Japanese Journal of Evaluation Studies

Vol. 8, No. 2, March 2008

English Version

CONTENTS

Japanese Journal of Evaluation Studies

Vol. 8, No. 2, March 2008
English Version

CONTENTS

Forewords

Special Issue: Capacity Development and Evaluation

Keiko Nishino
GLM Institute

Koichi Miyoshi
Ritsumeikan Asia Pacific University

Achieving positive societal change depends heavily on the capacity of government and the community in general, particularly in the context of sweeping economic, social, political and environmental transformation caused by both globalization and ecological turbulence. Given the policy trend towards decentralization, capacity development in developing countries has become an important item on the international development agenda and is essential to attain the benefits of globalization without worsening the level of social inequality. On the other hand, the severe budgetary restrictions and aging of Japanese society has weakened and thrown doubt on the administrative and governance capacity the local government. Thus, capacity development is being emphasized on a national scale in Japan.

In this connection, it becomes more urgent to evaluate the performance of government, community and society, and to utilize the evaluation results to upgrade their capacity. Responding to this situation, this special issue compiles four papers which describe capacity development and evaluation and provide food for further thought and discussion.

Matsuoka, Murakami, Aoyama, Takahashi and Tanaka explore social capacity assessment for environmental management. Their paper defines the basic concept of social capacity assessment and provides an analytical framework. Based on this framework, a possible design of an aid program to develop social capacity is presented. Fujikake employs qualitative data to evaluate the level of empowerment and constructs an empowerment evaluation model to visualize the empowerment phenomenon. Miyoshi and Stenning introduce a model of community capacity development and community policy structure, construct a theory-driven participatory evaluation method and emphasize the alternative development approach. Yokoyama examines the effect of the introduction of performance evaluation in local governments in Japan, the changed consciousness of the local government personnel and the long-term benefits brought about by the evaluation system.

Article　Invited Paper

Capacity Development and
Social Capacity Assessment (SCA)

Shunji Matsuoka
Professor
Graduate School of Asia-Pacific Studies
Waseda University
smatsu@ waseda.jp

Kazuma Murakami
Mitsubishi UFJ Research &
Consulting Co. Ltd.

Naoto Aoyama
COE Researcher
Hiroshima University (as of 2006)

Yoshi Takahashi
Associate Professor
Graduate School for International Development and
Cooperation, Hiroshima University

Katsuya Tanaka
Assistant Professor
Research Center for Sustainability and Environment
Shiga University

Abstract

This paper presents the basic design of the Social Capacity Assessment (SCA) studies in the case of environmental management. Firstly, it is defined basic concepts of SCA. In this model, the Social Capacity for Environmental Management (SCEM) is defined as the capacity to manage environmental problems in a social system composed of three actors, i.e., government, firms, and citizens and their interrelationships. The Social Environmental Management System (SEMS) is defined as interactions between the SCEM and institutions. According to these definitions, interactions between the SEMS, socio-economic condition, environmental quality and external factors shape a total system. Secondly, based on these concepts, we build the analytical methods of the SCA. The SCA is made up of the following five steps, (1) Actor-Factor Analysis, (2) Indicator Development, (3) Institutional Analysis, (4) Path Analysis, and (5) Development Stage Analysis. Finally, based on analytical methods provided above, the aid program for social capacity development is designed in order to achieve the aid effectiveness.

Keywords

Capacity Development, Capacity Assessment, Institutional, Analysis, Environmental Management

1. Introduction

During the 1990s, it became apparent that the Replacement Approach, i.e., the one-sided transfer of knowledge and technology from advanced countries to developing countries was insufficient to deal with the issues of international

development assistance. Moreover, a recent study conducted using the Capacity Development Approach (Fukuda-Parr *et al.* 2002) revealed that the self-efforts of the developing countries are necessary to improve their social capacity and enable them to achieve sustainable development performance. Although there has been some progress in the stakeholder and the institutional analyses (see Morgan and Taschereau, 1996; Lopes and Theisohn, 2003), there still exists a need to further intensify the research and development on Capacity Assessment.

In 2003, the Graduate School for International Development and Cooperation (IDEC) at Hiroshima University launched the 21st century Center of Excellence (COE) program, "Social Capacity Development for Environmental Management and International Cooperation" (Principal Researcher from July 2003 to March 2007 was Professor Shunji Matsuoka. In April 2007, he transferred to a professor at Waseda University.). The COE program is a five-year research project granted by the Japanese Government. It proposes a conceptual development model and indicators of Social Capacity for Environmental Management based on an environmental policy research, from technological and socio-economic perspectives. The purpose of this program is to design policy proposals for international cooperation and to achieve aid effectiveness in the field of environmental management. Final objective of this research is to encourage developing countries to evaluate and enhance their own social capacity for environmental management.

In 2004, this COE program in cooperation with several government agencies established the Japan Committee on Social Capacity Development (JCSCD). The objective of the JCSCD is to innovate Capacity Development frame work, based on the experience of East Asian Countries. The committee consists of Hiroshima University, the Japan International Cooperation Agency (JICA), the Japan Bank for International Cooperation (JBIC), the Institute of Developing Economies, the Japan External Trade Organization (IDE-JETRO), and the National Institute for Environmental Studies (NIES).

Since the launch of this COE program, the Social Capacity Assessment (SCA) model has been proposed in order to enable developing countries to achieve sustainable development. The research on the SCA has progressed on the following three consecutive levels: (1) the definition of concepts, (2) the establishment of the formal models, and (3) the development of the indicators (Matsuoka and Kuchiki 2003, Matsuoka, 2004, Matsuoka *et al.* 2004, Matsuoka 2007). In November 2005, during our joint seminar with the representatives from the World Bank in Washington D.C., a productive discussion regarding the design of our SCA was made. Based on the outcomes of our discussions during the joint seminar, we launched a pilot program in 2006. The pilot program applied the SCA methodology to Indonesian water quality environmental management case and Mongolian combating desertification case.

This research paper summarizes the studies carried out under this COE program in order to develop our SCA. The paper is divided into five sections. Section 2 introduces the concept and analytical methods of the SCA. Section 3 provides a detailed description of the following analytical methods: 1 Actor-Factor Analysis; 2. Indicator Development; 3. Institutional Analysis; 4. Path Analysis; and 5. Development Stage Analysis. Section 4 discusses the program design for social capacity development based on the analytical approaches described in section 3. Finally, section 5 presents the summaries and conclusions of our analysis.

2. Social Capacity Assessment (SCA)

Determining the target capacity level and obtaining information about the system factors of capacity development, i.e., socio-economic factors, environmental quality, and external factors, are the initial problems faced during the assessment of social capacity. Since the SCA has to be applied by the developing countries, it should be inexpensive, simple, and based on scientific research. Moreover, the development of the self-assessment ability of a developing

country must also be considered, in order to enable the country to assess its own social capacity.

The Social Capacity for Environmental Management (SCEM) is defined as the capacity to manage environmental problems in a social system composed of three social actors, i.e., government, firms, and citizens and their interrelationships (see Figure 1). The Social Environmental Management System (SEMS) is defined as the system of interaction between the SCEM and institutions (see Figure 2).

Figure 2 also shows the interrelationships between the SEMS, the socio-economic condition, the environmental quality, and the external factors in the total system. The SEMS of a country is constrained by the existing socio-economic conditions and the condition of the environmental quality. Furthermore, here we observe the inter-

Figure 1 Social Capacity for Environmental Management
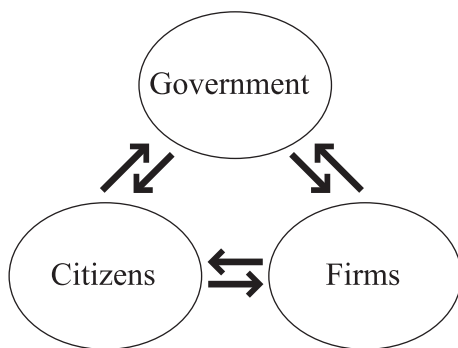
Source: Matsuoka and Kuchiki (2003)

Figure 2 Social Environmental Management in Total System
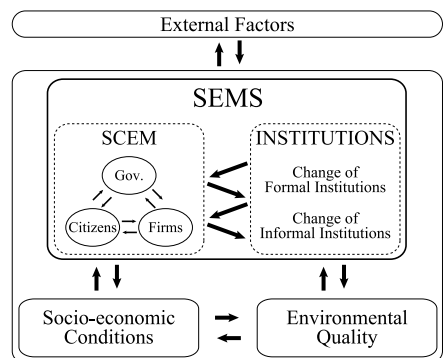
Source: Matsuoka (2005)

Figure 3 SCEM and Social Capacity Assessment

Source: Matsuoka (2005)

Shunji Matsuoka　Kazuma Murakami　Naoto Aoyama
Yoshi Takahashi　Katsuya Tanaka
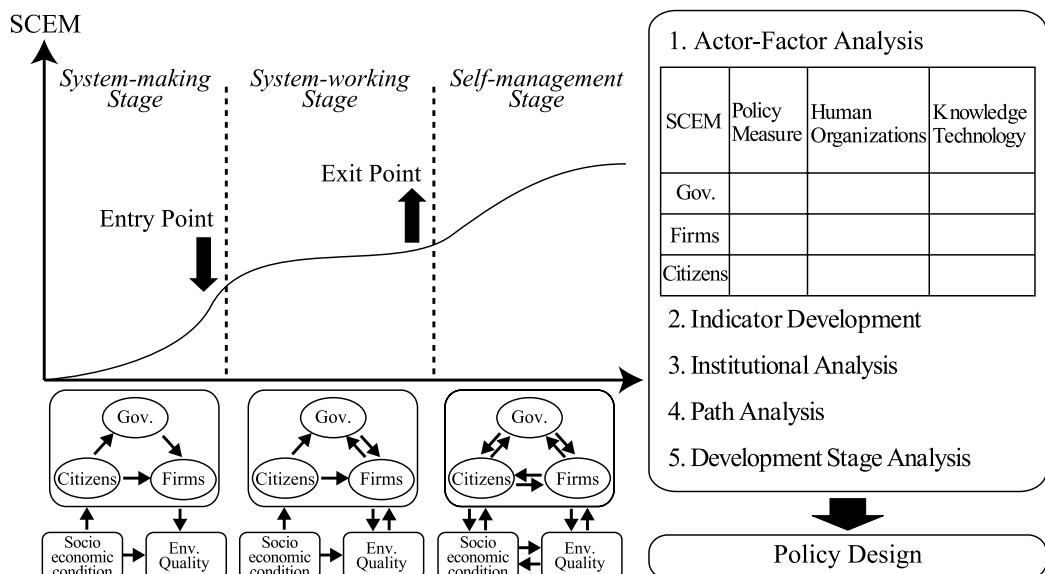
prescribing relations between environmental quality and socio-economic conditions (See, e.g., Matsuoka and Kuchiki 2003 and Matsuoka *et al.* 2004).

As evident in figure 3, the SCA is designed to analyze the interactions between the SEMS, the socio-economic condition, and the environmental quality of a total system. Apart from this, it is also designed to analyze the social capacity of each actor and the interactions between all the social actors. Thus, the SCA reveals the current social capacity and the development path of a particular region and/or a country. The SCA includes the following five steps: 1. Actor-Factor Analysis, 2. Indicator Development, 3. Institutional Analysis, 4. Path Analysis, and 5. Developing Stage Analysis. The next section provides a brief introduction to these steps.

2.1 Actor-Factor Analysis

The actor analysis evaluates the social capacity at a given time period, by analyzing the capacity of social actors (i.e., government, firms, and citizens) and their interrelations. This analysis also provides information regarding "critical minimum or benchmarks", that is, the capacity level each social actor has to satisfy that the social system functions. The factor analysis, on the other hand, focuses on the factors of social capacity, i.e., policies and measures, human and organizational resources, and knowledge and technology. It provides information on the existing condition of each factor and its critical minimum.

2.2 Indicator Development

This develops the indicators that carry summary information regarding the Social Capacity for Environmental Management. Based on the actor-factor analysis, two different statistical approaches are proposed.

2.3 Institutional Analysis

The institutional analysis investigates the institutions that form the basis of social capacity. The analysis deals with formal institutions (e.g., legal system) as well as informal institutions, and their interactions. The results of the institutional analysis reveal information regarding the reformation of current institutions for the development of social capacity.

2.4 Path Analysis

Based on the targets set by prior analyses, the path analysis concentrates on the development path of social capacity in order to achieve the targets. The path analysis also investigates the development path of the social capacity level, the socio-economic background, and the environmental performance. Thus, based on the relationship between the social actors and its development path, the path analysis provides the respective capacities of the social actors.

2.5 Development Stage Analysis

Based on a three-stage development of social capacity, i.e., system-making, system-working, and self-management, the development stage analysis reveals information regarding the current development stage, the next development target, and the approach to achieve the target (along with the path analysis). Thus, the development stage analysis can also be considered as a strategy for providing aid assistance.

Through the five steps analysis mentioned above, the SCA enables us to measure the current capacity level, the development path, the current state of development, and the institutions necessary to improve the capacity for environmental management. The next section presents a detailed explanation of each analysis.

3. Social Capacity Assessment Approach

3.1 Actor-Factor Analysis

The actor-factor analysis reveals the level of social capacity by combining the results of both the actors and factors approaches. This provides us with a concrete estimation of the social capacity. The results obtained by the actor-factor analysis enable us to design suitable programs for international development assistance.

In order to appropriately conduct the actor-factor analysis, we propose an actor-factor matrix (see Table 1) of 3 actors and 3 factors, i.e., a 3× 3 matrix. The data used to construct this matrix is obtained from statistical tables and through the interview and survey of each social actor. The cells of this matrix indicate the level of social capacity attained by each social actor. Table 1 displays the information regarding the programs and projects designed to compensate for the capacity gap, i.e., the difference between the actual social capacity and the critical minimum of social capacity established for each social actor's contribution to the designated factors.

The critical minimum that is obtained for each factor and is assumed to yield good results in terms of the environmental performance is distributed among the actors proportional to the roles they perform in their respective societies. However, this distribution is not always fixed, and changes in the initial situation might induce changes in the distribution of the critical minimum. These changes depend on institutions such as a political system and the relationship between the actors, historical path dependency, and the characteristics of the environmental problem. Moreover, the time required for the transition to the next development stage might also induce changes in the distribution of critical minimum among the social actors. (For details, see section 3.5).

