、マクロ経済的インパクトを考慮したプロジェクト選択のためには、そのプロジェクトが重点分野の案件であることが必要条件ではあるが、十分条件ではない。つまり、プロジェクト選択のためにはプロジェクト・コストの比較、組織・制度評価、環境影響評価、ミクロ経済分析、裨益者意識調査や実施機関の実施能力調査などミクロ的な定量分析や定性調査の結果も取り込んだ複眼的評価が必要である。

したがって、今後の課題は、マクロおよびミクロデータを用いた計量経済学的分析手法とアンケート調査などの定性的な分析手法などを組み合わせた包括的かつ体系的な評価手法の開発、そしてそれに基づく事例研究の積み重ねである。

注) 本研究は国際開発センターとコーエイ総合研究所が行った外務省委託研究報告書「政策レベルおよびプログラム・レベル評価の手法等に関する調査研究」(2000)の分担執筆部分の一部に、加筆修正を行ったものである。

注記

- 1 各種マクロデータの変換手法については、コラム1参照。

参考文献

- 援助評価検討部会・ODA評価研究会(2001)「ODA評価研究会」報告書
 ODA改革懇談会(2001)「第2次ODA改革懇談会」中間報告書
 刈屋武昭監修、日本銀行調査統計局編(1996)『計量経済分析の基礎と応用』、東洋経済新報社
 和合肇・伴金美(1995)『TSPによる経済データの分析 第2版』、東京大学出版会
 アジア経済研究所(1999)『テキストブック 開発経済学』、有斐閣
 ADB (Asian Development Bank) (1999, 2000). *Key Indicators of Developing Asian and Pacific Countries*. ADB.
 Economic Relations Division, Ministry of Finance (2000). *Flow of External Resources into Bangladesh*. The Government of Bangladesh.
 Damodar N. Gujarati (1995). *Basic Econometrics* 3rd ed. Singapore: McGraw-Hill, Inc.
 World Bank (1998). *Assessing Aid: What works, What doesn't, and Why*. New York: Oxford University Press. (小浜・富田訳(2000)『有効な援助』、東洋経済新報社)

略 語

ODA タイプ

CA	資金協力
TA	技術協力

ODA セクター

AGR	農業
RDI	農村・組織開発
WR	水資源開発
IND	工業
PO	電力
OGN	石油・ガス・天然資源開発
TR	運輸交通
COM	通信
PWS	住宅・施設、上水道整備
ED	教育
SC	スポーツ・文化
HPF	保健医療・家族計画
MM	マスメディア
SWY	社会保障、女性・青少年開発
PA	行政組織強化
ST	科学・技術研究
LM	職業訓練

GDP セクター

AG	農業
MI	鉱業
MAN	製造業
EGW	電気・ガス・水道
CON	建設
TRA	商業
TC	運輸・通信
FIN	金融
PA	公共サービス
OTH	その他

その他

EX	輸出
----	----

コラム1 データ変換方法について

本稿での分析に用いたマクロ経済および援助額の各種データは基礎データを以下のように変換したものである。

1. マクロ経済データ

セクター別名目GDPおよび輸出統計（出所：ADB、“Key Indicators 1999, 2000”）の変換は以下の通り（例：農業生産（AG））。

1982	1990	1999(year)
AG:104.6	295.1	554.7(billion Taka)

1982	1990	1999(year)
AG: 35.4	100.0	188.0(1990=100)

1982	1990	1999(year)
AG: 3.57	4.61	5.24(対数変換)

2. ODAデータ

セクター・タイプ別援助統計（出所：MOF、“Flow of External Resources into Bangladesh 2000”）の変換は以下の通り（例：水資源開発セクター(WR)）。

1982	1990	1999(year)
WR:10800.0	46049.0	78750.0

(Taka in Lakh)

1982	1990	1999(year)
WR:10800.0	251479.0	753051.3(累積額)

1982	1990	1999(year)
WR: 4.3	100.0	299.4(1990=100)

1982	1990	1999 (year)
WR: 1.46	4.61	5.70 (対数変換)

添付1

対バングラデシュODA評価, 1983-99

従属変数	AG	MI	EGW	CON	TC	FI	OTH	EX
(ODAタイプ別推計結果)								
定数項			4.082 (56.83)	1.786 (5.739)				2.824 (23.52)
CA	0.160 (7.387)	0.527 (3.029)	-0.084 (-4.102)	0.455 (5.527)	0.161 (9.197)	0.543 (8.612)	0.253 (15.91)	0.262 (5.582)
TA	0.124 (8.242)	0.635 (5.233)	-0.068 (-8.288)	0.198 (8.318)	0.156 (13.15)	0.400 (10.20)	0.158 (16.32)	0.228 (6.774)
自由度修正済決定係数	0.921	0.765	0.741	0.815	0.953	0.853	0.944	0.909
観測数	170	187	85	51	170	255	204	68
モデル	固定	固定	ランダム	ランダム	固定	固定	固定	ランダム
(ODAセクター別推計結果)								
定数項	3.529 (41.55)	-0.102 (-0.306)	3.916 (59.85)	2.379 (12.21)			3.158 (79.88)	2.938 (20.56)
AGR	-	-	-0.224 (-3.081)	-	-0.254 (-3.444)	0.786 (1.191)	-	-
RDI	0.121 (0.532)	-4.431 (-5.561)	-	-	-0.293 (-6.585)	3.566 (3.364)	0.034 (0.300)	0.276 (2.626)
WR	-0.019 (-0.084)	2.290 (3.790)	-	-	-	-	-0.091 (-1.254)	-
IND	-0.081 (-0.505)	-	0.072 (1.109)	0.719 (6.732)	-	-0.835 (-1.745)	-	-
P O	-0.008 (-0.035)	-1.427 (-3.413)	-	-	-	-0.776 (-0.953)	-	-
OGN	-0.067 (-0.377)	-	-	-	0.069 (1.265)	-0.318 (-0.575)	-	-
TR	-	-0.013 (-0.171)	-	-	-0.103 (-1.229)	0.562 (0.856)	0.268 (3.200)	-
COM	-0.099 (-1.005)	1.874 (6.162)	-	-	0.050 (1.353)	0.976 (2.495)	-0.141 (-2.930)	-
PWS	0.189 (2.207)	-0.778 (-0.979)	-0.072 (-1.049)	-	0.002 (0.031)	-1.838 (-1.729)	-0.182 (-2.463)	-
ED	0.141 (1.151)	1.492 (4.600)	-	-	0.221 (8.695)	0.719 (1.392)	0.154 (2.423)	0.162 (1.656)
SC	0.035 (0.618)	0.158 (0.684)	-	-	-	0.109 (0.537)	0.107 (3.099)	-
HPF	-	-0.232 (-0.853)	-	-	0.057 (3.035)	-0.084 (-0.452)	0.078 (2.363)	0.066 (0.599)
MM	-	-	-	-	0.047 (2.300)	-	-	-0.043 (-0.568)
SWY	-	2.856 (4.793)	-	-	0.475 (7.318)	1.367 (3.066)	0.091 (1.055)	-
PA	0.035 (0.618)	-0.795 (-1.986)	-	-	-	-0.356 (-0.912)	0.095 (1.902)	-
ST	-	-	0.116 (3.080)	-0.231 (-2.764)	-	-	-0.077 (-2.861)	-
LM	-	-	-0.004 (-0.288)	0.033 (0.892)	-	-0.346 (-2.110)	0.000 (-0.018)	-
自由度修正済決定係数	0.950	0.946	0.842	0.799	0.996	0.975	0.994	0.891
観測数	34	34	34	34	34	34	34	34
モデル	ランダム	ランダム	ランダム	ランダム	固定	固定	ランダム	ランダム

注: カッコ内はt値

(2002.4.16受理)

An Idea for Strategic Evaluation of Official Development Assistance

Nobuhiro Mori Mitsuru Aotsu

KRI International Corporation
a1351@n-koei.co.jp a5648@n-koei.co.jp

Abstract

The purpose of this study is to show an idea for sector-based ODA evaluation. The author examines what aid works, and what doesn't with the case study of Bangladesh.

ODA strategy is defined as long-term basic design of ODA project. It is critical for feasibility of the strategy that corroboration of resource allocation is incorporated in the design. Key of the strategy is to create good ideas for sectoral ODA allocation based on selection and focus, and for combination of ODA type as well. Recent investigations have demonstrated that accumulation of each project-based evaluation is not equivalent to ODA evaluation due to fungibility of project aid. In addition, project-based evaluation is usually concluded by itself. Accumulation of project-based evaluation, that is to say, tactical level evaluation does not lead to strategic ODA evaluation. The author would like to emphasize that policy makers should make clear priority of each ODA sector in perspective at first. Then, select projects in each sector from the compound viewpoint.

Keywords

Strategic ODA Evaluation, Impact Evaluation in the Light of Macro-economy,
Sector/Type-based ODA Evaluation, Selection and Focus, Resource Allocation

【実践・調査報告：依頼原稿】

Reflections on a Career in Evaluation

Charles Gene Lyle

Ramsey County Community Human Services Department
St. Paul, Minnesota

Gene.Lyle@CO.RAMSEY.MN.US

Abstract

The author reflects on his 27 year history as an internal evaluator in a large government social service agency, with particular emphasis on how and why they began doing evaluation, the evolution of their evaluation methodology over time, the impact of external forces on that evolution, and some of the specific distinctions between their original approach and their more recent systems. A "lessons learned" discussion may have particular value to evaluators beginning their careers or moving into evaluation in government settings. Emphasis is placed on interpersonal skills and consensus-building, taking advantage of technology, and the growth in understanding that evaluation has become an expected part of service administration and management.

Keywords

Utilization-Focused Evaluation, Evaluation client, Outcome evaluation,
Management Information Systems, Evaluation history, Accountability

Background

I entered the evaluation field around 1974 before the field really existed. I had spent a fair amount of time in my undergraduate years at the University of Minnesota doing research in education and with children with serious physical disabilities. After coming to Ramsey County, I worked as a Vocational Counselor helping people on welfare get jobs. However, because of my background in research I was asked by my agency to take on the task of evaluating certain social service programs under contract to the agency to provide these services. The incentive for this came not from the fact that people were "catching on" to the idea of doing evaluation - we hadn't progressed that far yet - but from a concern regarding a particular contract agency about which our County Board of Commissioners were receiving complaints from citizens. The Commissioners wanted to know what was going on in that program and what should be done about it.

Because of this incident, and because of a budding interest in evaluation in various other agencies and governmental units in Minnesota, my agency realized that it needed a process for determining if the services we were buying from community agencies, as well as the services we provided ourselves, were working. Were the people we served getting the intended benefit from those services? To answer these questions we developed the Office of Research and Evaluation. It was my good fortune to help establish that office and to continue contributing to it today. Our initial efforts, and our continued work, I should point out, are heavily influenced by Dr. Michael Quinn Patton's Utilization-Focused Evaluation teachings.

Initially, we outlined three major purposes for our work:

- Establish accountability. Be sure that programs are accomplishing what they are supposed to accomplish.
- Improve program performance. Provide information that can help programs be more effective.
- Inform planning and funding decisions. Assist others in making decisions by providing results about effectiveness and other information that can help determine service levels, funding, etc.

Our first step after setting these goals was to develop and pilot-test a methodology for gathering evaluative information from our contracted service providers as well as from our own internal programs. Fundamentally, this methodology involved the establishment of goal statements for each program, a description of the services they were to provide and the clients who were to receive the services. Specific outcome objectives were developed, measures for these objectives defined, and expected outcomes were set. A key feature of this approach is that each program was evaluated by measuring its actual performance against the performance expected for that program. Evaluation reports were produced generally every calendar quarter. The system was designed specifically to treat each program as unique. It was deliberately designed to NOT compare separate programs against one another or against some outside benchmark. We did it this way simply because we felt we did not have the technology, resources, or knowledge to enable cross-program comparisons. Later in this article I will touch on this early model again and contrast it to the approach we are taking now.

System Evolution

This approach held up well during the first few years. It was seen as credible and reliable by our administration and our Commissioners. Annual outcome reports were produced that summarized evaluation outcomes, along with the inclusion of activity and effort information such as counts of clients served, etc. But as the number of contracts and programs increased, this individual approach - developing unique evaluations for each and every contracted and direct service - became unwieldy and expensive. It was also suffering from the inevitable boredom and routine that sets in after time in any system. Because of these and other issues, it became clear that it was time for a change.

Let me pause here for a moment and mention that my years in the field have convinced

me that nothing stands still. There are subtle and obvious changes occurring all the time and one cannot allow oneself to be forever committed to a particular evaluation methodology or system. Factors over which evaluators have little or no control, such as changes in the law, administration, and technology, will inevitably create a need for change in the way we do evaluation.

As we began to realize that our original system was gradually losing its relevance and ability to supply useful information, we began to redefine our approach. The following table summarizes the differences between our original approach and the one we are taking now.

The Old	The New
Effort & Outcome	Process & Outcome
Program Based	Client Based
Paper & Pencil	Electronic
Non-Relational	Relational
Data Source: Others	Data Source: Us
Negotiated	Mandated
Wide but Shallow	Potentially Deep
Inflexible	Flexible
Aggregate Level of Analysis	Individual Level of Analysis
Individualized Measures	Standardized Measures
Program-level Decisions	Policy-level Decisions
Static	Dynamic

With apologies for some redundancies in the following, let me expand a bit on these differences.

Effort & Outcome/Process & Outcome: In the past we could look at numbers served, dollars spent, and expected outcomes achieved. In the new system the focus is more on how services are provided and how that influences outcomes. Outcomes for the client still remain at the core of the system, however.

Program Based/Client Based: Although we described our original system as being client outcome oriented, the outcomes were nearly always stated and reported for programs as a whole rather than for individual clients. The new system will allow client level analysis.

Paper & Pencil/Electronic: Electronic resources were seldom available when we began. We now have the resources necessary to collect and maintain data in this fashion, and this adds considerable power to our work. This is a major benefit of the growth of technology in government services in recent years.

Non-Relational/Relational: Our future capacity will be built around a comprehensive relational data base, or more accurately, a set of data bases that will all "talk" to one another, that will incorporate virtually everything we need to know about our clients and what happens to them. This will allow us, for example, to pull data from our accounting system and tie it to information

from our child placement data base, to do comparative cost analysis, such as between children in foster care versus children in group homes. This capacity is really at the heart of the new approach.

Data Sources: Others/Data Sources: Us: Our original approach was highly dependent on obtaining information directly from the programs providing services, generally in the form of aggregate reports. The new system collects client data independently and this data exists within our own computer systems. We can associate that data with individual programs so we do not need to rely on those programs to supply it.

Negotiated/Mandated: We will continue to require some programs to provide information directly to us to supplement data we can generate in our own system. And to a much greater extent now we will be requiring similar programs to collect the same information rather than negotiating separate measures for each program. In many cases we will use standardized and widely available instruments for this purpose.

Wide but Shallow/Potentially Deep: In the past we could look rather superficially at all of the services we provided or purchased but couldn't easily "drill down" into any single program or target group for more information. We will be able to do this now, principally because there will be so much more data available to us.

Inflexible/Flexible: Our original system locked us into a rigid process for using information, mainly because it often did not give us direct access to data on individual clients. Our new system will give us great access to this data and will allow us to manipulate it in nearly unlimited ways.

AggregateLOA/ Individual LOA: Judgements about the value, merit, or worth of services were limited in our original system to the program level. In the new system we will be able to get deeper, right down to the client level. This means that we will be able to do things like track changes in client status over time, no matter what combination of services they may receive. This is a huge advantage over the original approach.

Individualized/Standardized: As noted earlier, most evaluations were individualized to each program, and it was not unusual to find similar programs using very different measures. The new system will require more consistency across programs and services. Data on children in substitute care, for example, will be very similar even if they are placed in different settings.

Static/Dynamic: The new system can be more responsive to changes in information needs since essentially all the data will be at our disposal; it will just be a question of how we manipulate it. Changing the old system, on the other hand, could not be accomplished rapidly.

Program Level/Policy Level Decisions: We hope that the information available in the new system will be more valuable for making decisions at the policy level. In the old system, decisions about individual program performance could be made but often the data available could not be effectively incorporated into higher level decisions because it could not always be well integrated

with other data.

To sum it up, the major difference between the old and the new is our ability, thanks to new technological resources, to evaluate at the client level, rather than at the program level. We can now follow clients as they move through our system, look at the mix of services they receive, and assess the overall impact of these services on their lives.