For the purpose of our analysis, we assume the government (G), firms (F) and citizens (C) as the social actors. However, it is also possible to consider a collection of scientists and media as the fourth social actor (Zhang *et al.* 2004). Furthermore, we define the SCEM as the environmental management capacity stipulated by the capacity levels of the social actors and the correlation between them. Table 2 shows the classification of actors that are targeted for assessment. Among previous researches that have contributed to our understanding of the factors of environmental management capacity, the joint work by the UNEP and WHO, which focused on the air quality management capacities in cities, is worth a mention (UNEP/WHO 1996). The above-mentioned study assumes that the capacity for air quality management comprises four elements (see Figure 4). However, the targets in this study were limited to the capacity of the government and the local administration for managing the air quality. Thus, we focus on extending

Table 1 Actor-Factor Analysis: The Actor-Factor Matrix

Factors / Actors	Policy & Measure	Human & Organizations	Knowledge & Technology
Gov.	↕ Existing Capacity	↓ Critical Minimum	Project
Firms	↕ Capacity Gap		Project
Citizens			
G - F			
G - C			
F - C			
G - F - C			

Source: Matsuoka (2007)

Table 2 Classification of Actors in the Actor-Factor Analysis

	Classification
G: Gov.	Central government
	The government offices concerned
	The sections concerned
	The government
	The government offices concerned
	The sections concerned
F: Firms	Industry
	Industry fields (Major groups, Medium groups)
	Firms (Big business, Small and medium-sized businesses)
	Industrial unions
C: Citizens	Civil Society Organization (NGO, NPO, CBO)
	Citizens
G: Gov. ─ F: Firms	
G: Gov. ─ C: Citizens	
F: Firms ─ C: Citizens	

Source: Matsuoka (2007)

Shunji Matsuoka   Kazuma Murakami   Naoto Aoyama
Yoshi Takahashi   Katsuya Tanaka

Figure 4 Air Quality Management Capacities
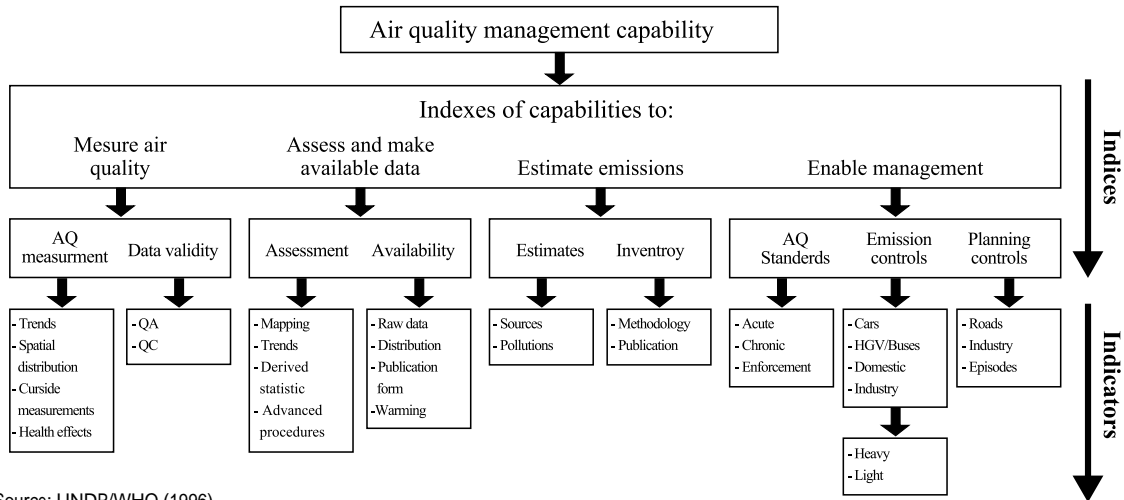
Source: UNDP/WHO (1996)

Table 3 Actor-Factor Analysis: Air Quality Management in China

		P: Policy & Measure	H: Human & Organizations	K: Knowledge & Technology
G		**Critical Minimum** Command and control - environmentals law are developed.	Organization -environmental Administration is developed.	Research, investigation -air pollution monitoring stations are installed. -environmental information is disclosed.
		1979 Environmental protection law (trial version) 1987 Air pollution control law 1989 Environmental protection law 1995 Environmental protection law 1996 The ninth five year plan	1988 National Environmental Protection Administration (NEPA) 1998 State Environmental Protection	1990 The China Environmental Yearbook 1995 Upgrade its quality
		In the mid-1990s, Critical Minimum was achieved (*System-working*)		
F		**Critical Minimum** Command and control -obey the law	Equipment, facilities -install end-of-pipe technology	research, investigation -self-monitoring for emission source
		a questionnaire etc.		
C		**Critical Minimum** Command and control -lodge a complaint, make demands, lobbying	Organization -NGO, NPO activity	Research, investigation -recognize air quality
		a questionnaire etc.		

Source: Matsuoka (2007)

this parameter of analysis by including the capacities of firms and citizens. Table 3 shows an example of the results of an assessment using the actor-factor analysis for air quality management in China. Considering the capacity of the government in China, we find that the critical minimum for the capacity for air quality management had been achieved during the mid 1990s.

3.2 Indicator Development
We develop two SCEM indicators using the following different statistical approaches: (1) Frontier/Tobit approach and (2) Factor Analysis approach. This section describes the methodology and the empirical applications of both these approaches.

3.2.1 Frontier/Tobit Approach

This approach is based on the Total System conceptual framework. In this framework, the SCEM as well as socio-economic conditions are included as a single component influencing the environmental performance (see Figure 2). Our analytical framework is as follows: First, the directional distance function estimates the emission-based environmental efficiency as environmental performance (of air quality). The Tobit model is then applied and the estimated environmental efficiency is used to identify the SCEM variables affecting the efficiency scores. Finally, the SCEM indicator is calculated as the weighted average of the SCEM variables.

We begin our analysis with the measurement of the environmental efficiency. Figure 5 illustrates the relationship between production (y) and the corresponding SO_2 emissions (b). Suppose that the current level of production of a firm i is y, while the observed SO_2 emission level is b. However, if this firm incorporates and operates with the best practice technology, then the SO_2 emission can be reduced to b^* with the output remaining constant. The production frontier line indicates the efficient (i.e., minimum feasible) SO_2 emission at the given output. We define environmental efficiency as the distance between observed and efficient levels of SO_2 (b, b^*); the smaller the distance the greater is the efficiency. In this study, the environmental efficiency is empirically estimated by using the directional distance function (Fare *et al.* 1994).

Once the environmental efficiency is estimated, the next step is to evaluate the role of the SCEM using the Tobit model. In this study, the Tobit model selects one SCEM variable for each of the three actors, The identified variables are used to construct the indicator for the SCEM. This is defined as follows:

$$S_{it} = \left(\omega_g \widetilde{G}_{it} + \omega_f \widetilde{F}_{it} + \omega_c \widetilde{C}_{it} \right) \tag{1}$$

where S_{it} is the level of SCEM for province i in year t. G_{it}, F_{it}, C_{it} represent the environmental management capacities of the government, the firms, and the citizens, respectively. ω_g, ω_f, and ω_c represent their weights. These are adjusted such that $\omega_g + \omega_f + \omega_c = 1$. Thus, our indicator proves to be a convenient measure because it always ranges between 0 and 1.

An empirical application of this framework is conducted by using the province-level data of China's manufacturing industry from the period 1994-2002. Using the Tobit model, we identify the total number of monitoring stations as the government's capacity and the ratio of SO_2 reduction as the firms' capacity. However, due to limited data, we are unable to include the citizens' capacity as a part of our model. Thus, the SCEM in this application refers only to the capacities of the government and the firms.

Figure 5 Production Frontier and Environmental Efficiency
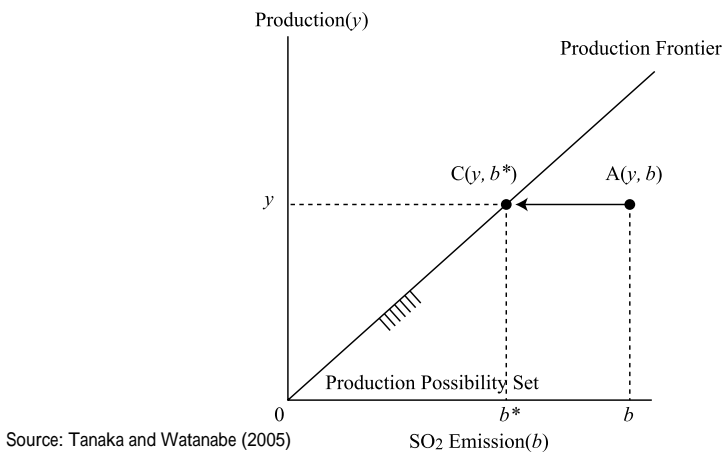

Source: Tanaka and Watanabe (2005)

Figure 6 SCEM Indicator and SO₂ Emission in China's Manufacturing Sector
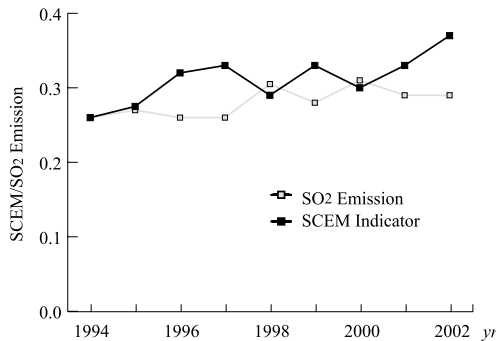
Source: Tanaka and Watanabe (2005)

Figure 7 Actor-specific Environmental Management Capacity
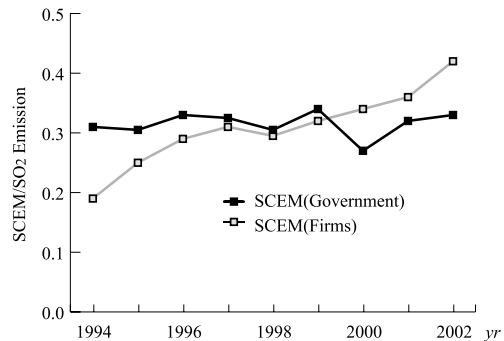
Source: Tanaka and Watanabe (2005)

Figure 6 depicts the SCEM and the normalized SO_2 emissions in China's manufacturing sector for the period 1994-2002. The figure indicates a significant increase of nearly 40% in the SCEM - from 0.25 in 1994 to 0.35 in 2002 - during the estimation period. In addition, the SO_2 emission is shown to be fairly responsive to the SCEM. Figure 7 illustrates the environmental management capacities for the government and the firms during the same estimation period. The firms' capacity (SO_2 reduction rate) increased from 0.19 in 1994 to 0.42 in 2002 - an increase of more than 120%. On the other hand the government's capacity (total number of monitoring stations) development rate improved by a mere 8%, i.e., from 0.31 in 1994 to 0.34 in 2002. Thus, the SCEM development in this period is largely due to an improvement in the firms' capacity, while the contribution by the government is rather limited.

In this section, we developed the indicator for the SCEM using the Frontier/Tobit approach. We observed a rapid increase in the SCEM in China for the period 1994-2002. Moreover, the results indicated a significant contribution of the firms in the development of the SCEM, while suggesting a limited contribution of the government. However, in order to provide future suggestions and recommendations, a further interpretation of these results is required. Finally, this approach can be extended to conduct an international comparison using international panel data. In future studies, we will use the same approach to analyze the SCEM development in Asian countries.

Table 4 Factor Loading and Contribution of Factor Loading (1) (Kitakyushu-city)

Data	factor 1	factor 2	factor 3	factor 4	elements of capacity
budget for the Environmental Research Center (ERC)	0.933	-0.182	0.000	0.058	
budget for Environmental protection (City)	0.819	-0.080	0.380	0.342	"policy resource"
number of personnel * average employment period (ERC)	0.733	0.310	0.411	0.347	management
number of monitoring stations	0.692	0.172	0.502	0.408	
number of personnel * average employment period (City)	0.096	0.915	-0.076	0.216	
number of investigations into emission source (City)	-0.229	0.855	0.167	0.024	"command and control" policy enforcement
number of inspections of a sample from emission source (ERC)	0.133	0.707	-0.033	0.450	
amount of finances provided by gov. to the firms for air pollution control (City)	-0.198	0.073	-0.818	-0.100	"financial support" policy
number of finances provided by gov. to the firms for air pollution control (City)	-0.571	-0.372	-0.603	-0.286	enforcement
number of presentations in academic society (ERC)	0.394	0.253	0.170	0.864	provision of "scientific
number of articles published in academic journal (ERC)	0.193	0.420	0.271	0.526	knowledge"
eigenvalue	3.363	2.508	1.821	1.594	
contribution(%)	52.0	21.4	6.9	4.2	
cumulative contribution(%)	52.0	73.4	80.2	84.4	

Source: Murakami and Matsuoka (2005)

Table 5 Factor Loading and Contribution of Factor Loading (2) (Osaka-city)

Data	factor 1	factor 2	factor 3	factor 4	elements of capacity
number of monitoring stations	0.971	0.189	-0.034	0.002	
number of personnel * average employment period (City)	0.832	0.443	0.220	0.216	"policy resource" management
number of personnel * average employment period (ERC)	0.687	0.629	0.225	0.213	
budget for the Environmental Research Center (ERC)	0.665	0.604	0.257	0.317	
budget for Environmental protection (City)	0.613	0.542	0.381	0.409	
amount of finances provided by the gov. to the firms for air pollution control (City)	-0.225	-0.952	-0.046	-0.049	"financial support" policy enforcement
number of finances provided by the gov. to the firms for air pollution control (City)	-0.492	-0.827	-0.052	-0.108	
number of articles published in academic journal (ERC)	0.068	-0.029	0.992	-0.091	provision of "scientific knowledge"
number of presentations in academic society (ERC)	0.389	0.496	0.580	0.230	
number of investigations into emission source (City)	0.020	-0.383	-0.503	-0.568	"command and control" policy enforcement
number of inspections of a sample from emission source (ERC)	-0.212	-0.103	0.043	-0.489	
eigenvalue	3.538	3.479	1.995	0.891	
contribution(%)	42.6	30.3	10.7	4.9	
cumulative contribution(%)	42.6	72.9	83.5	88.4	

Source: Murakami and Matsuoka (2005)

3.2.2 Factor Analysis Approach

Factor analysis is a statistical analysis technique that is used to uncover the latent relationships between many observed variables. This approach allows numerous correlated variables of air quality management policy to be summarized by fewer dimensions, i.e., factors. In the context of this research, the factors are interpreted as the elements of capacity for air quality management that contribute to the environmental performance. Murakami and Matsuoka (2005) estimate the factors of government capacity for air quality management in Kitakyushu and Osaka cities by using the factor analysis. In this study, the capacity for air quality management is assumed to be equal to the factor scores and to the contribution of factor loadings that are estimated by using the data on air quality management policies in Kitakyushu and Osaka cities from 1970 to 2000. Tables 4 and 5 show the results of factor analysis for each city. The screen test for factor analysis reveals four elements of capacity in each city. The four elements are further arranged into three factors, i.e., Policy & Measure, Human & Organization, and Knowledge & Technology (see Table 6).

By using the factor scores and the contribution of factor loadings, we estimate the weighted average for all the four elements. This is assumed to be an indicator of the capacity for air quality management in each city. The contribution of factor loadings is assumed to be the weights for capacity elements. The average weights of the factors of capacity of the two cities are as follows: Knowledge & Technology is 7.5%, Human & Organization is 47.3%, and Policy & Measure is 31.8%.
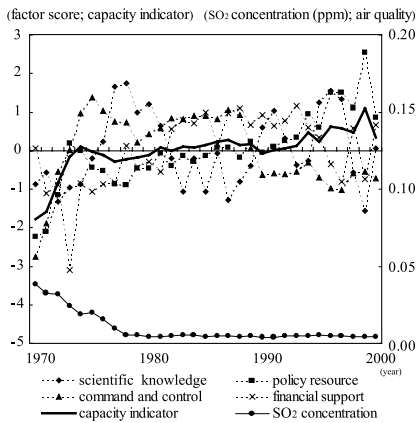
Table 6 Correlation of the Three Actors and Critical Minimum

Source: Matsuoka (2007)
Note: The State of Correlation of the Three Actors has an effect on
        the Critical Minimum Level.

Figure 8 Trend of Government Capacity for Air Quality
Management (Kitakyushu- city in Japan)

(factor score; capacity indicator)   (SO₂ concentration (ppm); air quality)

Figure 9 Trend of Government Capacityfor Air Quality
Management (Osaka-city in Japan)

(factor score; capacity indicator)   (SO₂ concentration (ppm); air quality)
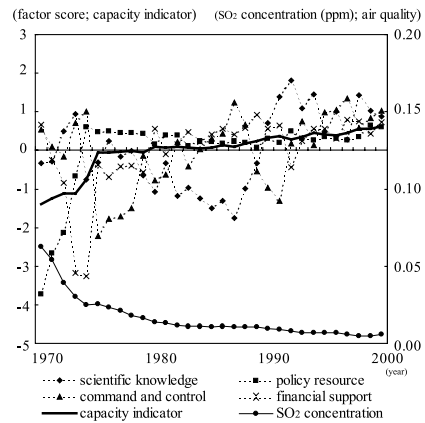
Source: Murakami and Matsuoka (2005)   Source: Murakami and Matsuoka (2005)

Figures 8 and 9 show the change in the government's capacity for air quality management from 1970 to 2000. It can be observed that the rapid improvement in government capacity in the early 1970s resulted in a dramatic reduction in the SO_2 concentration. Additionally, the effects of each indicator of capacity on the SO_2 concentration are estimated by a simple regression analysis.

3.3 Institutional Analysis

The institutional analysis of the SCA investigates a group of institutions (see, e.g., Aoki and Okuno, 1996) that constrain social actors' activities and capacities. It also regulates the current capacity level and affects the future formulations of social capacity. Therefore, this study will focus on the role of the individual institutions and the group of institutions as well as the processes of transitions among them. For this purpose, we will classify the institutions into two categories: principal institutions and secondary institutions, and then, we will classify each category into two subcategories, i.e., formal and informal institutions.