I hasten to point out that certain characteristics of our original system are carried over. Most importantly, the utilization focus remains in place. It is our belief that it is pointless to do evaluation in the first place if the result of your work is of little or no use. We also remain strongly committed to an outcome focus. We still need to address fundamental questions like "Are clients benefiting from services?" "Are services achieving what we want them to achieve?" "Are we getting the most value for our money?" We continue to produce regular reports as we have in the past so that continuous attention is paid to those issues in which management and staff are concerned. And, as much as possible, we still rely on the involvement of those who actually provide the services to aid in developing and reviewing our efforts.

Lessons Learned and Lessons to be Learned

So what is the point of all of this history, and what have I learned from it? It has taught me that as you move through a career in evaluation you will always need to be flexible, to be aware of changes going on around you that impact your work, and to be ready to take advantages of opportunities that arise for you to enhance your own skills and effectiveness. It has also taught me that one of the most important skills an evaluator can have is the ability to interact with people in an open and friendly manner. Programs that have little or no exposure to evaluation are often suspicious of the motives of evaluators and anxious about what the evaluator is going to do "to" them rather than "for" them. This is particularly true where the requirement for evaluation is coming from outside the program itself, such as from a government or funding agency. Defusing this fear and anxiety requires patience, cooperation and understanding. The Utilization-Focused approach, and others that emphasize interaction, negotiation, and clarity such as Empowerment Evaluation, are particularly effective approaches to take in cases like this.

The time spent in reaching understanding and consensus in the early stages of evaluation development will pay off in the long run with a system that all parties understand and support. This is true whether or not the evaluator is someone coming in from outside the organization or is an internal evaluator. The same anxieties and suspicions exist in both cases.

A second lesson I have learned is to take advantage of technological advancements as they become available. In my setting the technological resources we have today, for example, were not available twenty years ago and they have added immensely to our ability to efficiently acquire, analyze, and use data. It is almost a given, now that electronic resources will be available to evaluators. Be prepared to work closely with those responsible for developing and maintaining these resources. The connection between management information systems and evaluation will soon emerge as an important one in our field. It could even become a sub-specialty.

Something else is available today that didn't exist in earlier years and that is the expectation that evaluation will actually take place. I believe you are seeing this now in your own country, especially with regard to evaluation in education. As in Japan, much of the incentive for this in the US comes from government and legislative requirements. But I believe it also comes from the fact that evaluation has demonstrated its usefulness in understanding and demonstrating how programs work. Not all decision-makers understand or use evaluation, and certainly no decisions are made solely on the basis of evaluation information. But it is a tool that did not exist twenty or thirty years ago and as a tool it has become part of every knowledgeable modern decision-makers tool kit. It is our responsibility as evaluators to make sure those tools are the best ones for the job.

(2002.4.16受理)

編集委員会

1. 会員からの活発な投稿をお願いいたします。

学会誌は会員間の重要な研究成果や評価調査経験の交流の場です。平成14年4月からは、「行政機関が行う政策の評価に関する法律」(行政評価法)(平成13年法律第86号)が施行されるようになるなど、評価が行政活動の中でますます重要な位置を占めるようになってきており、日本でも評価文化が徐々に育ってきていることを実感いたします。このような状況を考えますと、学会誌の充実がわが国の評価の研究水準とともに評価の実践の水準を向上させるものであり重要なことでもあります。そのためにも、会員の皆様からの一層活発な良質な投稿をお願いいたします。原稿の種類は、「総説」、「論文」、「調査報告」、「研究ノート」のカテゴリーがあります。詳しくは、学会誌の執筆要領をご覧ください。

学会誌は、年2回発行いたします。第2巻2号を9月発行しますので、不規則ながらも年2回発行することに努めました。第3巻からは3月と9月の規則的な発行体制が整いつつあります。来年3月発行の第3巻1号の投稿締め切りは2002年9月末日です。皆様ふるって投稿をお願いいたします。

また、学会誌掲載原稿は、評価学会会員の研究成果と評価調査経験を広く問うとともに、これらを基に広く範囲に亘る評価の議論を期待しホームページにも掲載しております。

なお、第2巻2号から投稿原稿以外に特集を組み、少しでも学会誌の内容の充実に努めることにいたしました。第2巻2号は「政策評価制度の現状と展望」を、第3巻1号は「ODAにおける評価の新しい潮流(仮題)」を組むことにしております。評価の実践の動向を踏まえつつ、テーマを選択していきたいと考えております。特集は基本的には特集テーマにあった専門家に原稿を依頼していくことにしております。

2. 執筆要領と審査プロセスについて

学会誌の投稿締め切りはおおよそ発行の6ヶ月前になります。具体的には、学会ホームページに掲載いたします。投稿する方は、学会誌巻末の執筆要領をよくお読みください。また、執筆要領の修正は適宜学会ホームページに掲載しお知らせいたします。なお、英文原稿の英語を一定水準に保つために、投稿前に英文校閲など英文チェックを済ませておくことをお勧めします。

審査にあたっては、常任編集委員会は1論文ごとに担当編集委員を決め、原則2名の査読者を選定し、執筆者の名前を伏せて査読をお願いしております。担当編集委員は査読結果を踏まえた判断を常任編集委員会に提案し、審議を経て、採否あるいはリライトの要請を決めております。リライト原稿は、要請が満たされない場合には採用されません。審査結果の通知あるいはリライト要請は、原稿締め切りの約2ヶ月後になります。採用されなかった原稿についても、コメントに基づき再度リライトされ再投稿されることをお勧めいたします。

なお、編集委員についても投稿を認めておりますが、この場合も一般投稿と同様の扱いがされ、投稿者である当該編集者はその論文の審査プロセスにはまったく関与しない体制をとっております。また、特集テーマでの依頼原稿についても、原則査読を行うこととしております。

3. 投稿資格について

投稿するには日本評価学会会員であることが必要です。(筆頭執筆者が学会会員であれば可) 非会員が投稿される場合には、まず会員申し込みを行い投稿してください。皆様の知り合いの非会員が投稿を考えているような場合には、投稿される方にそのようにお伝えください。

4. 会員の皆様への査読依頼について

投稿原稿については、前述のように査読者を2名お願いし学会誌に適合しているか審査しております。その際、お願いいたします会員には、事前のご連絡なしに査読依頼状と査読対象原稿をお送りしております。依頼がありました場合には、会員の責務とお考えいただき、お引き受けいただきますようお願い申し上げます。担当編集委員として広く専門家の方に査読をお願いできるように、また、査読者として自由にご自分のご意見を表明できるように、加えて、編集業務の簡素化と期間短縮のため、事前了解をお取りしておりません。この点は是非ご理解いただきますようお願い申し上げます。なお、事情により査読が不可能な場合には、同封の返信用封筒(切手つき)で至急ご返送下さるようお願い申し上げます。2名の査読者からの査読結果を基に担当編集委員が、常任編集委員会に判定について提案し、審議を経て採否を決めております。リライトがある場合には、担当編集委員は2名の査読者の査読結果を最大限にいかして整合性のあるリライトのためのコメントを作成いたします。提出されたリライト原稿は、担当編集委員が審査を行い、常任編集委員会の合意を得て、投稿者ご本人に通知いたします。査読者の査読結果は、判定のための最も重要な意見となりますが、最終判断は常任編集委員会が行うという点をご理解ください。従いまして、査読者の方には査読結果受領のご連絡をお送りいたしますが、審査結果についてはお知らせすることをいたしておりませんので、この点あらかじめご理解いただくようお願い申し上げます。

5. 編集委員長等の交代と常任編集委員会の再構成

本年3月25日に開かれた日本評価学会の第12回理事会で研修委員会の設置が決定し、長尾編集委員長が研修委員会の委員長になられました。これに伴い、下記の通りの編集委員長の交代と常任編集委員会の再構成を行いました。

新常任編集委員会の構成：

委員長：三好皓一（国際協力事業団）

副委員長：古川俊一（筑波大学）

副委員長：村松安子（東京女子大学）

常任委員：牟田博光（東京工業大学）

常任委員：山谷清志（岩手県立大学）

当体制は、編集委員の任期は2年（全国大会を節目と考える）であるため、今期の残り期間である2002年12月7日、8日（第3回全国大会）までといたします。

編集委員長 三好 皓一

研修委員会

活動方針

本年3月、新たに研修委員会が設置された。この委員会では、「評価の普及および評価に関する人材育成に寄与すること」を目的として、各種セミナーの開催、研修プログラムの開発、教材の作成、講師の派遣、その他人材育成に関わる活動の実施支援を行う予定である。

当面の人材育成の対象として「自治体事業評価人材」と「ODA事業評価人材」の二つを掲げ、事業現場のニーズに則したセミナーの実施、研修プログラムの開発等を推進していく。また、委員会の活動は評価の学習や実習を主たる内容とするので、学生会員の積極的な参加を促すとともに、個別事業の実施に際しては広く当該分野・テーマを専門とする学会員各位の協力を求め、人材育成の側面から日本社会へ「評価文化」を発信する学会としての役割を果たしていくこととしたい。

当面の活動予定

第1回評価セミナーの開催 テーマ：「組織改革における評価の活用について」

研修委員会の設置を記念する評価セミナーを、国際協力銀行と広島大学教育開発国際協力研究センターとの共催で6月5日に開催した。講演者として、評価のエキスパートであるDr. John M. Owen（メルボルン大学）とDr. Elizabeth Jane Davidson（ウエスタン・ミシガン大学）を迎え、組織改革における内部評価と外部評価の組み合わせについて論じてもらうとともに、組織改革の手段としての評価研修に関し学会員との意見交換を行うものである。

研修モデル・プログラムの開発

自治体およびODA事業評価研修のモデル・プログラムの開発を予定している。自治体に関しては、モデル自治体を選定し評価人材養成プログラムを共同で開発するとともに、関係者との意見交換を通して自治体の評価をとりまく現状に対する理解を深めることも目的としている。

ODA事業に関しては、今年度は技術協力を取り上げ、評価調査に携る人材の実践的知識習得をめざした研修プログラムの開発を行う。そのプロセスにおいて、ODA事業実施主体（評価委託元）と評価受託者（コンサルタント等）の間の意見交換を目的としたワークショップも予定している。

研修委員会共同委員長 長尾 眞文
源 由理子

国際交流委員会

活動状況

国際交流委員会では、昨年夏に開設の「国際交流委員会メーリングリスト（ML）」も定着し活動方針を含む種々の議論がなされている。基本方針である「評価手法の精緻化に資すること」および「評価における国際交流・協力の日本側窓口となること」に従って、関連文献の整理に加え、最も基本となる評価用語を整備するため、ML参加者の協力を得て以下の共同作業が進行している。

- (1) 経済協力開発機構開発援助委員会評価用語集暫定版1の翻訳
- (2) 用語集翻訳に伴う関連用語の整備

分野別研究会の立ち上げについては、これまでに提案がなされていた5つの分野（環境、教育、保健医療、農業、評価手法）にこだわることなく各参加者の立場から自由に発言頂き、議論の牽引役からいくつかの問題提起もなされた。なおこのMLは、第二回全国大会総会にて紹介後も試験運用継続状態にあり、参加者が開設当初の40名程度にとどまっているのが現状である。今回、改めて紙面にてこのML（他機能も有するグループシステム）を紹介し、広く学会員に参加を呼びかけたい（日本評価学会員限定の登録制、投稿記事等は外部非公開にて運用）。

日本評価学会国際交流委員会グループシステム（以下URLより参加登録）

- ・ <http://www.egroups.co.jp/group/iJES-ml>
- ・ ML（同報メール）、データベース（用語集、IDEASに対する質問・提案・コメント）、予定表（学会スケジュール、イベント予定等自動配信）、ファイル共有（文献リスト、配布資料等保管）、その他（投票機能、関連情報へのリンクなど）

なお、廣野が参加している国際開発評価学会準備運営委員会（IDEAS/ISC: International Development Evaluation Association / Interim Steering Committee）では、昨年7月のワシントンの世界銀行における第1回会合と本年1月におけるニューヨークのUNDP本部における第2回会合にて、国際開発政策・プログラム・プロジェクト評価手法の科学的精緻化・体系化を促進し、国際標準の樹立に努めること、および、開発途上国の開発政策・プログラム・プロジェクト評価能力体制の強化を支援することを基本目的とし、これら事業に貢献できるIDEASの定款、学会理事選出規定、年次事業計画と予算計画の基本方針の作成、さらにIDEAS設立総会の議題の設定が進められている。国際交流委員会ではISCの支援活動に協力頂ける方を随時募集しており、各位の積極的参加を期待している。上記MLにて各人が個々に有する関連情報を提供・共有頂き、これらを元にIDEAS/ISCに対する質問や提案を頂ければ誠に有り難い。

今後の活動について

今後もML参加者の協力を得て継続的な活動を展開し、進行中の共同作業等については全て学会員に公開する。共同作業による成果物については基本的に一般公開する方向にて関係各方面と協議調整を行う予定である。引き続き、学会員の積極的な参加と、支援・協力を期待したい。

国際交流委員長 廣野 良吉

日本評価学会第3回全国大会について

日本評価学会は、今年12月7～8日成蹊大学にて第3回全国大会を開催することとなりました。いずれの学会でも全国大会は会員の研究活動に発表の場を提供することと、会員総会を通じて学会の運営に対する会員の総意を汲み、学会の今後の発展を期すために開催されるものです。本学会の第3回全国大会でも第1回、第2回と同様に、共通論題・自由論題セッションを設けて、2日間という限られた時間内であるが、出来るだけ多くの会員に研究の成果を世に問うて、切磋琢磨できるようなプログラムの作成に意を尽くしました。その結果、本年度の全国大会では自由論題のみならず、共通論題でも会員の皆さんの研究発表募集に踏み切りました。さらに、今年は学会のもう一つの役割である「社会への問題提起」という面で工夫して、大会の共通論題として、「1990年以降における日本の行政・経済・社会改革の評価と展望」を取り上げました。

昨秋来第3回全国大会実行委員会としては当初、今年は第2次世界大戦後の占領から解放されて政治的独立を果たした1952年から丁度50周年を迎えることを記念して、戦後50年間の日本の政治・経済・社会改革の実態を評価して、次の50年間に日本が国内外にて進むべき方向を見出すという企画を持っていました。しかし50年というのは2日間の大会で討議するには余りにも長期ということで、1990年以降の行政・経済・社会改革の成否を厳しく評価して、今後を展望することとしました。そのために、全国大会初日にシンポジウムを開催して、日本社会の主要な構成員である政界、官界、産業界、学界、市民社会を代表する人々に参集して貰い、日本評価学会の会員の皆さんとの討論は勿論のこと、一般市民にも開放して広く民意を問い、会員の皆さんとの意見交換の場を提供することにしました。

さらに、会員の皆さん各自にあっては自分が選択する共通論題セッションに出来るだけ参加できるよう、共通論題セッション時には、自由論題セッションを一切設けないことにしました。また、自己選択によって他の共通論題セッションには参加できなかった会員の皆さんのために、総ての共通論題セッションの座長による報告会を午前午後と2回に分けて設けるよう工夫しました。なお、2日間の全国大会運営では、総ての学会理事の皆さんに共通論題セッション司会や自由論題セッション座長という大役をお引き受けいただきましたことに心から御礼申し上げます。日本評価学会会員の皆さんには、成蹊大学で開催される第3回全国大会では、自由・共通論題セッションでの研究発表への応募は勿論のこと、シンポジウム、自由・共通論題セッション、総会、懇親会等総てに積極的に参加なされんことを希望して止みません。

日本評価学会副会長 / 第3回全国大会実行委員長 廣野 良吉

評価関連ウェブサイトリスト

2002年5月15日現在

1. 各省における政策評価・行政評価のリンク

内閣府

- ・内閣府のHP
<http://www.cao.go.jp/>
- ・内閣府本府の政策評価について
<http://www8.cao.go.jp/hyouka/index.html>
- ・内閣府本府政策評価基本計画（平成14年4月）
http://www8.cao.go.jp/hyouka/h14_kihon.pdf