The method of classifying an institution as a principal or a secondary institution is based on analyzing them according to the level of their incentive or disincentive, i.e., the upper levels are principal institutions, and the lower levels are secondary institutions. Further, in order to classify the institutions into the subcategories, i.e., formal and informal institutions, this study follows North's study (1990) and defines formal institutions as public formalized rules, such as state laws, and informal institutions as unspoken rules, such as social norms and customs that influence the behavior of social actors.

While investigating informal institutions, we pay close attention to the changes in the relationships between the social actors. Figure 10 indicates the basic concepts for analyzing the informal institutions. Based on these concepts, we identify three types of relationships between the social actors: one-side (or direct) relationships, mutual relationships, and multilateral relationships (partnership). As shown in table 6, each relationship has an effect on the critical minimum capacity of each actor. Thus, the next step is to analyze the impact of each relationship between the actors on their critical minimum capacities.

In order to conduct this analysis, we introduce a case study wherein we have analyzed the institutional changes in Ube City. Ube City, often referred to as the " Ube Model" or the" Ube System" (Nose, 1996), is a model Japanese city that has succeeded in effectively managing the problem of air pollution. The most important characteristic of the

Figure 10　The Benchmarks for the Social Actor's Relationship
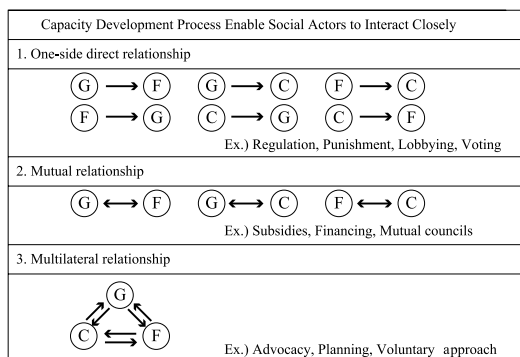
Source: Matsuoka (2007)

Table 7　The Benchmarks for The Social Actor's Relationships

	Principal Institution	Secondary Institution
Formal Institution	Ube Model	
Informal Institution	Institution based on the specific culture and customs of Ube City	

Source: Matsuoka (2007)

Table 8　Environmental Policy and the Characteristics in Ube City

	Dust pollution	SO$_2$ pollution
Main Events	1949 (Dust control measure was initiated) Ube City Dust Fall Control Committee 1952 - 54 The citizens held large-scale Anti-dust pep rallies 1956 Ube Pozzoran Cement 1957 The mayor and important business owners set numerical targets for dust control measures and each factory decided to make a plan, including time limits and expenditures, in order to accomplish the goals that were laid down.	1960 Ube City Air Pollution Control 1962 SO2 monitoring devices were set up in 19 area 1968 Enactment of the Air Pollution Control Law 1969 The first official warning was announced in Ube City 1970 The first air pollution alarm in Yamaguchi prefecture was officially announced in Ube City 1971 Ube city concluded the pollution control agreement 1972 The full-scale work on SO2 measures began after finalizing the enforcement details of the pollution control agreement
Axis Institution	Ube Model	Pollution Control
Characteristics	The dust control measures were adopted promptly and social capacity was formed.	The institutions of the Ube model did not function efficiently for the SO2 control measures. Social capacity did not improve and sufficient pollution control measures were not adopted. Eventually, institutional change in Ube City accelerated under the external pressure of the increased restrictions that were instituted at the national level. The improvement of social capacity was achieved through the institutional change that was instituted after the finalization and implementation of the pollution control agreement in 1971. This resulted in an improvement in the efficiency of the SO2 control measures.

Source. Matsuoka et al. (2004)

Ube model is that the decision-making process is not solely dependent on government regulations; rather, it is a joint exercise carried out by a committee comprising representatives from industry, government, educational system, and general population. It is therefore believed that the spirit of the Ube model can be replicated by formally institutionalizing the informal institutions, however, keeping in mind, the specific culture and customs of a city (see Table 7).

Table 8 indicates the environmental policies and their characteristics in Ube City. Figure 11 shows the relationship between institutional changes (formal and informal) and the SCEM of Ube City, while figure 12 shows the systemic change and the formulation of SCEM in Ube City. Thus, we observe that as compared with the policy

Figure11 Institutional Change and Social Capacity Environmental Management
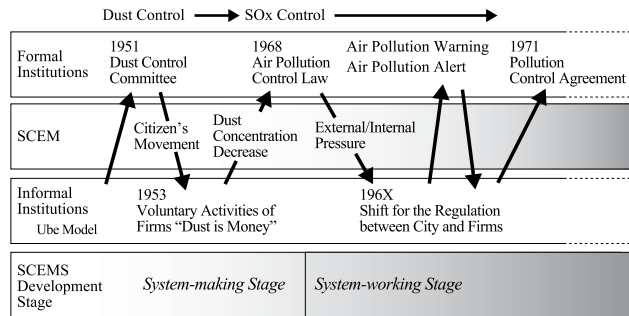
Source: Matsuoka *et al.* (2004)

Figure 12 Social Capacity for Environmental Management in Ube City
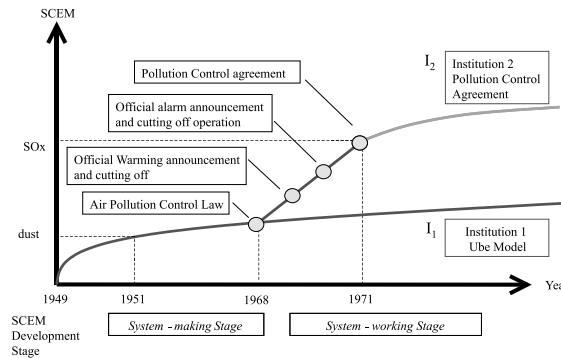
Source: Matsuoka *et al.* (2004)

for dust pollution, the measures for controlling SO_2 in Ube City were delayed until the enactment of the pollution control agreement in 1970. According to this investigation, we conclude that (1) the knowledge and technology were not sufficient to control SO_2 pollution in Ube City, and (2) the characteristics of the Ube Model. Thus, these conclusions highlight the following:

    (1) The institutions needed for controlling pollution differ on a case by case basis and depend on the type of pollution;

    (2) The efficiency of the performance of the institutions is closely related to the SCEM in the region.

Therefore, in order to achieve a higher capacity level for a country, it is important to analyze the nature of the existing institutions, i.e., whether they are principal/secondary and formal/informal. Moreover, it is also important to ascertain whether the actors' capacities of environmental management satisfy the efficient performance requirement of the institution.

3.4 Path Analysis

The path analysis clarifies the information and the conditions that are prerequisites for setting a rational capacity level target. Moreover, an analysis of the path (strategy or program) adopted for the current social capacity level helps in identifying the ideal path toward achieving the set target.

Figure 13   Transition of SCEM, Socio-Economic Conditions, and Environmental Performance:
The Case of SO₂ in Japan
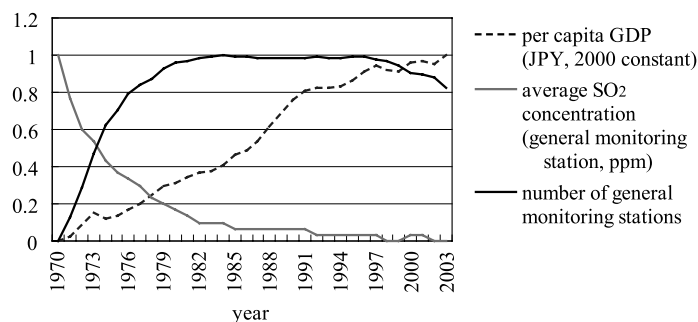
Source: Ministry of Environment, Japan (2005)

Figure 14   Path Analysis
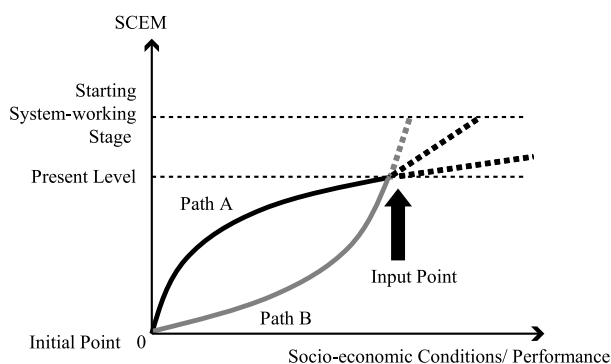
Source: Matsuoka (2007)

As discussed in the previous section, social capacity is developed through the interactions between the actors and the institutions. In a broader sense, we can consider the capacity level as defined by the interrelationship among the capacity level, the socio-economic levels and the performance levels (environmental quality). First, the path analysis deals with the development process of the total system, which consists of three components.

Figure 13 informs us about the indicators pertain to SO₂ in Japan. We adopted the SO₂ general monitoring stations as the capacity level, per capita GDP as the socio-economic level, and the performance level as the average monitoring data at the stations; although, due to limited information, this data was compiled after the peak of the observed SO₂ value. According to the figure, we observe that until the mid 1980s all the three components improved (capacity and socio-economic level increased, while the performance level decreased). However, post the 1980s the socio-economic level continued to improve, while the capacity level remained almost constant and the performance level stabilized at a low level. Based on this information, it can be said that until the mid 1980s the system operated efficiently resulting in an improvement in the environmental performance. However, since then the system continues to operate at a necessary minimum capacity, irrespective of any improvement in the socio-economic level.

By conducting a thorough analysis of the cases of different countries and their environmental issues, we can identify the characteristics responsible for the improvement of the environmental performance in each case. For example, figure 14 clearly demonstrates the differences between the cases wherein the adopted path changes from

Figure 15 Interrelationship between SCEM, Socio-Economic Conditions, and Environmental Performance
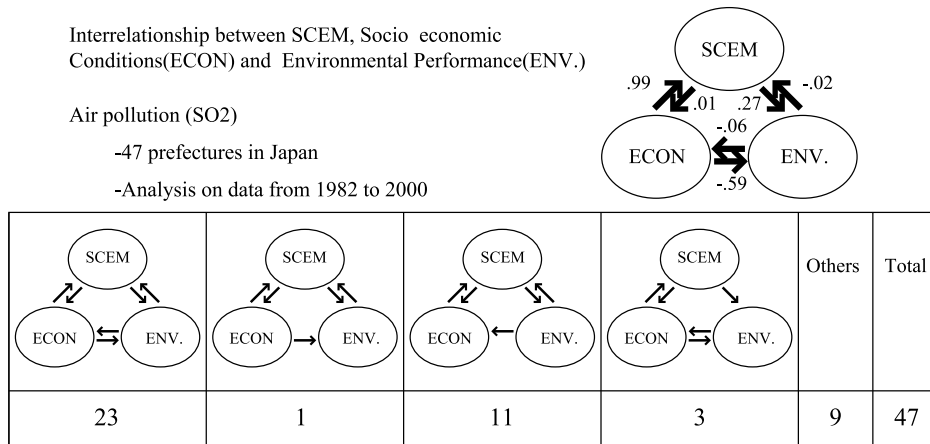
Source: Honda *et al.* (2004)

SCEM-led to socio-economic conditions-led and vice versa. Moreover, such a path analysis enables us to identify the course that we must adopt for improving environmental performance in the future.

Thus far, we have focused on the change in the level of the three components of the total system. However, in order to understand the development process of the system, it is necessary to bear in mind that these changes do not occur independently; rather, they undergo a transition in the context of the interrelationship between the three components. Honda *et al.* (2004) analyzed the relationship between these three components for 47 prefectures in Japan. From among these analyses related to several environmental issues, let us present the case of SO_2. The analysis is carried out using the Granger Causality Test and is based on the data for the period ranging from 1982 to 2000. Figure 15 confirms the existence of interrelationships between the three components for 23 out of 47 prefectures. In order to complete the path analysis, we need to verify the hypothesis that the change would occur from a state of weak or partial interrelationship at an early stage to that of a strong interrelationship with an interactive impact on all the three components (we do not exclude the possibility of plural paths to achieve the target). Thus, we shall now investigate the methodology and pursue these analyses.

In addition, the development processes of the capacities of social actors and their relationships also form a part of the path analysis' targets. In this case, we assume a certain level of substitutability among the actors; for instance, part of the government's role can be borne by a firm or a citizen. Future efficient capacity development paths are different for cases with different paths, such as government-led and citizen-led; however, they have the same level of social capacity as a whole. Regarding aid policy, this proposition implies that there should be cases wherein firms or citizens would not rely on the government to government approach and would be the direct beneficiaries of the aid.

3.5 Development Stage Analysis

The development stage analysis that is conducted on the basis of the actor/factor analysis, the indicator development, the institutional analysis, and the path analysis, aims at specifying the development stage based on the benchmarks and then presenting the development process and the direction for further development. The analytical results highlight certain preconditions that clarify appropriate quantity, quality, and timing of input in order to enable development and aid policies to be implemented as programs.

Matsuoka and Kuchiki (2003), bearing in mind industrial pollution, assumed the following three development

Table 9 The Stages and Benchmarks of Social Environmental Management System

	System-making stage	*System-working stage*	*Self-management stage*
Definition	Period in which the bases of SEMS, especially governmental institutions, are developed.	Period in which the regulations between the government and firm sectors become stronger through the setting the incentives for pollution abatement and industrial pollution improves after reaching its peak.	Period in which a comprehensive environmental policy is needed, since new types of environmental issues emerge, and the firms and citizens sectors take leading roles through voluntary approaches for environmental management. Harmonious relations between government, firms, and citizens accelerate the efficient social environmental management.
Environmental Issues	Poverty related issues and issues related to industrial pollution.	Issues related to Industrial pollution.	Consumption-related issues.
Issues related to Industrial Pollution	Degradation.	Turning point (peak of the Environmental Kuznets Curve).	Improvement.
The Role of the Three Actors	-Government (system-making) -Firms (efforts for pollution reduction) -Citizens (pressure on the government and firms and research cooperation)	-Government (pollution control regulation) -Firms (pollution reduction) -Citizens (pressure on the government and firms and research cooperation)	-Government (proposal of comprehensive policy) -Firms (voluntary approach) -Citizens (voluntary approach)
The Relationship between the Three Actors	Government - Firms Government - Citizens	Government - Firms Government - Citizens Firms - Citizens (through government)	Firms - Citizens Government - Firms Government - Citizens
Benchmarks (Essential)	-Environmental Law -Environmental Administration -Environmental Information (Monitoring Data)	-Regulation -Reaching the peak of pollution level and improvement	<First phase> (In the case of developing countries) -Graduation / Independence from ODA <Second phase> -Comprehensive Environmental Management
Benchmarks (Important)	- Negotiations between Government-Firms, Government-Citizen - Mass media	- Negotiation, adjustment, and cooperation between Firms and Citizens	Voluntary approach of Firms and Citizens (Environmental Accounting, Environmental Reporting, Green Consumption, and Advocacy Planning)

Source: Matsuoka and Kuchiki (2003)

stages for the SEMS: system-making stage, system-working stage, and self-management stage. Table 9 indicates the stages and the benchmarks of SEMS.

The system-making stage focuses on the development of the fundamental functions of the SEMS. Since this stage particularly focuses on the capacity development in the government sector, the benchmarks in this stage should be the development of the environmental law (basic law and acts for specific pollution control mechanisms), environmental administration, and environmental information. With regard to the environmental information benchmark, it is important to arrange the data by networking, understanding the environmental status, and then presenting the policy measures. Thus, we use not only the number of monitoring stations but also the first publication of the State of the Environment and the like as specific evaluation indicators.

In the system-working stage, the system actually starts functioning to improve the environmental quality. This occurs in response to the improvement of the basic environmental administrative institutions. As the pollution trend changes from increasing to decreasing a turning point of the so-called environmental Kuznets curve is observed.

With reference to this, we focus upon the results of the implementation of government regulation (reduction of pollution by firms) and the consequent change to a decrease in pollution levels. In order to evaluate the achievement of pollution reduction measures, the standard achievement ratio of SO_x　a typical industrial pollutant　will be observed as the indicator. If the achievement ratio for all the monitoring stations in the country is higher than 90%, then it is considered to be an indication of the end of SO_x pollution. In developed countries, the Command and Control (CAC) has played a significant role in pollution reduction at the system-working stage. The CAC requires the government to utilize its administrative capacity in order to understand the state of pollution, set regulation standards, and ensure that those responsible for pollution are complying with the regulations. It is observed that as compared to the governments of developed countries, the governments of developing countries lack this administrative capacity and are therefore ineffective in implementing the CAC. However, pollution reduction can be realized efficiently by effectively introducing the market based instruments (MBIs) for environmental regulation and utilizing the market mechanism (Matsuoka, 2000).

　　　The self-management stage is the stage wherein the system develops in a sustainable manner through the strong interrelations between the government, firms, and citizens, and a comprehensive environmental policy is enforced. At this stage, firms and citizens voluntarily adopt and participate in initiatives for environmental management. For instance, firms voluntarily upgrade their facilities in order to obtain the ISO 14000 certification as an in-house environmental management program, and in order to increase the efficiency of environmental management, they adopt environmental accounting. Moreover, they highlight their environmental management achievements in order to court consumer appreciation and thus gain a competitive advantage in the market. With regard to international cooperation, at this stage, a developing country becomes less dependent on donor's assistance and utilizes its own financial and human resources.

　　　As a country experiences the development of SEMS, the roles and relationships of the three actors also evolve. The government sector plays an important role in managing and coordinating issues at the system-making and system-working stages; however, at the self-management stage, its responsibility evolves to supporting the firms and the citizens by designing a framework for comprehensive environmental management.

　　　Figure 16 shows the development of SCEM with the stages and benchmarks mentioned above (China's case).

Figure 16　Development stages of SCEM: The case of China

Source: Japan Society for International Development (2004)
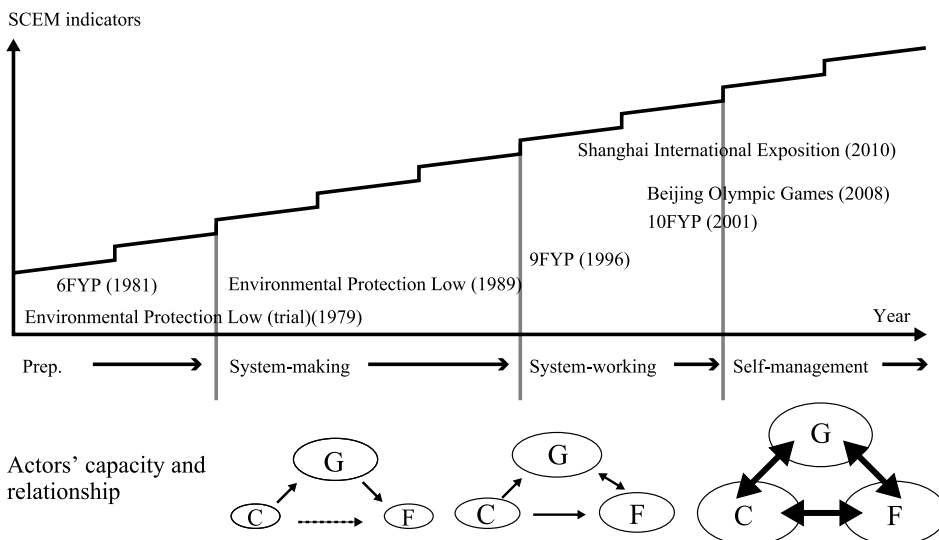
Figure 17 Social Capacity Development in Trade: Malaysian Case
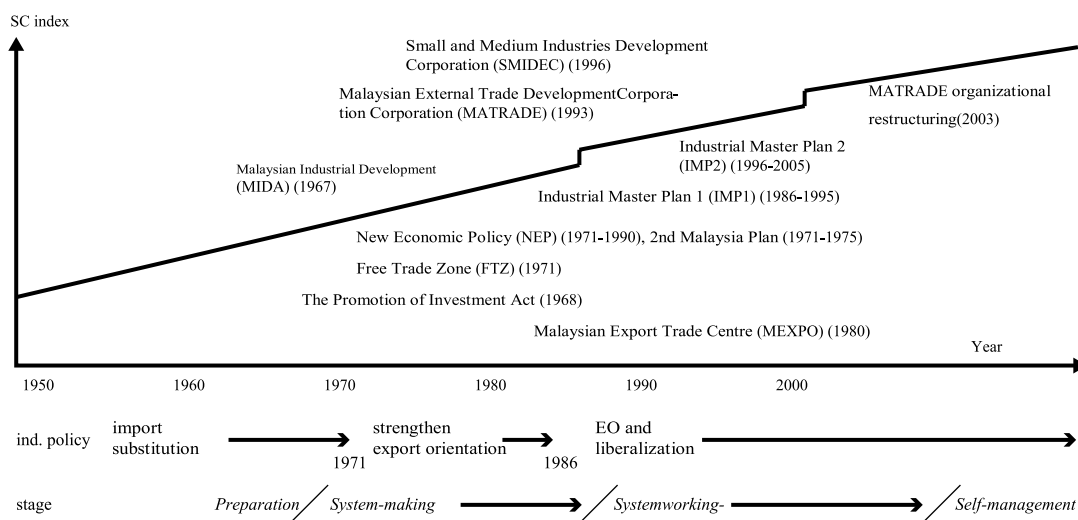
Considering economic indicators or passage of time as the horizontal principal and SCEM index (a group of indicators) as the vertical principal, it can be presumed that, by and large, China adopted the capacity development process that is shown in the figure. After the enactment of the Environmental Protection Law as the starting point of system-making, China entered a full-scale system-working stage during the 9th Five Year Plan (1996-2000). The 10th Five Year Plan (2001-2005) further accelerated this process. It is expected that China will be able to lay the foundation for initiating the self-management stage between the period of the Beijing Olympic Games in 2008 and the Shanghai Expo in 2010.

In terms of relationship between the three actors, the SEMS in China has changed drastically. As shown in figure 16, the government had exclusively performed all the functions and roles at the system-making stage. However, during the system-working stage, although the government continued to institute vigorous steps, the firms did render some important tangible contributions to curtail pollution. In addition, the relationships between the actors, particularly between the government and the firms grew stronger. Based on this, we can expect that during the self-management stage, a more balanced relationship, entailing the promotion of environmental industry and self-sustained growth of an environment-oriented market will be developed.

Moreover, we have also begun to apply the development stage analysis beyond the field of environmental management. Figure 17 describes the development stage analysis of social capacity development for trade (particularly export promotion) in Malaysia. The research is conducted for the JICA evaluation project (Thematic Evaluation: Economic Partnership). We observe that it is possible to conduct the analysis based on a similar format of benchmarks and stage setting; nevertheless, the trade capacity has its peculiar characteristics, such as the limited role of citizens and the vulnerability of performance level to external conditions.

This section introduced and discussed the basic designs of specific analytical methods that form the components of the SCA. The methodology enabled developing countries themselves to understand the current state of pollution and the problem of social capacity. Adopting the analytical method mentioned here as a precondition, the final section deals with the following question: How to transform development and aid policies into effective

programs for attaining the capacity level that developing countries regard as their target.

4. Designing the Program for Social Capacity Development

This section describes the program design for social capacity development. Based on the SCA framework, we develop the program approach to identify the target level of capacity, and to provide specific strategies to achieve the target. The program presents an overall package consisting of: (1) the relationship between social actors, (2) the input resources their quantity and timing, and (3) the institutional changes.

The program approach differs from the conventional stand-alone projects in many respects. This approach considers the following: (a) wide and systematic approach; (b) recognition of mutual dependence of society, economy, and culture; (c) long-term project implementation; (d) the harmonization of system development and its process; (e) focus on the capacity of the recipient countries; and (f) cost reduction by avoiding redundant aid projects (Bolger, 2000). Table 10 shows a detailed comparison between stand-alone projects and the program for social capacity development.

Sector-wide approaches (SWAPs) for social capacity development can be classified as one of the approaches of the program. The SWAPs are primarily carried out in basic education and healthcare sectors in the African countries. Jones and Lawson (2000) characterizes the SWAPs as follows: (i) the harmonization of policies between the donor and recipient countries (policy alignment), (ii) efficiency improvement in internal and external resource allocations, (iii) developing partnerships with local stakeholders, and (iv) emphasis on ownership. This characterization, however, is insufficient. We define the program as a program involving three actors (government, firms, and citizens) and three factors (policy and measure, human and organizations, and knowledge and technology). Thus, our social capacity development approach always takes the form of the program.

The program design begins with social capacity assessment based on the actor-factor matrix presented in figure 18. When analyzing the pollution problem, the matrix is used to evaluate: (1) the current capacity for pollution abatement, (2) the critical minimum capacity during the system-working stage, and (3) the gap between current and critical minimum capacities.

It should be noted that we assume that capacities are substitutable between the actors, but not between the factors, i.e., the capacities are complements between the factors. For instance, suppose that the critical minimums for

Table 10 Programs and Stand-alone Projects

Principles	Stand-Alone Projects	Programs
Local Ownership	Projects are often supply-led.	Based on locally owned programs, involving a community of stakeholders.
Donor Coordination	Limited donor collaboration, leading to inefficiency.	A high level of donor coordination, ideally involving all of the donor community, under national leadership
Partnerships	Projects are often managed directly by executing agencies or project implementation units.	Programs are intended to involve movement towards the use of local procedures and controls.
Attention to institutional development, governance issues, and civil society participation	Projects attempt to ensure success by establishing project-specific control mechanisms. They thus attempt to bypass, rather than solve, certain institutional weakness.	Attention is brought to bear on institutional, governance, and participation issues necessary to ensure success and the accountability of local institutions to their constituents.
Results-based Approach	Attention is focused on the success of the projects themselves, even though other conditions necessary to the achievement of development results may not be met.	The focus is on results at the program level such as those identified in the Millennium Development Goals or in the PRSs.

Source: Lavergne and Alba (2003)

Figure 18   SCA and Program Design

1. Assess the Current Capacity, and Required Capacity to Work Program

Factors \ Actors	Policy & Measure	Human & Organizations	Knowledge & Technology
Gov.	↕ Existing Capacity	↓ Critical Minimum	
Firms	↕ Capacity Gap		
Citizens			Project
G - F			
G - C			
F - C			
G - F - C			

2. Design Projects which cover the Capacity Gap in each factor level.
   Capacity can be a complement or substitution among actors.

Capacity Gap   Project

P&M   H&O   K&T
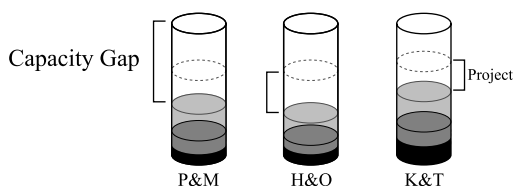
Source: Matsuoka 2007

policy and measure, human and organizations, and knowledge and technology are 30, 50, and 10, respectively. Then, the critical minimum of policy and measure (30) can either be accepted solely by the government or it can also be accepted by the government and the firms jointly. Any combination of actors is possible in order to achieve the critical minimum; however, this is not true in the case of factors. Thus, the "Substitutability of actors" and "complementarities of factors" are equally important in our framework. The capacity gaps identified through the actor-factor matrix are expected to be filled by the projects. These projects are the ones based on the program (referred to as program-based projects) and are different from the conventional stand-alone projects.

Entry and exit points of the program and the projects can be determined through the development stage analysis of social capacity. Figure 19 illustrates the brown issue example. The figure shows the following institutional

Figure 19   Entry, Exit Point, and the Development Stages

SCEM

Exit Point

Entry Point

Comprehensive
environmental
management

Compliance with Regulation

Reach the peak of pollution
level and improve

Environmental Information

Environmental Administration

Environmental law

| Stage | System-making → | System-working → | Self-management → |
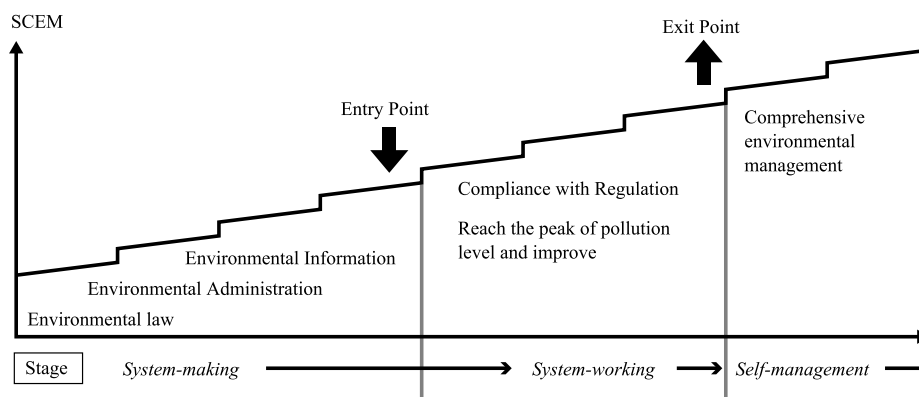
Source: Matsuoka (2007)

milestones during the system-making stage: (1) enacting environmental law, (2) the establishment of environmental administration, and (3) environmental information disclosure. Technical aids, such as the environmental center, are commonly provided by the JICA and can be effective in the latter half of the system-making stage (i.e., developing the system of environmental information disclosure).

In the system-working stage, it is important to focus on environmental business planning, resource allocation and organizational development, and research and development pertaining to pollution reduction. In addition to these, the pollution control management certification system, compliance with regulations, and financial assistance for developing environmental technologies are also important. Aid programs/projects can generally reach their exit point when the level of pollution decreases as per the target. In this stage, the environmental cooperation is horizontal, such as technology exchange, research exchange, and civil exchange. At the same time, the environmental policy measures take the form of economic instruments and self-regulation. Once this is achieved, the recipient countries will gradually move toward the self-management stage.

5. Conclusions

Based on the basic design of the SCA studies conducted under the current 21st COE program, this paper provides the definition of the SCEM as a total system and specific analytical methods, which are the components of the SCA, and the program design. Further, the case studies and the details of the following analytical methods are also provided: (1) Actor-Factor Analysis, (2) Indicator Development, (3) Institutional Analysis, (4) Path Analysis, and (5) Development Stage Analysis.

Hereafter, in order to design programs for achieving sustainable development in East Asia, we intend to continue studies on the SCA in East Asia and also intend to develop the models of the SCA.

Acknowledgement

References

Aoki, M., and Okuno, M. eds. (1996). *Comparative Institutional Analysis of Economic Systems.* Tokyo University Press (in Japanese).

Bolger, J. (2000). Capacity Development: Why, What and How. *Capacity Development Occasional Series*, No. 1, Canadian International Development Agency.

Evaluation Team on Environmental Cooperation, Japan Society for International Development (JASID) (2004). *Thematic Evaluation (Environment) by External Party, Environmental Center Approach: Development of Social Capacity for Environmental Management in developing Countries and Japan's Environmental Cooperation*, Japan International Cooperation Agency (in Japanese).

Fare, R., Grosskopf, S., Norris, M., and Zhang, Z. (1994). Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. *American Economic Review*, 84 (1) :66-83.

Fukuda-Parr, S. ed. (2002). *Capacity for Development: New Solutions to Old Problems*. Earthscan.

Hiroshima University-Mitsubishi Research Institute Inc. Joint Venture (2005). JICA Asia Region Thematic Evaluation "Economic Partnership". Briefing Material for 4th Review Meeting, (in Japanese).

Honda, N., Matsuoka, S., and Tanaka, K. (2004). Causal and Structural Analysis on Social Capacity Development for Environmental Management in Japan's Air Pollution Problems. *HICEC Discussion Paper Series* Vol. 2004-7 (in Japanese).

Jones, S. and Lawson, A. (2000). Moving from Projects to Programmatic Aid. OED *Working Paper Series*, 5.

Lavergne, R., and Alba, A. (2003). *CIDA Primer on Program-Based Approaches*. Canadian International Development Agency (CIDA).

Lopes, C., and Theisohn, T. (2003). *Ownership, Leadership and Transformation: Can We Do Better for Capacity Development*. Earthscan.