総務省

- ・総務省のHP
<http://www.soumu.go.jp>
- ・行政機関が行う政策の評価に関する法律について
（法律の骨格）
<http://www.soumu.go.jp/kansatu/hourituan.htm>
（全文）<http://www.soumu.go.jp/kansatu/houritu.htm>
- ・行政評価局トップページ（最新ニュース等）
<http://www.soumu.go.jp/kansatu/index.htm>
- ・政策評価に関する標準的ガイドライン（平成13年1月15日）
<http://www.soumu.go.jp/kansatu/gaido-gaiyou1.htm>
- ・政策評価制度の法制化に関する研究会報告（平成12年12月）
<http://www.soumu.go.jp/kansatu/houseika.htm>
- ・政策評価制度の在り方に関する最終報告（平成12年12月11日）
<http://www.soumu.go.jp/kansatu/matome1.htm>
- ・政策評価の手法等に関する研究会
<http://www.soumu.go.jp/kansatu/kaisai-moku.htm>

法務省

- ・法務省のHP
<http://www.moj.go.jp>
- ・法務省の政策評価及び実施庁の評価
<http://www.moj.go.jp/KANBOU/HYOUKA/hyouka01.html>
- ・法務省政策評価に関する基本計画
<http://www.moj.go.jp/KANBOU/HYOUKA/hyouka07.html>

外務省

- ・外務省のHP
<http://www.mofa.go.jp>
- ・外務省における政策評価
<http://www.mofa.go.jp/mofaj/annai/shocho/hyouka/index.html>

- ・ 外務省における政策評価の実施要領

<http://www.mofa.go.jp/mofaj/annai/shocho/hyouka/youryou.html>

- ・ ODA評価報告

http://www.mofa.go.jp/mofaj/gaiko/oda/siryo/siryo_3/siryo_3f.html

- ・ 「ODA評価研究会」報告書（平成13年2月）

http://www.mofa.go.jp/mofaj/gaiko/oda/siryo/siryo_3/siryo_3_2.html

- ・ 「評価研究作業委員会」報告書（平成12年3月15日）

http://www.mofa.go.jp/mofaj/gaiko/oda/siryo/siryo_3/siryo_3_1.html

財務省

- ・ 財務省のHP

<http://www.mof.go.jp>

- ・ 財務省の政策評価

<http://www.mof.go.jp/jouhou/hyouka/14nendo/top.htm>

- ・ 財務省の「平成14年度政策評価実施計画」と「政策評価に関する基本計画」について

<http://www.mof.go.jp/jouhou/hyouka/14nendo/point.html>

文部科学省

- ・ 文部科学省のHP

<http://www.mext.go.jp>

- ・ 文部科学省の政策評価

http://www.mext.go.jp/a_menu/hyouka/index.htm

- ・ 平成14年度文部科学省政策評価実施計画

http://www.mext.go.jp/a_menu/hyouka/jissi/020301.htm

厚生労働省

- ・ 厚生労働省のHP

<http://www.mhlw.go.jp>

- ・ 厚生労働省の政策評価

<http://www.mhlw.go.jp/wp/seisaku/index.html>

- ・ 平成14年度厚生労働省実施計画（厚生労働省における事後評価の実施に関する計画）

<http://www.mhlw.go.jp/houdou/2002/04/dl/h0403-2c.pdf>

農林水産省

- ・ 農林水産省のHP

<http://www.maff.go.jp>

- ・ 農林水産行政に係る政策評価の実施について

<http://www.maff.go.jp/mlet/191-2.html>

- ・ 農林水産省政策評価基本計画（H14.3.29決定）

<http://www.maff.go.jp/soshiki/kambou/kikaku/hyoka/kihonkei.pdf>

- ・ 農林水産省政策評価実施計画（H14.3.29決定）

<http://www.maff.go.jp/soshiki/kambou/kikaku/hyoka/jishikei.pdf>

経済産業省

- ・ 経済産業省のHP
<http://www.meti.go.jp>
- ・ 経済産業省の政策評価
http://www.meti.go.jp/policy/policy_management/main_04.html
- ・ 経済産業省の技術評価
http://www.meti.go.jp/policy/tech_evaluation/toppuindex.html
- ・ 経済産業省政策評価基本計画
http://www.meti.go.jp/policy/policy_management/kihonkeikaku/kihonkeikaku.pdf
- ・ 平成14年度経済産業省事後評価実施計画
http://www.meti.go.jp/policy/policy_management/kihonkeikaku/jigohyoukakeikaku.pdf

国土交通省

- ・ 国土交通省のHP
<http://www.mlit.go.jp>
- ・ 国土交通省の政策評価
<http://www.mlit.go.jp/hyouka/>
- ・ 国土交通省政策評価基本計画（平成14年度～平成18年度）
<http://www.mlit.go.jp/hyouka/pdf/kihon.pdf>
- ・ 平成14年度事後評価実施計画
<http://www.mlit.go.jp/hyouka/pdf/h14jisshi.pdf>

環境省

- ・ 環境省のHP
<http://www.env.go.jp>
- ・ 環境省の政策評価
<http://www.env.go.jp/guide/seisaku/index.html>
- ・ 評価結果の反映状況
<http://www.env.go.jp/press/press.php3?serial=2929>

会計検査院

- ・ 会計検査院のHP
<http://www.jbaudit.go.jp>
- ・ 会計検査院研究
<http://www.jbaudit.go.jp/gar/>

2．ODA評価のリンク

国際協力事業団（JICA）

- ・ 国際協力事業団のHP
<http://www.jica.go.jp/>
- ・ JICAの評価
<http://www.jica.go.jp/evaluation/index.html>

- ・事業評価年次報告書（平成13年度）
<http://www.jica.go.jp/evaluation/general13/index.html>
- ・事後評価報告書（全件リスト）
<http://www.jica.go.jp/evaluation/after/index.html>
- ・事業事前評価表（平成13年度）
<http://www.jica.go.jp/evaluation/index.html>
- ・「事前事業評価表」について
<http://www.jica.go.jp/evaluation/before/>

国際協力銀行（JBIC）

- ・国際協力銀行のHP
<http://www.jbic.go.jp/>
- ・円借款案件事後評価報告書2001（全文版・要約版一覧）
http://www.jbic.go.jp/japanese/oec/post/A56/2001_index.php
- ・事業事前評価（平成13年度）
<http://www.jbic.go.jp/japanese/oec/before/index.php>
- ・「事業事前評価表」の作成、公表について～円借款事業の透明性向上のために～
<http://www.jbic.go.jp/autocontents/japanese/news/2001/000034/index.htm>

3. 自治体リンク

- ・全国の都道府県の政策評価・行政評価のリンク集
<http://www.idcj.or.jp/jes/todoufukun.htm>
- ・総務省行政評価ライブラリー「地方公共・非営利組織団体における行政評価の取組状況」
<http://www.mha.go.jp/click/003.html>

4. 世界の評価学会リンク（ウェブページを開設している学会のみを掲載しました）

このリストは、全米評価学会のThe International & Cross-Cultural Evaluation Topical Interest Group（I&CCE）が作成したリストを参照した。（<http://home.wmis.net/~russon/icce/eorg.htm>）

南北アメリカ

- ・全米評価学会(American Evaluation Association)
<http://www.eval.org/>
- ・カナダ評価学会(Canadian Evaluation Society)
<http://www.evaluationcanada.ca/>

ヨーロッパ

- ・ヨーロッパ評価学会(European Evaluation Society)
<http://www.europeanevaluation.org/>
- ・イギリス評価学会(UK Evaluation Society)
<http://www.evaluation.org.uk/>
- ・イタリア評価学会(Associazione Italiana de Valutazione)
<http://www.valutazione.it/>

- ・ドイツ評価学会(Deutsche Gesellschaft für Evaluation)
<http://www.degeval.de/>
- ・フランス評価学会(La Société Française de l'Évaluation)
<http://www.sfe.asso.fr/>
- ・スイス評価学会(Swiss Evaluation Society)
<http://www.seval.ch/>
- ・ロシア国際プログラム評価ネットワーク(Russia/NIS)
(International Program Evaluation Network (Russia/NIS))
<http://ipen21.org/ipen/en/default.html>

アジア（大洋州含む）

- ・オーストラリアン評価学会(Australasian Evaluation Society)
<http://www.aes.asn.au/>
- ・マレーシア評価学会(Malaysian Evaluation Society)
<http://www.angelfire.com/ab/mes/>
- ・スリランカ評価学会(Sri Lanka Evaluation Association)
<http://www.naresa.ac.lk/sleva/profile.htm>
- ・日本評価学会(Japan Evaluation Society)
<http://www.idcj.or.jp/jes/JES.htm>