Matsuoka, S. (2000). Environmental policy in Developing Countries: Instruments and Efficiency of Regulations. *Journal of International Development Studies*, 9.2: 17-37 (in Japanese).

Matsuoka, S., and Kuchiki, A. (2003). Social Capacity Development for Environmental Management in Asia: Japan's Environmental Cooperation after the Johannesburg Summit 2002. *IDE Spot Survey*, IDE-JETRO.

Matsuoka, S., ed. (2004). *International Development Studies*. Toyo-Keizai-Shinpo-sha, (in Japanese).

Matsuoka, S., Okada, S., Kido, K., and Honda, N. (2004). Development of Social Capacity for Environment and Institutional Change. *Journal of International Development Studies*, 13(2): 31-50, (in Japanese).

Matsuoka, S. (2005). Social Capacity Development for Environmental Management and *International Cooperation, in the second research report. Incubation Research Project 31 "Total Changes in Environment and building the society for regional environmental cooperation in East Asia*" (in Japanese).

Matsuoka, S. ed. (2007). Effective Environmental Management in Developing Countries: Assessing Social Capacity Development. Palgrave-Macmillan, London.

Ministry of Environment, Japan (2005). Quality of the Environment

Morgan, P., and Taschereau, S. (1996). *Capacity and Institutional Assessment; Framework, Methods and Tools for Analysis*, Canadian International Development Agency (CIDA).

Murakami, K., and Matsuoka, S. (2005). Evaluation of Social Capacity for Environmental Management, *Submitted to the Japanese Journal of Evaluation Studies* (in Japanese).

North, D.C. (1990). *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press.

Nose, Y. (1996). *Ecological Community-Roles of Public Health Scholars for the 21st Century*. Kindai-bungeisya (in Japanese).

Tanaka, K., and Watanabe, M. (2005). Social Capacity for Environmental Management and its Indicator Development: Environmental Efficiency Approach, *Presented at the Japan Society for International Development, 16th Conference*, November 26-27, 2005, Kobe, Japan (in Japanese).

UNDP/WHO (1996). *Air Quality Management and Assessment Capabilities in 20 Major Cities*, the Monitoring and Assessment Research Center.

UNDP (1998). *Capacity Assessment and Development in a Systems and Strategic Management Context*, UNDP/BDP/Management Development and Governance Division.

Yeager, T. J. (1999). *Institutions, Transition Economies, and Economic Development*, Westview Press.

Zhang, J., Fujiwara, A., Tanaka, K., Senbil, M., and Nagashima, K. (2004). Methodology of Developing Social Capacity Indicators for Environmental Management: State-of-Art and Future Directions. *Proceedings of the Second Expert Meeting on SCEM Capacity Indicators and Development Process Modeling* (in Japanese).

Article　Invited Paper

Qualitative Evaluation: Evaluating People's Empowerment

Yoko Fujikake

Tokyo Kasei-Gakuin University
yoquita@kasei-gakuin.ac.jp

Abstract

It is challenging to employ quantitative data to evaluate people's empowerment since the criteria for "richness" and "happiness" of people in the area where development practice is undertaken can differ between individuals/groups and can also differ from the criteria of the evaluators. It is yet more difficult for outsiders to evaluate whether gender relationships have or have not changed in people and society as a result of the developmental project.

The aim of this paper is to visualize the phenomena of people's empowerment through the construction of an empowerment evaluation model known as the "Fujikake Model". First, this paper summarizes some of the current trends in qualitative evaluation and empowerment evaluation. Second, it illustrates qualitative changes regarded as empowerment by the target people who participated in the "Improving Living Standards Project", and visualizes the process of the Fujikake Model which is derived from the narratives and daily practices of rural women in Paraguay. Third, this paper studies the application of this model to the "Micro-Business Promotion Project" of the rural women in Honduras, and examines the model's obstacles and possibilities.

Keywords

Narratives, Rural Women, Qualitative Evaluation, Empowerment Evaluation

1. Introduction: Considering Qualitative Evaluation[1]

Qualitative evaluation[2] is the collection and analysis of qualitative data and is classified in two ways: (1) using qualitative data to complement quantitative data so as to offer a broader evaluation of the target society (Mutually Complementary Qualitative Evaluation: MCQE); (2) using qualitative data to evaluate phenomena that cannot be thoroughly explained using quantitative data (Independent Qualitative Evaluation: IQE) (Fujikake 2007)[3].

IQE is composed of data that is difficult to understand quantitatively and cannot be extracted until a rapport has been established between researchers and informants. For example, the analysis of culture and gender issues (domestic violence, 'honor' killings, dowries, female genital mutilation, male suicide, etc.) or issues of religion and gender (family planning and illegal abortions in Catholic countries, etc) would be meaningless if only quantitative data were examined and evaluated. It is necessary to understand phenomena in the target society by examining a few

cases and on that basis generalize about the whole.

The verified cases of IQE are considered just the tip of the iceberg and it may be argued that the data is not objective as there are so few cases. Employing IQE, it is possible to obtain data from different angles, but still relevant to the issue, by interviewing more people. That is, one could use the triangulation method by collecting data from different sources.

The argument of universality when there are many subjects and non-universality and non-validity when there are few subjects, is based on the confusion between "what can be studied through the subject" and "to study the subject" [4]. Even though there is little data and it is qualitative, it can be regarded as objective if the data expresses the reality of the target society.

2. People's Empowerment and Empowerment Evaluation

2.1 People's Empowerment

The author initiated to consider an empowerment of rural women in the Republic of Paraguay in 1993 - a time when the concept of empowerment, though clearly defined within the "*Women in Asia and the Pacific 1985-1993*" (United Nations, 1994), was generally not as widely known as it is today. Currently however, there is increasing interest in evaluating empowerment phenomena in international cooperation projects[5].

It has been encouraging to witness the degree to which the "Beijing Platform for Action (BPFA)", adopted at the 1995 Fourth World Conference on Women, became an "agenda for women's empowerment" and elevated interest in empowerment issues today. As a case in point, the "Millennium Development Goals" (MDGs), which include among them the attainment of gender equality and the empowerment of women, were adopted by the United Nations General Assembly after endorsement by 189 Heads of State in September 2000.

2.2 Definition of Empowerment[6]

In the academic area there naturally exists a variety of definitions for empowerment. To begin, Caroline Moser has given her definition to be:

' *Empowerment is the capacity of women to increase their own self-reliance and internal strength. This is identified as the right to determine choices in life and to influence the direction of change, through the ability to gain control over material and nonmaterial resources*' (Moser cited in Rowlands 1997, p. 15).

McWhirter defines empowerment as:

' *A process whereby women become able to organize themselves to increase their own self-reliance, to assert their independent right to make choices and to control resources which will assist in challenging and eliminating their own subordination...The process by which people, organizations or groups who are powerless (a) become aware of the power dynamics at work in their life context, (b) develop the skills and capacity for gaining some reasonable control over their lives, (c) exercise this control without infringing upon the rights of others and (d) support the empowerment of others in the community*' (ibid, p. 15).

Rowlands argues that:

' *Using the conventional definition, of 'power over', empowerment is bringing people who are outside of the decision-making process into it, this puts a strong emphasis on participation in political structures and formal decision-making and, in the economic sphere, on the ability to obtain an income that enable participation in economic decision-making. Individuals are empowered when they are able to maximize*

the opportunities available to them without constraints. It must also include the processes that lead people
to perceive themselves as able and entitled to make decisions' (ibid, pp.13-14).

These three researchers emphasize that the key to empowerment is increasing women's self-reliance by bolstering their ability to organize for their right to make choices, gain confidence through self-determination, and exercise control over both material and non-material resources in order to alter their subordinate status. It is also safe to say that the researchers have reached a broad agreement to view empowerment as a *process* (Fujikake & Kuroda 2008).

2.3 Evaluating the Phenomenon of Empowerment

"Empowerment Evaluation" developed by David Fetterman (Head of the Evaluation Department, Stanford University, School of Medicine/Anthropology) is representative of empowerment evaluations. Fetterman presented 'Empowerment Evaluation" in 1994 and the "*Foundation of Empowerment Evaluation*" in 2001. He defined evaluation as a process to attain self-determination through self-evaluation and reflection in individuals (or groups) who attempt to improve an undertaking in which they are involved and to improve their situation under their own initiative (Fetterman 2001). Though the empowerment evaluation he proposed is designated as "evaluation", it is in fact a method for a series of processes of self-evaluation and self-reflection of the subjects (ibid., Fetterman Wandersman 2005). Minamoto argues that this evaluation can be applied to social development projects undergoing Development Assistance (Minamoto 2003).

The empowerment evaluation team of the World Bank performs research to assess the level of empowerment in five countries (Alsop & Heinsohn 2005). The key concept here is "Agency" [7]. In this way, attempts to evaluate empowerment phenomenon have been fast developing.

The author has been involved in the "Improving Living Standards Project" of rural women in Paraguay as a development practitioner since 1993, and through this has identified numerous phenomena in women's daily practices that can be interpreted as empowerment. Therefore the author reviewed empowerment and changes in the consciousness of people participating in small-size of social development projects of four NGO's in India, Bangladesh, and Mali (Fujikake 2000, Fujikake ed. 2003).

Though not considered an initial goal, as a result of these four case studies the following conscious and behavioral changes were observed in women: 1) they began to offer their own opinions concerning reproductive health/rights, 2) they gained self-confidences and voiced their opinions, 3) they took action under their own initiative---such as participating in village meetings, 4) they began to have new aspirations, 5) they were able to take pride in themselves, and so on. These changes were also found in the rural women of Paraguay observed from 1993 to 2006.

Each individual views empowerment differently. However, the women who participated in the development projects also regarded the above-mentioned phenomena as 'empowerment'. The author believes that to some degree the phenomenon known as empowerment has universal tendencies. Therefore, on the basis that we must understand the values and reality of their society, a systematic analysis of comments and qualitative data was undertaken and the empowerment evaluation model mentioned in Chapter 3 was elaborated (Fujikake 2000, 2001, Fujikake ed. 2003).

Evaluating the qualitative changes in people's empowerment is a challenging task. Since empowerment phenomena are not countable, and it is difficult for outsiders to evaluate whether gender relationships, in particular whether the personal and social constructions experienced change due to the development project, it is important to establish a framework to aid agency and people in the target society in visualizing qualitative changes.

3. Empowerment Evaluation Model derived by Abstracting a Case Study of the Rural Women of Paraguay[8]

3.1 Overview of the Republic of Paraguay and background of Projects in Village‘ S’
3.1.1 Overview of the Republic of Paraguay
The Republic of Paraguay (Paraguay) has an area of 406,752km² and an approximate population of 6.7 million people, 47% of whom live in rural areas. Located in South America, Paraguay is landlocked, predominantly agricultural, and gained its independence from Spain in 1811. Ninety seven percent of the population is Mestizo, a blend of Guarani and Spanish and the result of government policy during the colonial period. Ninety eight percent profess to be Catholic - another visible result of Spanish rule during the colonial period. The official languages are Spanish and Guarani, with the rural population most commonly using Guarani in their daily lives.

Paraguay was under the dictatorship of Alfredo Stroessner for 35 years from 1954 to 1989, one of the longest reigns in the history of this now democratic country (Fujikake 2002). During this period the people had no freedom of speech, ideology or expression (Arditi 1989).

Based on the World Bank's Poverty Assessment of 2003, *Census Bureau Household Survey*, 41.2% of people in rural areas lack a monthly income sufficient to cover basic necessities, whereas in urban centers this figure is 27.6%. The top 10% of the population hold 43.8% of the national income, while the lowest 10% have only 0.5%.

The economic recession has detrimentally affected income equality, most notably in rural areas where the Gini Index has risen from 0.56 in 1995 to 0.66 in 1999. Similarly, the concentration of land ownership in the Paraguayan countryside is one of the highest in the world: 10% of the population control 66% of the land, while 30% of the rural people are landless (Marío 2004). This inequality has caused a great deal of tension between the landless and the elite (Nagel 1999).

Based on research conducted between 1993-1999 using quantitative and qualitative data, Fujikake concluded that women in rural areas had only limited opportunities to participate in projects due to the structure of male-dominated society (Machismo). Domestic violence against wives, the exploitation of women, and sex-discrimination in employment were evident in rural and urban areas alike (Ziogas 1987, Fujikake 2000, UNIFEM/SM/PRP 2001).

3.1.2 Background Overview of Projects in Village‘ S’
The author was sent to the Agricultural Extension Services (Extension Services) of the Ministry of Agriculture and Livestock in Paraguay (referred to by its Spanish initials MAG) as a volunteer from the Japan Overseas Cooperation Volunteer (JOCV) under JICA (Japan International Cooperation Agency) from January 1993 to February 1995 and supported the Increasing Consumption of Vegetables Project with the goal of improving the lives of the rural people in the area of Colonel Oviedo city (Oviedo) in the province of Caaguazú.

In the years 1993-1995, the women from village‘ S’ wished to acquire more knowledge about vegetable cultivation and nutrition, skills such as cooking, sewing, and knitting; better educational opportunities for the village children, and learn how to generate income by making and selling processed foods. These were the practical gender interests of the women in village‘ S’ during this time.

Figure1 is problem analysis tree from the point of view of the rural women in village S, shows the concerns of the village women and the inter-relationships between their concerns. Here they expressed the desire to learn about nutrition, vegetable cultivation, the production and marketing of processed foods, and schooling. As a result, the following projects (see 3-2) for the" Improving Living Standards Project" were initiated in village‘ S’.

The author analyzed both individual and group changes of consciousness and behavior that emerged from the self-implementation of the" Improving Living Standards Project" at the micro level on the basis of comments and

Figure 1 Problems Analysis Tree from the Point of View of the Rural Women in Village S, Paraguay
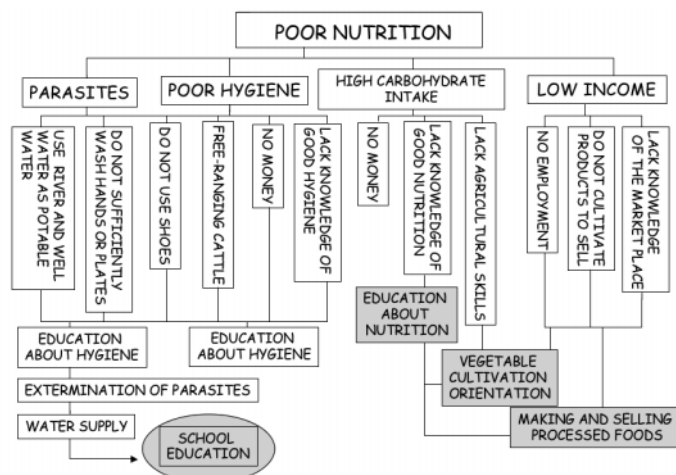
Source: Fujikake (2000)

Figure 2 Fujikake Model 1:THREE TYPES OF RESULTS
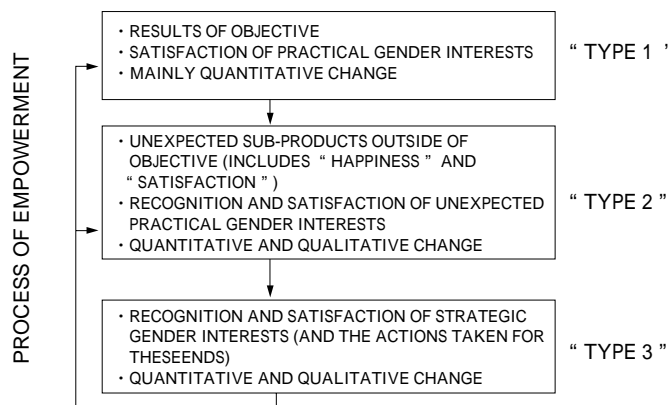
Source: Fujikake (2000, 2001)

practices of the rural women in village 'S' from 1993 to 1999.

Concurrently, comments and practices of the people close to the women (partners, mothers-in-law, fathers-in-law, daughters, sons, people from outside the village, etc.) and indications of realignment to the societal and gender structure were also analyzed (Fujikake 2000, 2004). From the gender viewpoint the author verified factors that encouraged women's empowerment, and at the same time studied the societal level impact of women's empowerment on political practice.

The procedure for the construction of the Fujikake model was as follows: the comments obtained in the research were converted to text, classified according to person and research period, and compiled to compare the frequency the same word or expression appeared so as to extract common traits. The relationship between the women's comments and the aims of the project were converted into the schematic concept drawing shown below (Figure 2) [9]. Follow-up research was conducted in 2001, 2004, and 2006.

3.2. The "Improving Living Standards Project" of the Rural Women of Paraguay.

The initial aims of the "Improving Living Standards Project" (comprised of 3 sub-projects) developed in village 'S' in 1993 were as follows:

  "Increasing Consumption of Vegetables Project"
  a. To increase knowledge of hygiene and nutrition.
  b. To diversify menus.
  c. To increase cultivated vegetable varieties.
  d. To generate income.
  "*Mitai Roga* (A Place for Children) Establishment and Management Project"
  a. To construct and manage an area in which to offer village children the same opportunity to learn Spanish as their urban children.
  b. To use as a multipurpose salon
  "Jam Factory Management and Marketing Project"
  a. To manage a jam factory.
  b. To generate income by selling processed foods

As a result of analyzing rural women's comments and checking them against the aims of the project mentioned above, it became obvious that not only were the initial aims achieved but many sub-products were also obtained.

Changes to the consciousness of the women were classified as Type , Type , or Type as shown in Figure 2. This is the tree types of results (Fujikake Model 1). The rural women personally and mutually evaluated these changes as: "very good" ; "I have changed" ; and "I am not like I was before" [10].

The comments and practices of the rural women were then examined within each sub-project and classified according to the following 12 indicators: 1.Participation, 2.Voicing opinions, 3.Change of consciousness, 4.Taking action, 5.Solidarity, 6.Cooperation, 7.Creation, 8.Setting new goals, 9.Negotiation, 10.Satisfaction, 11.Self confidence, 12.Administrative and economic management (Fujikake 2000, Tables 7-1, 7-2 material: pp.47-48). These were defined as "empowerment indicators" (Figure 3 is Fujikake Model 2:Empowerment Evaluation using 12

Figure   Fujikake Model 2 Empowerment Evaluation using12 indicators
Women's group's Empowerment in the case of village S in Paraguay:1994-1999
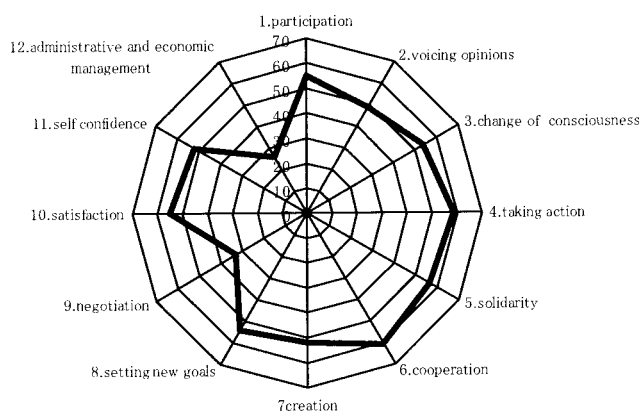
Source: Fujikake (2000, 2001)

Figure 4 Fujikake Model 3:Four core elements of empowerment: women, family and community
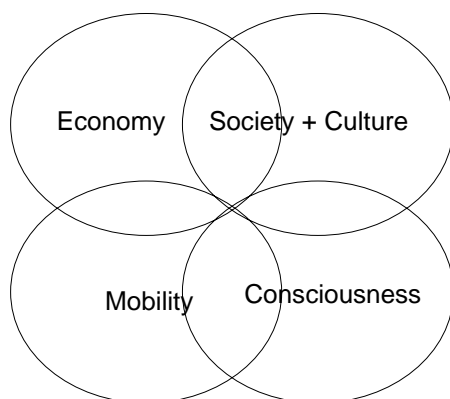
Figure 5 Fujikake Model 4:Level of Empowerment

Economy

Society + Culture

Mobility

Consciousness

International Level

Macro Level: National Policy, National Machinery

Meso Level: Networking with Prefectures, Cities, Other Organizations

Micro Level:Village

Education   Water

Health   Development   Mayor

Individual, Organization

Source: Fujikake (2005, 2006)
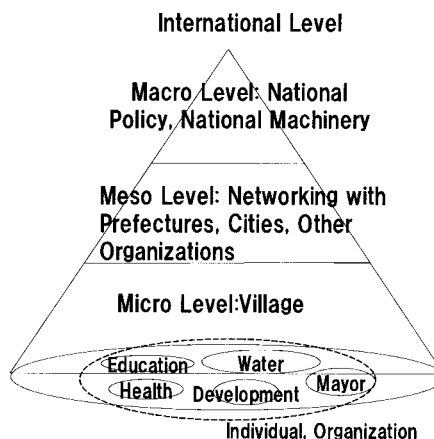
Source: Fujikake (ed.) (2003)

indicators).

The women of village'S'perceived their relative social positioning by leaving their village, negotiating with women in other towns, and by having contact with diverse people such as development assistants and aid fieldworkers (Fujikake 2004). And they began to utilize their past personal history and experience as social capital.

3.3 Visualization of the empowerment process

Having first been converted to text, women's comments found to be related to empowerment indicators were assigned a value of 1 point. These points were totaled for each empowerment indicator to obtain Figure 3 (Women's group's Empowerment in the case of village'S'in Paraguay). In the same way, individual empowerment was visualized using pie charts.

The purpose of quantifying women's empowerment was not to define a low level as" not good" (Fujikake 2000). The purpose was to understand the peculiarities in the individual or group empowerment of village'S', and to analyze how women perceived it (ibid.). This was solely an operation to understand the phenomenon known as empowerment that occurs in the individual, the group, and the target society's consciousness---the differences due to the accumulated past experience of an individual were not analyzed. (ibid.) [11].

Empowerment embodies conscious, social, cultural, economic, and behavioral aspects which are mutually related as shown in Figure 4 as Fujikake Model 3: Four core elements of empowerment (Fujikake 2005, 2006c). The same tendencies were observed in research done in Guatemala (Fujikake 2006c).

As shown in Figure 5 is level of empowerment; Fujikake Model 4, it was clarified in the case of Paraguay that people's empowerment attained different levels: micro, mezzo, and macro. That is, rural women have extended their micro daily and political activities ("Improving Living Standards Project" begun in 1993) to a more macro level [12].

Fujikake Model (1-5) as the Empowerment Evaluation Model has the following characteristics:

(1) Visualizes qualitative and gender relationship phenomena that are observed in the evaluation of the target people (good, bad, undesirable, etc.); (2) The ability to evaluate sensitive gender issues that can be confirmed only through IQE; (3) The analysis of comments can examine not only an individual within and outside of the family, but also changes within the community and people's political practices; (4) Visualizes women's economic and social practice in several way; (5) Can be applied to social development projects in other regions.

4. Application of "Empowerment Evaluations: Fujikake Model" [13]

4.1 Application to the "Micro-Business Promotion Project" for Rural Women in Honduras

The "Promotion of Self-management Enterprises of Women in Rural Areas in Honduras" (the "Micro-Business Promotion Project") is conducting from 2003 to 2008 in Lempira and Copán province, each having different social and gender situations, as a technical cooperation project (JICA, Social Development Department 2006b).

The purpose of the "Micro-Business Promotion Project" was for "women beneficiaries of the project to establish and manage micro-businesses appropriate to local resources", and the overall goal was "to achieve empowerment of the people in the target society" [14].

From the viewpoint of achieving empowerment of the people in the target society, the evaluation of the results of the project was undertaken applying the Fujikake Model.

The author was sent as a Short-Term Expert to guide the management of the gender empowerment indicators of the "Micro-Business Promotion Project". On the basis of research at the project sites, the author modified the following indicators from the Fujikake Model obtained from the Paraguay project: Changed indicator (1) to "Participation" because the concept in Japanese means "to participate in planning" which is not clear in Spanish; Unify into one indicator Solidarity (5) and Cooperation (6); Create another indicator (12) Decision-making [15].

4.2 Results of Evaluation Research

In the "Micro-Business Promotion Project" are 16 target groups with 148 women working in micro-businesses. Seventy-six women from 12 groups were selected for the final evaluation conducted in August 2006 [16]. The effective answer rate was 100.0% and approximately 2 hours were spent on each interview.

These 12 groups scored highly on all 12 empowerment indicators and the conclusion was drawn that empowerment emerged in these groups (see figure 6: Final Evaluation of Project of Promotion of Self Management Enterprises of Women in Rural Area in Honduras) (Fujikake 2006b, JICA Social Development Department 2006b).

Although particular regional characteristics and the gender situation differ between Lempira and Copán provinces, and even vary within the province, the obtained results are noteworthy (ibid.).

Of the 76 women who were individually evaluated, the emergence of empowerment was observed in all participants. Through participation in the training courses of the project, many women were able to earn income to contribute to family finances and thus buy food and medicines for the family. By satisfying their practical gender interests, a change in gender relationships was observed. In other words, the recognition and satisfaction of strategic gender interests enabled women to protest against spousal violence and contribute to family planning and so on.

On the other hand, woman who lacked schooling showed little tendency for change in 11 (Administrative and economic management) and 12 (Decision-making), which are related to strategic gender interests. This tendency was also observed in the case of rural women in Paraguay.

Among the 12 indicators "Administrative and economic management" showed strong growth in the third monitoring (Kuroda 2007), and is thought to be due to conscious reinforcement of members having micro-business experience, the result of intensive training in administrative subjects (ibid).

With these results it is possible to map the course of future follow-ups of the project. For example, a training program can be developed to address 11 (Administrative and economic management) or 12 (Decision-making), dependent upon the empowerment level of each women's group. These evaluation results have had a positive effect on the Poverty Alleviation Policy of Honduras (Fujikake & Kuroda 2008).

Figure 6 Final Evaluation of Project of Promotion of Self Management Enterprises of Women in Rural Area in Hounduras (July. 2007)
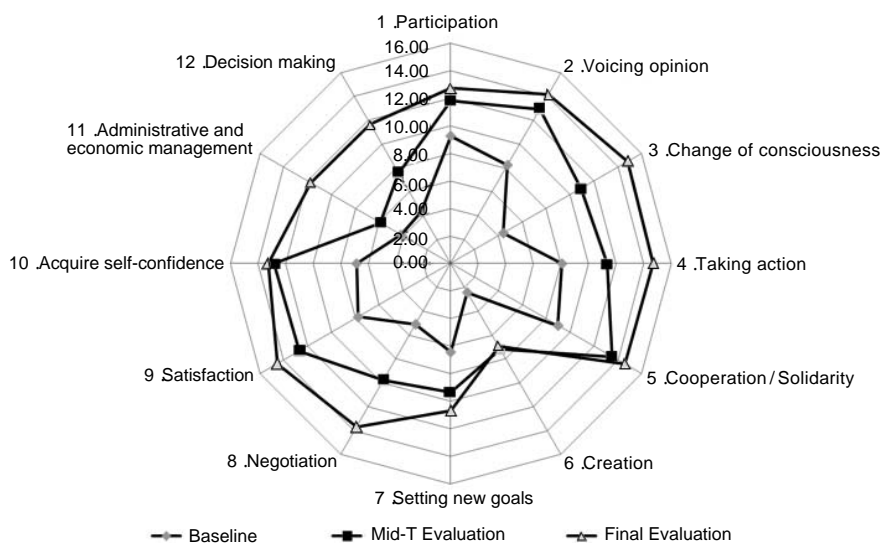
Source: Based on Fujikake Model 2(2000,2001), Project of Promotion of Self management Enterprises of Women in Rural Area in Honduras, JICA (2007)

5. Obstacles and Prospects of the Empowerment Evaluation Model by using Fujikake Model[17]

5.1 Obstacles

To apply the Fujikake Model (especially No.2) in empowerment evaluation, it is necessary to collect and analyze qualitative data, particularly commentaries. However, there are three obstacles and these are elaborated on below.

The first obstacle is that the collection and analysis of qualitative data is a lengthy process.

The second is the necessity of training researchers. "Gender viewpoint" training is essential because among strategic gender interests, matters relating to reproductive health/rights and domestic violence are delicate topics during an interview. In some cases these matters cannot be broached until the interviewer has an understanding of matters relating to gender and a rapport has been established.

The third obstacle, as was the case in Paraguay and shown to apply in Honduras too, are differences due to the past accumulated experience of an individual.

A division of labor occurred as individuals/groups gained experience in the Micro-Business Promotion Project, or a point decrease in the indicator "Participation" was observed when the efficiency of the organization improved[18]. It is necessary to discuss how these situations can be examined using the empowerment indicators. For example, regarding the ending of participation, when it is confirmed how the person concerned feels it is possible that the individual concludes it is better not to participate. This is seen as a "good decision" for the person. It is an issue to be addressed whether in future case studies outsiders evaluate this point as positive or negative.

As a characteristic of the Fujikake Model, it is important that the opinions of the person being evaluated are reflected in the evaluation. In the future it will be possible to address this issue by improving the questions in the interview by making them more detailed.

Figure 7 Fujikake Model 5: Changes in the Scope of
Economic Activities in Rural Women, Paraguay
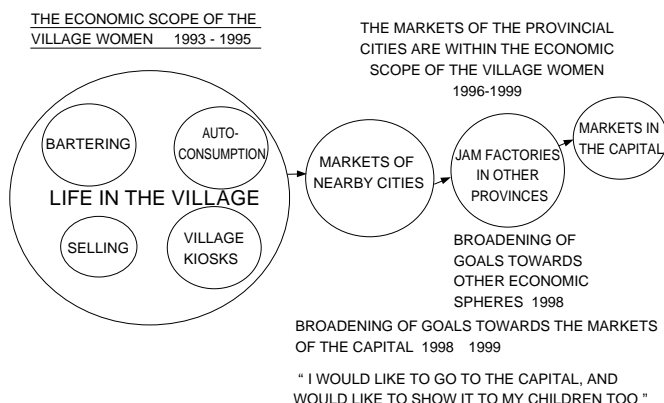
Source: Fujikake ed. (2003)

5.2 Prospects

The Fujikake Model(includes No.1-5), which was an abstract creation from the narratives and/or comments of rural women in Paraguay, South America, could be applied to the interim and final evaluations of the Micro-Business Promotion Project for Rural Women in Honduras, Central America. Therefore it is believed that it was possible to some degree to visualize the process of women's empowerment through participation in the project.

Using this model to evaluate the empowerment of women including behavior changes (see Figure 7), the course of future project follow-ups can be mapped. For example, a training program could be elaborated relating to 11 (Administrative and Economic Management) or to 12 (Decision-making), depending on the level of empowerment of each women's group.

It was shown in the application of this model that there were to some degree common tendencies between South and Central America. In the future it will be necessary to examine the possibility of applying this model in Asia, Middle and Near Eastern countries, and Africa.

6. Conclusion: People's Empowerment and women's Political Practice

Not only are quantitative evaluations required in social development projects but also evaluations of the qualitative aspects. Depending upon the phenomena in the target society, the reality of the society can best be conveyed using the MCQE, or the IQE may be necessary.

If the phenomenon known as people's empowerment is a process, and regarded as changes to gender relationships, power relationships and one of social reform, it is then in other words women's or people's political practice. The Fujikake Model, which was an abstract creation from the comments of rural women in Paraguay, influenced the Poverty Alleviation Policies of Honduras. This model is a new framework that reflects the voice of people in policy and may be regarded as political practice at the grass-roots level (Fujikake 2007a, Fujikake & Kuroda 2008).

The Fujikake Model discussed in this paper was verified by JICA's" Improving Research Methods and Empowerment Indicator Classification" Committee which confirmed that this model has validity, reliability (internal

consistency and high re-productivity) and can be applied to similar projects in the future (Social Development Cooperation Department, JICA 2008). The author thus believes that the reliability of the 12 indicators was ascertained by this study.