アフリカ

- ・アフリカ評価学会(African Evaluation Association)
<http://www.geocities.com/afreval/>
- ・ナイジェリア・モニタリング&評価ネットワーク(Reseau Nigerien de Suivi et Evaluation)
<http://www.ird.ne/rense/introuk.html>

5. 世界の評価機関リンク

- ・経済協力開発機構 開発援助委員会 評価作業部会(OECD DAC Working Party on Aid Evaluation)のトップページ
<http://www.oecd.org/dac/Evaluation/index.htm>
- ・世界銀行 Operations Evaluation Department
<http://www.worldbank.org/oed/>
- ・アジア開発銀行 Operations Evaluation Office
<http://peo.asiandevbank.org/>
- ・UNDP Evaluation Office
<http://www.undp.org/eo/>
- ・DAC加盟国の評価部署のリンク
<http://www.oecd.org/dac/Evaluation/htm/evallinkmem.htm>
- ・DAC評価5項目
<http://www.oecd.org/dac/Evaluation/htm/evalcrit.htm>
- ・PARC(Performance Assessment Resource Centre) (DfIDが設立した評価機関)
<http://www.parcinfo.org/default.asp>

日本評価学会規約

2000年9月25日設立総会承認

第1章 総則

(名称)

第1条 本会は、日本評価学会と称し、英語名は、The Japan Evaluation Society (略称JES) とする。

(所在地)

第2条 本会の本部は、東京都江東区に置く。ただし、その他の地に支部を置くことができる。

(目的)

第3条 本会は、評価に関する研究及び応用を促進し、会員相互及び関連学・協会との情報交換を図るとともに、この分野の学問の進歩発展及び評価に携わる人材の育成を通じ、評価活動の向上と評価の普及を目的とする。

(事業)

第4条 本会は、前条の目的を達成するために次の事業を行う。

- 1) 評価に関する学術研究会、講演会、国際シンポジウムなどの開催
- 2) 評価に関する研究助成、報奨
- 3) 評価に関する調査及び研究並びに研修
- 4) 学会機関誌及びその他の刊行物の発行等の普及啓発活動
- 5) その他、本会の目的を達成するために必要な事業

第2章 会員

(会員の種別及び入会)

第5条 本会の会員の種別は、次の通りとする。総会での議決権は、正会員のみが持つものとする。

- 1) 正会員 本会の目的に賛同して入会した個人。
 - 2) 学生会員 本会の目的に賛同して入会した、原則として大学以上の学生で、学生会員を希望する者。尚学生会員は、卒業と同時に正会員となることができる。
 - 3) 賛助会員 本会の目的に賛同し、その事業を後援する者。
 - 4) 名誉会員 本会に功労のあった者及び広く評価分野に関連ある分野における学識経験者で理事会の推薦に基づき総会の承認を経た者。
- 2 本会の会員になろうとする者は、入会申込書を会長に提出し、理事会の承認を得なければならない。また、名誉会員として推薦された者は、入会の手続きを要せず、本人の承諾をもって会員となるものとする。

(会費)

第6条 会員は、総会の定めるところにより、会費を負担しなければならない。

(退会)

第7条 会員は、退会しようとするときは、事前にその旨を書面をもって会長に届け出なければならない。

2 会員は、次の各号のいずれかに該当するときは、本会を退会したものとみなす。

- 1) 正会員又は学生会員にあっては、会費を2年以上滞納した場合

- 2) 本会の名誉を毀損し又は設立の趣旨に反する行為を行なったことにより、総会において、除名すべきものと認められた場合
- 3) 前各号に掲げるほか会員たる資格を喪失した場合
- 3 本会は、会員がその資格を喪失しても、既に納入した会費その他の拠出金品は返還しない。

第3章 役員等

(役員の種類)

第8条 本会には、次の役員を置く。

- 1) 理事 20名から25名(うち会長1名、副会長若干名)
- 2) 監事 2名

(選任)

第9条 理事及び監事は、総会において、正会員の中から選任する。

- 2 会長及び副会長は、理事会において、理事の互選により定める。
- 3 理事及び監事は、相互に兼ねることができない。

(職務)

第10条 理事は理事会を構成し、会務の執行を決定する。

- 2 会長は、本会を代表し、会務を統括する。
- 3 副会長は、会長を補佐し、会長に事故あるとき又は欠けたときは、会長があらかじめ指名した順序で、その職務を代行する。
- 4 監事は、会務の執行及び会計を監査する。

(任期)

第11条 役員の任期は、2年とする。ただし、連続3期までの再任を妨げない。

- 2 補欠又は増員により就任した役員の任期は、前項本文の規定にかかわらず、前任者又は現任者の残任期間とする。
- 3 役員は、任期満了の場合においても、後任者が就任するまで、その職務を行なわなければならない。

(幹事)

第12条 本会に幹事若干名を置く。

- 2 幹事は、会長が理事会の同意を得て任命する。
- 3 幹事は、共同して会務の執行を補佐する。
- 4 幹事の任期については、第11条第1項の規定を準用する。

(顧問)

第13条 本会に顧問若干名を置くことができる。

- 2 顧問は、会長が理事会の同意を得て委嘱する。
- 3 顧問は、本会の運営に関し、会長の諮問に答え、又は意見を述べることができる。
- 4 顧問の任期については、第11条第1項の規定を準用する。

第4章 総会

(招集)

第14条 会長は、毎年1回通常総会を招集する。

- 2 理事会が必要と認めた場合、臨時総会を招集することができる。

（開催及び議決）

第15条 総会は、正会員総数の3分の1以上の出席が無ければ開催することができない。

2 総会の議事は、この規約に別に定めるもののほか、出席正会員の過半数の同意でこれを決し、可否同数の場合は、議長の決するところによる。

3 総会の議長は、会長をもってこれにあてる。

（書面表決等）

第16条 やむを得ない理由のため、総会に出席できない正会員は、書面又は代理人をもって表決権を行使することができる。

2 前項の代理人は、代表権を証する書面を総会毎に議長に提出しなければならない。

3 第1項の場合において、正会員は、表決内容等について、総会の議長に一任することができる。

4 第1項及び前項の規定により、表決権を行使する場合は、当該正会員は総会に出席したものとみなす。

第5章 理事会

（構成及び機能）

第17条 本会に理事会を置く。

2 理事会は、理事をもって構成する。ただし、監事、顧問、幹事ならびに第30条第2項に定める事務局長は、理事会に出席し意見を述べることができる。

3 理事会は、この規約に別に定めるもののほか、次の事項を議決する。

1) 総会の議決した事項の執行に関すること

2) 総会に付議すべき事項

3) その他総会の議決を要しない会務の執行に関すること

（開催及び招集）

第18条 定例理事会は、毎年3回開催する。

2 前項にかかわらず会長が必要と認めた場合には臨時理事会を開催することができる。

3 理事会は会長が招集する。

4 会長は、緊急に理事会を招集する必要がある場合において、やむを得ない事情によりこれを開催できないときには、理事の承諾を得て、書面により議決を得ることができる。この場合、理事会は開催されたものとみなす。

（議決等）

第19条 理事会は、理事現在数の3分の1以上の出席がなければ、開会することができない。

2 理事会の議決は、この規約に別に定めるもののほか、出席理事の過半数の同意でこれを決し、可否同数のときは、議長の決するところによる。

3 理事会の議長は、会長をもってこれにあてる。

（書面表決等）

第20条 やむを得ない理由のため、理事会に出席できない理事は、書面又は代理人をもって表決権を行使することができる。

2 前項の場合において、理事は、表決内容等について理事会の議長に一任することができる。

3 前2項の規定により表決権を行使する場合は、当該理事は理事会に出席したものとみなす。

第6章 資産及び会計

(資産の構成及び管理)

第21条 本学会の資産は、次に掲げるものをもって構成する。

- 1) 財産目録記載の財産
- 2) 会費
- 3) 寄付金品
- 4) 資産から生ずる収入
- 5) 事業に伴う収入
- 6) その他の収入

2 本会の資産は、会長が管理し、その方法は理事会の議決による。

(経費の支弁)

第22条 本会の経費は、資産をもって支弁する。

(事業計画、収支予算、事業報告及び収支決算)

第23条 本会の事業計画書及び収支予算書は、会長が作成し、理事会の議決を経た後、毎事業年度の開始前に総会の議決を得なければならない。ただし、やむを得ない事情により、当該事業年度開始前に総会を開催できない場合にあっては、理事会の議決によることを妨げない。この場合、当該事業年度の開始の日から90日以内に総会の議決を得るものとする。