Considering the circumstances of the analysis of Paraguayan rural women's comments using cultural anthropology techniques (Fujikake 2000), it is thought the significance of the 12 indicators would diminish if the number of indicators were needlessly increased or decreased. It is possible circumstances necessitate the modification of indicators to some degree in accordance with the social gender situation of the target area; however the author does not agree with proposals, such as to cut by half the 12 indicators with the view of "improving the efficiency of the study". The model would lose its significance---which is the narrative analysis. To the author's gratification, the effectiveness and significance of the 12 indicators were quantitatively substantiated through statistical analysis by JICA research.

This evaluation model differs from Fetterman's empowerment evaluation in that it is more focused on gender relationships. It is necessary to examine this more deeply on another occasion. People's empowerment and changing gender relationships within the social structure and social movement are closely related. The Fujikake Model could be a framework to evaluate these phenomena and affects policy proposals that influence development practice.

Acknowledgement

Notes

1　Extracted from Fujikake, Yoko (2007b)

2　Evaluation can refer to policy evaluation, program evaluation, and project evaluation. In this research paper "evaluation" refers to "project evaluation", except where otherwise noted.

3　Analysis and evaluation using qualitative data can be classified by the quantification of qualitative data, visualization, and hypothesis proposal.

4　Ibid.

5　In September 2006, an evaluation seminar was held in Nepal to increase the efficacy of activity evaluation under the sponsorship of the South Asia Regional Office of UNICEF with more than 100 participants, including government officials from South Asian countries and donor staff, handout of Kusagou, Takayoshi (2007) "Zizokuteki na kaihatsu shien", *Dai nanakai ODA Hyouka Kenkyukai*, held in Tokyo, World Bank, March 13, 2007.

6　This section is reprinted from Fujikake & Kuroda (2007).

7　Please refer to practice theory (Agency) of Fujikake (2004).

8　This chapter is reprinted from Fujikake (2006c).

9　Please refer to Fujikake (2000) and Fujikake & Uemura (2003) for details concerning Figure 1.

10　Please refer to Fujikake (2000).

11　The author will study the issue of how differences in accumulated past experience between individuals can be uncovered.

12　Various analytical concepts are used because the following are deemed necessary: 1) the use of diverse data sources; 2) the implementation of theory triangulation; 3) the implementation of method triangulation (Frick 2002: 282-283).

13　This chapter is reprinted from Fujikake (2006c) and Fujikake & Kuroda (2007).

14　It was discussed the member of the project include author as one of a members of evaluation committee.

15 Discussion with Shihoko Kuroda, Expert of JICA, and Sara Elisa Rosales a local consultant. The Paraguayan case is a post-fact investigation, so that in this paper it is expressed in the past. However, the Honduran case is an ongoing project and is expressed in the present (with a few exceptions).

16 As support and activities were initiated in 2006, final evaluation was not conducted on the remaining 4 groups (as per JICA headquarters' decision) because the term of activity was too brief.

17 This chapter is refered from Fujikake (2006c), Fujikake & Kuroda(2008 pending).

18 From discussions between Mr. Itsuo Kuzasa, Project Expert (leader), Mr Hisao Odagiri (job training specialist), and Ms. Shihoko Kuroda (coordination, gender, and participatory development specialist).

References

Arditi, Benjamín (1989). Adiós a Stroessner; nuevo espacios, viejos problemas, *Nueva Sociedad,* 102, 24-32.

Alsop, Ruth & Heinsohn, Nina (2005). *Measuring Empowerment in Practice: Structuring Analysis and Framing Indicators,* World Bank Policy Research Working Paper 3510(February 2005).

UNIFEM/SM/PRP(Fondo de Desarrollo de las Naciones Unidas para la Mujer y Secretaría de la Mujer, Presidencia de la República del Paraguay)(2001). *Beijin+5: informe nacional-evaluación de la implementación, plataforma de acción, cuarta conferencia mundial sobre la mujer,* Asunción: Paraguay.

Fetterman, David M. (2001). *Foundation of Empowerment Evaluation.* Thousand Oaks, California: SAGE Publications.

Frick, Uwe (2002). *An Introduction to Qualitative Research.* London: SAGE Publications.

Fujikake, Yoko (2000). *Nohson Josei no Empowerment ni Kansuru Kosatsu - Paraguay Kyowakoku "S" Mura no Jumin Josei ga Jisshi shita Seikatsu Kaizen Project no Jirei Yori - (Observations on the Empowerment of Rural Women - From the "Improving Living Standards Project" Carried Out by the Women of Village "S" in Paraguay -).* Master's thesis of Ochanomizu University Graduate School.

_____ (2001). "Project ga Jumin Josei ni Motarashita Shitsuteki Henka no Hyoka ni Mukete - Paraguay Kyowakoku Nohson no Josei ga Jisshi Shita Seikatsu Kaizen Project no Jirei Yori - (The Evaluation of Qualitative Change the Project brought to Resident Women - From the "Improving Living Standards Project" Carried Out by Resident Women in Rural Areas of Paraguay -), *The Japanese Journal of Evaluation Studies,* 1 (2), 29-44.

_____ (2004). *Paraguay ni okeru Campesina no Shutai Kochiku Katei ni Kansuru Kenkyu - Kenkyusha/Chosasha to Jissensha no Ohkan kara Mita Kaihatsu Kyoryoku (Research on the Autonomy Building Process of Campesinas in Paraguay - Development Cooperation from the Viewpoint of the Relationship Between Researchers/Investigators and Practitioner).* Doctoral dissertation of Ochanomizu University Graduate School.

_____ (2005). "Nohson Josei no Empowerment to Empowerment Hyoka - Nanbei Paraguay ni okeru Seikatsu Kaizen Project no Jirei kara Kangaeru - " (Empowerment of Rural Women and Empowerment Evaluation - Observations of the "Improving Living Standards Project" in Paraguay, South America -), *World Trend,* Sep. 2005, Institute of Developing Economies: 30-33.

Fujikake, Yoko (2006a). *Proyecto Fomento de Empresas Autosostenibles para Mujeres en Zona Rural: Evaluación del empoderamiento (The Micro-Business Promotion Project for Rural Women: Empowerment Evaluation),* Seminar (August 10[h], 2006 at the offices of JICA Honduras).

_____ (2006b). *Honduras Kyowakoku Chiho Josei no Tameno Shokibo Kigyo Shien Project: Empowerment Saishu Hyoka Hokokusho (The Micro-Business Promotion Project for Rural Women in Honduras: Final Report on Empowerment Evaluation),* JICA.

_____ (2006c). "Empowerment no Tsubunkasei ni Kansuru Kohsatsu - Paraguay, Honduras, Guatemala de Tenkai sareta Project no Jirei yori - " (Observations on the Cross-culturality of Empowerment - Cases from Projects Carried Out in Paraguay,

Honduras, and Guatemala -), *Summaries Papers from the 17th National Meeting of Japan Society for International Development*, Japan Society for International Development, 152-155.

_____ (2006d). Empowerment Hyoka Model: Seika Sanrui Kei Model ni Kansuru Kohsatsu (Empowerment Evaluation Models: Observations of Models Representing Three Types of Results), *Summaries of the 7th National Meeting of the Japan Evaluation Society - Regionally focused Evaluations -*, The Japan Evaluation Society, 41-46.

_____(2007): "Toward Post Gender and Development(GAD)," *GAD Locally-Diverse Gender Needs*, The International Society for Gender Studies: 18-19.

_____ (2007b). "Shitsuteki Hyoka" (Qualitative Evaluation), In Miyoshi, Kohichi (ed.), *Hyokaron wo Manabu Hito no Tameni (For People Learning the Evaluation Theory)*, Sekaishisosha.

Fujikake, Yoko ed. (2003). *Hitobito no Empowerment no tameno Gender Tohkei, Shihyo to Hyoka ni Kannsuru Kohsatsu - Teiseiteki Data no Katsuyo ni Mukete - (Observations on Statistics, Indicators, and Evaluations of Gender relating to People's Empowerment - Practical Use of Qualitative Data -)* Institute for International Cooperation, JICA (http://www. jica.go.jp/kokusoken/enterprise/chosakenkyu/kyakuin/200306_08.html for download),

Yoko Fujikake & Shihoko Kuroda (2008). "Extraction of an Empowerment Evaluation Model from a case study in Paraguay and its application in Honduras". In Keichi Kumagai (eds.), *Beyond the Differences: Repositioning Gender and Development in the Asian and Pacific Context*, Ochanomizu University.

JICA, Social Development Department (2006a). *Honduras Kyowakoku Chiho Josei no Tameno Shokibo Kigyo Shien Project Unei Shido (Chukan Hyoka) Chosa Houkokusho (Investigative Report on the Project of Promotion of Women's Self-Management Enterprises in Rural Honduras [Interim Evaluation])*, JICA.

JICA, Social Development Department (2006b). "*Honduras Kyowakoku Chiho Josei no Tameno Shokibo Kigyo Shien Project Unei Shido (Syuuryouzi Hyoka) Chosa Houkokusho (Investigative Report on the Project of Promotion of Women's Self-Management Enterprises in Rural Honduras [Final Evaluation])*, JICA.

Kuroda, Shihoko 2006. *Senmonka Gyoumu Kanryou Houkokusyo (JICA Expert Final report in Honduras*, JICA.

Mario, Estanislao Gacitua et al. (2004). *Paraguay: Social Development Issues for Poverty Alleviation.* World Bank Report. June 18, 2007.

Molyneux, Maxine (2001). "*Mobilization without Emancipation? Women's Interests, the State and Revolution in Nicaragua", Women's Movement in International Perspective.* Latin America and Beyond, New York: Palgrave, 38-59.

Minamoto, Yuruko (2003). "Enpawamento Hyouka no Tokucyou to Tekiyou no Kanousei", *The Japanese Journal of Evaluation Studies* 3 (2), 70-86.

Nagel, Beverly Y. (1999). "Unleashing the Fury': The Cultural Discourse of Rural Violence and Land Rights in Paraguay", *Comparative Studies in Society and History*, 41, 148-181, Cambridge University Press.

United Nations (2006). *The Millennium Development Goals Report 2006*, New York, United Nations.

World Bank (2003). *Census Bureau Household Survey.*

Rowlands, J. (1997). *Questioning Empowerment - Working with Women in Honduras*, Oxfam Print Unit.

Ziogas, Marilyn Godoy (1987). *Indias vasallas y campesinas: la mujer rural paraguaya en las colectividades tribales, en la colonia y en la república*, Asunción: Editorial Arte Nuevo.

Research Note   Invited Paper

Designing Participatory Evaluation for Community Capacity Development: A Theory-driven Approach

Koichi Miyoshi

Ritsumeikan Asia Pacific University

miyoshik@ apu.ac.jp

Naomi Stenning

Ritsumeikan Asia Pacific University

Abstract

This article is presented in response to the need for an alternative development that benefits rural communities. By conducting a participatory evaluation of a particular community activity it is possible to bring about changes in the thinking and behavior of the evaluation participants and stimulate an endogenous community capacity development. This is contingent, however, on the ability of the participants to formulate appropriate evaluation questions that specifically address community capacity related issues and processes. In order to do this, appropriate guiding theoretical framework and concepts are required. Thus this paper advocates for a theory-driven participatory evaluation based on a community capacity development and community policy structure model. The model is useful for conceptualizing development in an alternative manner that is more holistic and realistic than the conventional focus on the economic sphere. It will prove a useful framework for both development planning and evaluation, although further elaboration of the model remains as a future task.

Keywords

Community Capacity Development, Participatory Evaluation, Theory-driven Evaluation, Evaluation Design

1. Introduction

Despite significant economic development in many countries rural populations have been left behind and continue to find it difficult to compete with urbanized areas and achieve real improvements in living conditions. The persistent impoverishment of these rural communities around the globe requires a shift away from the conventional economic-centered approach to development. An alternative development is called for. In order to benefit rural communities, an alternative approach to development is grounded in reality and is holistic in its view. It takes into account not only the economic and formal (market) aspects of life but also acknowledges the social, political and informal realms.

 Participatory evaluation has the power to cause changes in society, and therefore affect development, through the influence of the evaluation process on participants in the evaluation (Miyoshi & Tanaka 2001; Minamoto 2008). However, the majority of participatory evaluations are still lacking in terms of theory. Evaluations that are not theory-

driven tend to be method-oriented, or 'black box' evaluations (Chen 1990). These usually focus on one of either experimental or naturalistic methods, but in both cases there is a distinctive lack of theory (Chen 1990). Chen strongly emphasizes the importance of theory in evaluation conceptual frameworks for more systematic identification of stakeholders' needs, selection of an appropriate approach and feedback into action (Chen 2005). Rogers *et al* attribute the use of theory to guide evaluations as resulting in a clearer understanding of why programs work or fail to work, allowing the attribution of certain outcomes to a program, and improving the program in question (2000). This is also true for participatory evaluations. Without theory on which to be based, participatory evaluations are vague and ambiguous and will probably fail to identify the important questions to be asked in the evaluation. This in turn makes it difficult to achieve the sought after changes in the evaluation participants and therefore desired societal change and development.

The authors argue that in order to conduct a participatory evaluation that responds to the need for an alternative development approach in rural communities, a theory-driven participatory evaluation is required. Through the integration of a community capacity and community policy structure theoretical model into participatory evaluation, evaluation participants are able to formulate questions purposely aimed at identifying, clarifying and improving both their capacity and the specific community policy structure in question. By asking the right questions in a participatory evaluation, a change in the consciousness and behavior of the evaluation participants is possible and, in turn, a positive change in society is achievable. Being able to ask the right questions is contingent upon having an appropriate guiding theoretical framework and concepts.

This article describes a theoretical model (community capacity and community policy structure model) to be used in participatory evaluations aimed at bringing about an increase in community capacity. Participatory evaluation design for stimulating community capacity development is also discussed. In order to illustrate that the model and evaluation design are practical and operational, a potential participatory evaluation design for community capacity development is then presented as a case study.

2. Community Capacity and Policy Structure Theory

2.1 Community Capacity Development and Policy Structure Model

The model in Figure 1 illustrates that a community uses its capacity to plan, implement and evaluate community policy structures. This framework allows identification, conceptualization and clarification of community processes through the inclusion of program theory, whilst simultaneously providing a basis for the analysis of community capacity. This model is a dual function model aiming at developing community capacity, and introducing and implementing a higher value added and better well-being policy structure, which consists of economic, social and political activities to change the life of the community's population. The model itself has been formulated based on Chaskin *et al* (2001), Friedmann (1992), Miyoshi *et al* (2003), and the results of research conducted through the implementation of training programs and preparation of training materials under the international cooperation program.[1] This model emphasizes operational aspects of its utilization and aims at providing concrete and practical concepts for the implementation of rural promotion and development. The model can be used in both community development planning and evaluation. The origins of the model and its specific components are elaborated below.

The model adapts the concept of community capacity from Chaskin *et al* who define community capacity as 'the interaction of human capital, organizational resources, and social capital existing within a given community that can be leveraged to solve collective problems and improve or maintain the well-being of that community' (2001, p.7). Chaskin *et al* also describe community capacity as operating 'through informal processes and/or organized

Figure 1 Community Capacity Development and Policy Structure Model

Outcomes

Outcomes of Community Change of Society)

Intermediate Outcomes Change of Target group

Community Policy Structure

Outputs Results of Activities

Activities

Economic
Formal/Market
Informal
Communal
Household

Social
Formal
Informal
Communal
Household

Political
Formal
Informal
Communal
Household

Inputs

Community Capacity

Function
Planning
Implementation
Evaluation

Characteristics of Community Capacity
- Sense of Community
- Commitment
- Ability to Set and Achieve Objectives
- Ability to Recognize and Access to Resources

Strategy or Elements
Human resources
Leadership
Organizations
Networks

Source: Formulated based on Chaskin et al (2001) , Friedmann (1992) and Miyoshi et al (2003)

efforts by individuals, organizations and social networks that exist among them and between them and the larger systems of which the community is a part' (2001, p.7).

Characteristics of community capacity presented by Chaskin *et al*'s study include: sense of community, commitment, problem-solving ability, and access to resources (2001, p.13). The model presented in Figure 1 above includes the characteristics of community capacity as described by Chaskin *et al*, except modifying for' problem-solving ability' which has been supplanted with 'ability to set and achieve objectives' and 'access to resources' with ' recognition of and access to resources'. This is because, in reality, a positive (also referred to as asset-based) approach to development is more viable. Setting realistic community objectives and striving to achieve them using the resources available is essentially a less difficult task then focusing on a negative aspect (problem) in the community and attempting to 'solve' it.