2 会長は、前項の事業計画書及び収支予算書を変更しようとするときは、理事会の議決を得なければならない。

3 本会の事業報告書及び収支決算書は、会長が毎事業年度終了後遅滞なくこれを作成し、監事の監査、理事会の議決を経た後、当該事業年度終了後90日以内に総会の承認を得なければならない。

(特別会計)

第24条 本会は、事業の遂行上必要がある場合には、理事会の議決を経て、特別会計を設けることができる。

2 前項の特別会計は、前条の収支予算及び収支決算に計上しなければならない。

(会計年度)

第25条 本会の会計年度は、毎年10月1日に始まり、翌年9月末日に終わる。

第7章 規約の変更

(規約の変更)

第26条 この規約は、総会において出席正会員の3分の2以上の議決を経なければ変更することはできない。

第8章 雑則

(年次大会)

第27条 本会は、毎年1回年次大会を開催する。

2 年次大会の開催は、大会実行委員長がこれを指揮統率する。

(委員会及び分科会)

第28条 本会は、事業の円滑な遂行を図るため、委員会及び分科会を設けることができる。

2 委員会及び分科会は、その目的とする事項について、調査及び研究し、または審議する。

3 委員会及び分科会の組織、構成及び運営その他必要な事項は、理事会の議決を経て、別

に定める。

（支部）

第29条 本会は、事業の円滑な遂行を図るため、支部を置くことができる。

2 支部には、理事会の同意を得て会長が委嘱する支部長を置く。

3 支部の位置、組織、運営その他必要な事項は、理事会の議決を経て、別に定める。

（事務局）

第30条 本会の事務処理のため、事務局を置く。

2 事務局には、理事会の同意を得て会長が委嘱する事務局長を置く。

3 事務局の場所、組織、職員及びその他必要な事項は理事会の議決を経て、別に定める。

附則（2000年9月25日）

1. この規約は、本学会の設立の日（以下「設立日」という。）から施行する。

2. 設立発起人は、第5条の規定にかかわらず、本学会設立当初の会員とする。設立日までに入会の申し込みを行なった者も同様とする。

3. 本学会の設立初年度の年会費は、第6条の規定にかかわらず、設立総会の定めるところによる。

4. 本学会の設立当初の役員は、第9条第1項及び第2項の規定にかかわらず、設立総会の定めるところとし、その任期は、第11条第1項本文の規定にかかわらず、2001/2002年度の事業報告及び収支決算の承認を審議する総会の日までとする。

5. 本学会の設立初年度の事業計画及び収支予算は、第23条第1項の規定にかかわらず、設立総会の定めるところによる。

6. 本学会の設立当初の事業年度は、第25条の規定にかかわらず、設立日から2001年9月30日までとする。

日本評価学会誌刊行規定

2001.9.9改訂

（目的および名称）

1. 日本評価学会（以下、「学会」という）は、評価に関する研究および実践的活動の成果を国内外の学界をはじめ評価に関心をもつ個人および機関に広く公表し、評価慣行の向上と普及に資することを目的として、「日本評価研究（仮名）」（英文仮名："The Japanese Journal of Evaluation Studies"、以下、「評価研究」という）を刊行する。（編集委員会）
2. 「評価研究」の編集は、後で定める「編集方針」にもとづいて編集委員会が行う。
3. 編集委員会は、学会会員15名以内をもって構成し、委員は学会理事会が選任する。編集委員の任期は2年とし、再任を妨げないものとする。
4. 編集委員会は、互選により委員長1名、副委員長2名および常任編集委員若干名を選出する。
5. 編集委員会は、最低年1回編集委員会を開き、編集方針、編集委員会企画、その他について協議するものとする。
6. 編集委員会は、その活動等について、随時理事会へ報告し、承認を受けるとともに、毎年1回学会年次大会の場で、過去1年の活動成果と翌年の活動計画に関する報告を行う。
7. 委員長、副委員長および常任編集委員は、常任編集委員会を構成し、常時、編集実務に当たる。

（編集方針）

8. 「評価研究」は、原則として、年2回刊行する。
9. 「評価研究」の体裁は、B5版とし、英文又は和文とする。
10. 「評価研究」に掲載する原稿（以下「論文等」という）の分類は、以下の5カテゴリーからなるものとする。
 - (1) 総説
 - (2) 研究論文
 - (3) 研究ノート
 - (4) 実践・調査報告
 - (5) その他
11. 「評価研究」への投稿有資格者は、学会会員および常任編集委員会が投稿を依頼した者とする。学会会員による連名での投稿および学会会員を主筆者とする非会員との連名での投稿は、これを認める。編集委員による投稿はこれを認める。
12. 投稿原稿を上記分類のどのカテゴリーとして扱うかは、投稿者の申請等をもとに常任編集委員会が、下記の「作業指針」に従って決定する。
 - (1) 「総説」は、評価の理論あるいは慣行について概観する論文とし、その掲載については編集委員会が企画・決定する。
 - (2) 「研究論文」は、評価の理論構築あるいは慣行の理解について重要な学問的貢献となると認められる論文とし、その採否については次項に定める査読プロセスを経て常任編集委員会が決定する。
 - (3) 「研究ノート」は、「研究論文」作成過程での理論的あるいは経験的な研究の中間的成果物に相当する論考で、その採否については次項に定める査読プロセスを経て常任編集委員会が決定する。
 - (4) 「実践・調査報告」は、評価事業の実践あるいは評価にかかわる調査の報告で、その採

否については次項に定める査読プロセスを経て常任編集委員会が決定する。

- (5)「その他」には、編集委員会が独自に企画する特集に掲載する依頼原稿や学会誌の刊行に関する編集委員会からの学会会員への連絡等が含まれる。

- 13 論文等は2名の査読者により査読することとし、その人選は編集委員会が行う。「研究論文」については、査読結果と編集委員会が査読者とは別に指名する担当編集委員1名の参考意見をもとに、編集委員会が掲載に関する決定を行う。「総説」、「研究ノート」、「実践・調査報告」および「その他」の論文については、査読結果にもとづき編集委員会が掲載に関する決定を行う。
- 14 編集委員が「評価研究」に投稿した場合には、当該委員はその投稿に係わる常任編集委員会あるいは編集委員会の議事に一切参加しないものとする。
- 15 上記いずれのカテゴリーの投稿についても、常任編集委員会による掲載の判断は可・不可の二者択一で行うこととする。但し、場合によっては編集委員会の判断で、小規模の修正による掲載も認める。「研究論文」としての掲載が適当でないと判断された場合でも、投稿者が希望すれば、常任編集委員会は「研究ノート」あるいは「実践・調査報告」としての掲載を決定できる。

(投稿要領の作成公表)

- 16 編集委員会は、上記の編集方針にもとづき投稿要領を作成し、理事会の承認を得て、広く公表する。

(配布先)

- 17 「評価研究」は、学会会員に無償で配布するほか、非会員に有償で提供する。
- 18 「評価研究」掲載論文等は、投稿者(原著者)の了承を得て全文をインターネット上で公開する。
- 19 「評価研究」に掲載された論文等の著作権は各投稿者(原著者)に帰属するものとし、編集権は本学会に帰属するものとする。

(事務局)

- 20 「評価研究」編集及び配布の事務は、それに関連する会計も含めて学会事務局が担当する。

(以上)

『日本評価研究』投稿規定

2002.3.25改訂

2001.9.9改訂

1. 『日本評価研究』(The Japanese Journal of Evaluation Studies) は、評価に関する論文、論考、調査報告等を掲載する。
2. 『日本評価研究』は、会員間の研究成果交流の場を提供し、内外における評価研究の一層の発展に資することを主目的として発行されており、原則として会員による寄稿を掲載する。なお、依頼原稿を除き、ファーストオーサーは学会員でなければならない。
3. 投稿された原稿は、編集委員会の責任において審査を行ない、採否を決定する。審査にあたっては、1原稿毎に2名の査読者を選定し、査読結果を参考にする。(査読者には、投稿者名を伏せて査読を依頼する。)
4. 原稿料は支払わない。
5. 『日本評価研究』に掲載された論文等は、その全文をインターネット上の本学会のホームページに掲載する。
6. 投稿にあたっては、投稿原稿が、研究論文、総説、研究ノート、実践・調査報告、その他のうち、どのカテゴリーに入るかを明記する。ただし、カテゴリーについての最終判断は、編集委員会で行なう。「研究論文」は評価の理論構築あるいは慣行の理解について重要な学問的貢献となると認められる論文、「総説」は、評価の理論あるいは慣行について概観する論文、「研究ノート」は「研究論文」作成過程での理論的あるいは経験的な研究の中間的成果物に相当する論考、「実践・調査報告」は評価事業の実践あるいは評価にかかわる調査の報告、「その他」は編集委員会が独自に企画する特集に掲載する依頼原稿等である。
7. 投稿方法
 - (1) 使用言語は日本語又は英語とする。
 - (2) 著者校正は原則として第一校までとする。
 - (3) 英文原稿については、ネイティブスピーカーによる英文チェックを済ませ、完全な英文にして投稿すること。
 - (4) ハードコピー4部(A4版)を提出する。その際、連絡先(住所、Tel、Fax、E-mail)と原稿の種類を明記すること。掲載可と判断された原稿については、必要なリライトを経た後に、最終原稿のハードコピー2部とDOS/Vフォーマットのフロッピーを用いたTEXTファイルを提出する。その際、オリジナル図表を添付すること。
 - (5) 刷り上がりは最大14ページとする。これを超える場合は、その経費は著者負担とする。
 - (6) 日本語原稿の最大文字数は以下のとおり。研究論文20,000字、総説15,000字、研究ノート及び実践調査報告10,000字、その他適宜。それぞれ和文要旨を400字程度、

英文要旨を150words程度、及び和文・英文でキーワード（5つ以内）を別に添付する。印刷は1ページ、20字×43行×2段（1,720字）とする。20,000字の原稿の場合、単純計算では英文要旨1ページを加えて合計13ページとなるが、図表の量によっては、それ以上のページ数となり得るので、注意すること。

- (7) 英文ではA4版用紙に左右マージン30mmをとり、10ポイントフォントを使用し、1ページ43行のレイアウトとする（1ページ約500words）。論文冒頭に150words程度のAbstractをつける。14ページでは、7,000words相当になるが、タイトルヘッド等を考慮して、最大語数を約6,000words（図表、注、文献込み）とする。図表の量によっては、ページ数が予想以上に増える場合もあり得るので、注意すること。

8. 送付先

〒135-0047 東京都江東区富岡2 - 9 - 11 京福ビル
財団法人 国際開発センター内 日本評価学会事務局
TEL: 03-3630-6994
FAX: 03-3630-8120
E-mail: suzuki.n@idcj.or.jp

『日本評価研究』執筆要領

2002.3.25改訂

1. 本文、図表、注記、参考文献等

(1) 論文等の記載は次の順序とする。

日本語原稿の場合

第1ページ：表題、著者名、所属先、Email

第2ページ以下：本文、謝辞あるいは付記、注記、参考文献

最終ページ：英文表題、英文著者名、英文所属先、Email、英文要約（150words程度）、
和文要約（400字程度）英文原稿の場合

第1ページ：Title; the author's name; Affiliation; Email address; Abstract (150 words)

第2ページ以下：The main text; acknowledgement; notes; references

(2) 本文の区分は以下のようにする。

例1（日本語）

1.

(1)

(2)

(3)

例2（英文）

1.

1.1

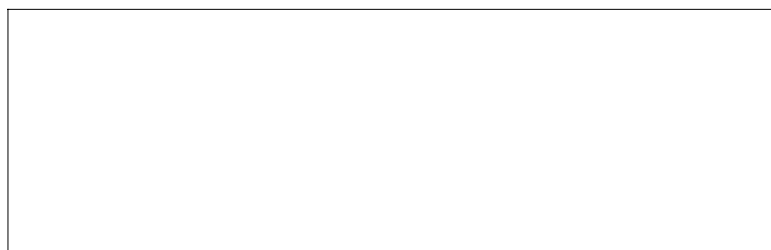
1.1.1

1.1.2

(3) 図表については、出所を明確にする。図表は原則として、筆者提出のものをそのまま写真製版するので、原図を明確に作成すること。写真は図として扱う。

例1：日本語原稿の場合

図1 州における生徒数の推移



(注)

(出所)

表1 州における事故件数

--

(注)
(出所)

例2：英文原稿の場合

--

Figure 1 Number of Students in the State of

Note:

Source:

--

Table 1 Number of Accidents in the State of

Note:

Source:

- (4) 本文における文献引用は、「...である（阿部1995、p.36）」あるいは「...である（阿部1995）」のようにする。英文では、（Abe 1995, p.36）あるいは（Abe 1995）とする。

- (5) 本文における注記の付け方は、(...である¹。)とする。英文の場合は、(...¹)とする。
 (6) 注記、参考文献は論文末に一括掲載する。

注記

- 1
 2

- (7) 参考文献の書き方については以下のようにする。

日本語単行本：著者（発行年）『書名』、発行所

（例）日本太郎（1999）『これからの評価手法』、日本出版社

日本語雑誌論文：著者（発行年）「題名」、『雑誌名』、巻(号)：頁 - 頁

（例）日本太郎（1999）「評価手法の改善に向けて」、『日本評価研究』、1(2)：3 - 4

日本語単行本中の論文：著者（発行年）「題名」、編者『書名』、発行所、頁 - 頁

（例）日本太郎（2002）「行政評価」、日本花子『評価入門』、日本出版社、16-28

英文単行本：著者（発行年）、書名、発行地：発行所、

（例）Rossi, P. H. (1999). *Evaluation: A Systematic Approach 6th edition*. Beverly Hills, Calif: Sage Publications.

英語雑誌論文：著者（発行年）、題名、雑誌名、巻（号）、頁 - 頁、

（例）Rossi, P. H.(1999). Measuring social judgements. *American Journal of Evaluation*, 15(2), 35-57.

英語単行本中の論文：著者（発行年）、題名、In 編者（Eds.）、書名、発行地：発行所、
 頁 - 頁、

（例）DeMaio, T. J., & Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Francisco, CA: Jossey-Bass, 177-196.

（注1）同一著者名、同一発行年が複数ある場合は、（1999a）（1999b）のようにa,b,cを付加して区別する。

（注2）2行にわたる場合は2行目移以降を全角1文字（英数3文字）おとして記述する。

編集後記

第2巻第1号を新しいデザインで発行することができました。しかし、第2巻第1号では採用原稿の数は伸びず、結局期待とは異なりかなり薄いものとなりました。第2巻第2号は、9月発行を予定しております。また、特集を組んでおりますので、本年度学会誌としては情報量としては何とかバランスはとれそうだと思っております。

また、本学会誌では、査読を行うことによって掲載原稿の水準を保とうとしております。採択されなかった原稿にも学会員で共有したい内容のものもありましたが、論理的展開の弱さや説得性などが不備で採択に至らなかったものもあります。是非、コメントなどを参照し再投稿されることをお勧めいたします。

学会誌を通して、できる限り多くの方の研究成果と評価調査の経験を会員間で共有し、評価の水準を向上させるとともに評価文化の確立を目指したいと思います。会員皆様の積極的な投稿を期待いたします。

編集委員長 三好 皓一

『日本評価研究』第2巻第1号

2002年6月27日

編集・発行

日本評価学会

〒135-0047 東京都江東区富岡2-9-11京福ビル

財団法人国際開発センター 企画広報室内

TEL : 03-3630-6994 FAX : 03-3630-8120

印

刷

株式会社三好デジタルプリンティング

©日本評価学会

The Japanese Journal of Evaluation Studies

Vol. 2, No. 1, June 2002

CONTENTS

Articles

James W. Altschuld

The Preparation of Professional Evaluators:
Past Tense and Future Perfect

Koichi Miyoshi

Utilization of Program Theory Matrix:
For Improving The Usefulness of Evaluation Analysis

Research Notes

Tatsuya Ono

Performance Measurement in Local Government Problems
and Prospects

Mariko Homma

Evaluation of JICA Projects from Socio-Gender
Perspective: Case Studies in African Countries

Nobuhiro Mori Mitsuru Aotsu

An Idea for Strategic Evaluation of Official Development
Assistance

Development

Charles Gene Lyle

Reflections on a Career in Evaluation

Activities

Editorial Board

Training Board

International Exchange Board

Others

3rd Annual Conference

List of Evaluation-related Web Sites