The first characteristic shapes the very existence of the community. It involves the members of the community being aware that they are a community. Important points of' sense of community' include the degree of connectedness and strength of relationships among community members and to what degree they recognize a mutuality of circumstance and share commonly held norms, values and vision. [2]

' Commitment' refers to the responsibility that individuals, groups, and organizations take for what happens in their community. This includes two aspects, both recognizing oneself as a stakeholder or beneficiary in the wellbeing of the community as well as being willing to participate actively as a stakeholder for the betterment of the community.

The' ability to set and achieve objectives' involves translating the abovementioned commitment into action. As mentioned above, the authors prefer to highlight an objectives-oriented approach to action in the community such as appreciative inquiry, as it is more realistic to make progress with rather than ' problem-solving' which denotes negation, criticism and making difficult changes to' fix' the problem. [3] The ability to set objectives is also emphasized. Setting objectives is an indispensable ability for achieving objectives and is required for guiding activities in the appropriate direction.

' Recognition of and access to resources' involves being able to identify and secure the productive use of

various available resources (social, human, economic, physical/environmental, political) both within and beyond the boundaries of the community. This includes being able to forge and make use of linkages between community members (groups, individuals, organizations) and actors in the broader system of which the community is a part (for example in the case of a hamlet or village these would include those in the city or province where it is located). The authors also emphasize the ability to recognize underutilized resources for development in order to encompass a wider scope and increased selection of possible resources.

The 'Strategies' box in the diagram describes both the elements of the community which hold and contribute to community capacity and the 'points of entry' or specific targets for community capacity development strategies. These include: human resources development (individuals); fostering leadership (formal and informal leaders); establishing and enhancement of community organizations (governmental and non-governmental, voluntary, private enterprise, social organizations); and social capital (networks among community members and between them and entities beyond the community boundaries). [4]

The elements of community capacity do not necessarily need to be kept uniform or static. As each community changes and evolves so too does that community's capacity and its elements. Thus it does not make sense to attempt to fix community capacity at a certain level or to aim to apply some kind of recipe as a panacea as each case will differ and the one community's situation will also differ over time. Community capacity should be understood and applied as both a diverse and a flexible concept.

Mainstream doctrine on international development tends to reduce the definition of development to a purely economic consideration, for example GDP per capita, and direct focus on capital accumulation at the expense of other important development issues (Friedmann 1992). This narrow view of development also results in the plight of the rural poor being largely ignored. In order for an effective 'alternative development' it is essential that a broader range of issues is included.

The policy structure part of the model acknowledges non-economic aspects by incorporating social and political activities and allowing for the incorporation of informal as well as formal activities. Community policy structure illustrates the process of community activities through the application of program theory. Community activities that it can be used to describe are many and varied, for example, it might be used to describe and clarify agricultural production processes, specific development initiatives, community events, or informal social processes. This part of the model helps to clarify these processes along the lines of a logical framework (end outcomes, intermediate outcomes, outputs, activities, inputs). [5]

End outcomes signify the eventual change in society due to a certain activity or process. Intermediate outcomes are the changes in the target groups' behavior or situation that are expected to lead to the end outcome. Outputs are the results of activities usually expressed in numerical terms (number of workshops conducted, for example). Activities, on the other hand, are the actual workshops themselves (for example), whilst inputs include any resources used for a certain policy structure such as funds, local human resources, external experts and so on.

The community's selection of policy structure and its successful implementation depends heavily on the community's capacity situation. Moreover, as the community develops or upgrades its capacity, the community will naturally transfer to or select a new and more sophisticated or value-added policy structure. The process of a community using its capacity to plan, implement or evaluate a certain policy structure may also contribute to developments in community capacity, particularly if these result in expected end and intermediate outcomes (changes in society and specific target groups).

2.2 Note on Community

At this point, clarification of how the model perceives the community is required. Community is a social system

confirmed by administrative boundaries and in which the members (organizations, groups, and individuals) recognize themselves and each other as belonging to the same community. The community definition utilized by many in the community development literature is comparatively narrow encompassing only the residents of a sub-area of a town or city, referred to as 'a neighborhood', which results in analysis of the community based solely on the situation of the residents of the neighborhood (see Chaskin *et al* 2001; Gittel & Vidal 1998). Analysis based on the model presented in this paper includes those wider stakeholders that exist in addition to the residents of a community. These include government institutions, civil society organizations, NGOs and NPOs, private enterprise, educational institutions and so on. With this broader and more inclusive definition of community and wider range of units of analysis it is possible to construct a more worthwhile policy-oriented argument. [6]

2.3 Community Capacity Development

The strengthening of community capacity is referred to as community capacity development. Community capacity development involves seeking out and presenting latent, or potential, advantages and opportunities in the community in order to promote 'positive neighborhood change' (Chaskin *et al* 2001). Capacity is developed through the community's attempts to develop or maintain these identified advantages and opportunities.

The latent advantages and opportunities within a community are represented by the potential of community components, their characteristics and their functions which lie dormant within a community and are able to be changed. Through identifying and focusing on these latent advantages and opportunities and by activating interactions between and among the various elements of community capacity, potential or latent advantages and opportunities can become real and active.

Emphasis should be given to community capacity development as a non-linear and continual process. Moreover, capacity development should be approached in a way that seeks to encourage endogenous development through the activation of interactions and synergies between the elements of capacity, rather than as an outside intervention. Community components, their characteristics and functions should not be simplified into a conceptualization of linear, one-way or easily categorized inter-relationships, as reality is not that straightforward. For example, a rise in the individual's capacity may contribute to the improvement of a community organization, whilst a rise in organizational capacity likewise may contribute to an improvement in abilities at the individual level.

3 Participatory Evaluation: A Theory-Driven Perspective

3.1 The Influence of Participatory Evaluation

By viewing the evaluation process as participatory[7], stakeholders are not taken as passive or disempowered subjects or objects of the evaluation; rather they are viewed as participants with an active role and stake in the evaluation process. [8] This is in line with the principles of participatory development, which has gained prominence in recent years.

In terms of the influence of evaluation, historically conventional non-participatory evaluation has heavily emphasized the results of the evaluation (the evaluation report) and how these results are utilized (Johnson 1998; Shulha & Cousins 1997 cited in Kirkhart 2000, p.9). Supplementing the evaluation results, participatory evaluation has an additional and very important element of influence; the actual process of conducting the participatory evaluation.

The evaluation process itself is becoming recognized as being independent from the evaluation results and as having the power to bring about changes in the people or organizations participating in the evaluation (Kirkhart 2000,

p. 10; Miyoshi & Tanaka 2001; Minamoto 2008). By involving stakeholders (including project staff, project participants, beneficiaries etcetera) it is possible to conduct an evaluation from the direct viewpoint of those most intimately involved in the project or program. Furthermore, through experiencing the evaluation process firsthand these stakeholders are given the opportunity to increase their consciousness of their own stake and influence in the project, thus strengthening stakeholder ownership and contributing to improvement of the project as a result. In this way, it is expected that the implementation of participatory evaluation will contribute to a societal change (Miyoshi & Tanaka 2001).

The influence of the evaluation process works through the interaction of answering the evaluation questions and the opportunity of the evaluation questions themselves in terms of awareness, and the emotional and political facets of the lives of stakeholders who participate in the evaluation. In terms of awareness, the focal point is promoting the understanding of the evaluation participants' own roles in the project or activity. On the emotional side, feelings of value or self-worth that individuals and groups gain from participating in the evaluation process are most important. In terms of the political facet, the evaluation process itself may be used to create a new dialogue, re-ignite interest in social issues, and also raise the stakeholders' awareness of the actual conditions of their rights and the prevailing political environment. In the case of administrative reform and where evaluation is regarded as an explicit intervention to an assistance activity, the political dimension of the evaluation process' influence has been particularly emphasized (Kirkhart 2000; Miyoshi 2001).

Effects of the evaluation process are not limited to the short-term. Influence of the evaluation process does not necessarily cause an instant change in the stakeholders of the project or activity. There are effects that occur later, in the long-term, and there are also those influences which will continue to accumulate to bring about an eventual change. And, although the influence can still be identified as coming from the evaluation process, there are instances where the effects grow stronger, little by little, over a period of a number of months or even years. The long-term influences of evaluation are those effects which develop over the course of time, covering a wider area in the process. The strongest influences of evaluation are often those which do not appear at the point in time when the evaluation is conducted and, therefore, are not necessarily obvious. In fact, it is the long-term effects of evaluation that become institutionalized by the organizations and society involved in the evaluation and which become more evident only over time. This reality requires thinking in this kind of context. When longer-term influences are the intention of the evaluation influence, it is important to adopt a long-term approach and frame of mind when designing the evaluation.

Participatory evaluation will produce an effective influence when the evaluation facilitator clearly defines the target group/s they would like the evaluation to affect, consciously recognizes the desired influence of the evaluation, and where the future outcomes of that influence are carefully envisaged. The shape of the desired, or expected, influence of the evaluation and who will be affected by such influence are, therefore, important points to consider when actually conducting the evaluation. Usually the evaluation is implemented with a certain intention in regards to the influence of the evaluation (to stimulate community capacity development, for example). [9]

3.2 Rationale for a Theory-driven Participatory Evaluation

In order to result in the desired influence, the participatory evaluation must be based on an appropriate and relevant guiding theoretical framework and concepts. The paragraphs above reveal that participatory evaluation has the power to cause changes in society, and therefore affect development, through the influence of the evaluation process on participants in the evaluation. However, as mentioned in the introduction section, the majority of participatory evaluations are still lacking in terms of theory, making it difficult to both define the desired influence of the participatory evaluation and to achieve it. Without theory on which to be based, participatory evaluations are vague and ambiguous. Without an idea of the desired influence and important issues in the participatory evaluation it is very

difficult to know what questions need to be asked and answered.

Through the integration of a community capacity and community policy structure theoretical model into participatory evaluation, evaluation participants are able to formulate questions purposely aimed at identifying, clarifying and improving both their capacity and the specific community policy structure in question. By asking the right questions in a participatory evaluation, a change in the consciousness and behavior of the evaluation participants is possible and, in turn, a positive change in society is achievable. Being able to ask the right questions is contingent upon having an appropriate guiding theoretical framework and concepts. The theoretical model presented in Section 2 (and Figure 1) of this article provides the guiding framework and concepts needed in order to guide the definition of the problem and to formulate questions that will both evaluate and result in improvements in community capacity.

The introduction itself of these community capacity and program theory concepts for discussion within the community may also entail long-term benefits for the community. Although here introduction of the concepts is advocated for in order to guide a participatory evaluation, it is probable that, through learning these concepts and discussion of these concepts through the course of the participatory evaluation, the community members will remember and utilize them in the planning and implementation of later projects (even if this is at a sub-conscious level). In this way it is envisaged that the introduction of the community capacity development and policy structure model will provide continuous and long lasting benefits for the community, even after completion of the participatory evaluation.

3.3 The Evaluation Questions

The power of the participatory evaluation to stimulate components of the community and cause community capacity to develop is contingent on the participants being able to formulate and answer meaningful evaluation questions. These questions should be concerned with the characteristics of community capacity (sense of community, commitment, community ability to set and achieve objectives, and the awareness of and access to community resources) and be related to the specific community policy structure that is the subject of the evaluation (the end outcomes, intermediate outcomes, outputs, activities, and inputs of the subject). They are to be formulated and answered by the participants in the evaluation after being provided with the relevant knowledge (guiding theoretical framework and concepts) by the evaluation facilitator, who will also assist community facilitators to facilitate question-formulating sessions. The following are representative of the type of questions that may be addressed:

Sense of community: To what degree do the community members have strong relationships with each other? To what degree do members hold common values and norms, and a shared awareness of the condition of the community including future goals?

Commitment: What kind of responsibility do individuals, groups and organizations belonging to the community take for what happens in their community? As stakeholders, in what way do community members view themselves in terms of community outcomes? Are community members willing to actively participate in community activities?

Ability to set and achieve objectives: What kind of awareness is there of the challenges that arise due to changes in the environment surrounding the community? What kind of awareness is there of objective setting methods/processes? What do people think of the existence of relevant mechanisms? Is there an awareness that these are functioning?

Awareness of and access to community resources: What kind of awareness is there of the existence of human resources, leadership, organizations, and networks in the community? What kinds of relationships are being built between the community and the larger social system? What kind of awareness is there of the utilization of

economic, human, physical, and political resources available within and outside of the community?

3.4 Participants in the Evaluation

Participants in the evaluation are members of the community which the evaluation concerns. They may include implementers of the activities/interventions and the beneficiaries or people affected by such activities, and individuals and representatives from every organizational level. In terms of organizations, these may include government agencies, public service agencies, social organizations and associations, private enterprises, religious organizations, health-related organizations, and civil society organizations such as NGOs or NPOs.

In terms of stimulating participation, focus has usually been centered on the administrative executer of project activities however the participation of civil society organizations such as NGOs/NPOs to implement evaluation is becoming an increasingly possible option. Here it is important that those who have the intention of developing community capacity, as the facilitator of the evaluation, take the main role in guiding the implementation of the evaluation. Furthermore, regarding the evaluation participants, it is necessary to take into account the existing human resources, leadership, organizations and networks and their capacity and willingness to participate. However, even if this kind of independence is secured, in order to conduct an evaluation, there are particular skills that are required and these will be supplied by the expert (most probably external) evaluation facilitator.

3.5 Role of the Evaluation Facilitator

The role of the external evaluation facilitator is important in order for the evaluation to be truly participatory and effective. A facilitator will usually have an academic background in the social sciences and is either a social science researcher or development practitioner (UNDP 1997). He/she should have a significant amount of field experience and also have the ability to: listen; guide and facilitate discussions; encourage trust; delegate tasks and responsibilities; plan actions to bring together the viewpoints of different stakeholders; and create an environment of reflection and sharing (UNDP 1997).

The facilitator will also need to introduce the concepts of community capacity development, community policy structure and participatory evaluation to the community counterparts, ensuring that at least the selected community counterpart facilitators have a good and thorough understanding of these. After the community counterparts are equipped with such knowledge the external facilitator will take a supportive role and 'act as a catalyst or stimulator, managing the evaluation process without being seen as directing it' (UNDP 1997). Thus the evaluation facilitator will assist the counterpart facilitators and manage the entire evaluation process, particularly in terms of the technical and practical aspects of the evaluation questions, participant identification and management, and the evaluation timeframe.

3.6 The Participatory Evaluation Timeframe

Regarding the time period of the participatory evaluation, rather than simply selecting when to answer the evaluation questions, it is important to concentrate on the potential contribution of the particular timeframe to community capacity formation. The advantages and opportunities that the community possesses should be identified and a period of time that appears to allow for the most progress in terms of a positive change and when conscious participation will also be strong should be selected. The community itself is a continuous organization and within it continuing events and activities are held. These community events or activities present a good opportunity for an evaluation period that allows the community to be informed and involved.

Events which have potential include music concerts, sports carnivals, festivals and so on which are held by the municipality, prefecture, or community-based groups/associations. Of these types of events there are many which do

not have specific results or outcomes and it is relatively simple to produce evaluation questions regarding community capacity. Many activities are conducted on an annual basis, year after year, and by utilizing this situation and through conducting ex-post evaluation and traversing the process of making clear the activity's outcomes, how it should be run, and its significance (through the process of evaluation) it is relatively simple to create an opportunity to answer the intended evaluation questions.

It may be especially fruitful to pick up development activities as the subject of the participatory evaluation. Development activities are limited to a certain timeframe of their own for implementation, and when the activities have been completed, the interest of members also diminishes. This is not limited to 'hard' infrastructure activities/projects, but also occurs in education, health, welfare and other such 'soft' activities. Thus, it is at the time of commencement or during the earliest stages of development activities that have the most potential to incite a high level of interest and willingness to participate that will, in turn, allow for concrete influence on community capacity.

An appropriate evaluation timeframe should be negotiated and decided with the community participants taking into account such factors as the amount of time and frequency that participants are able and willing to dedicate themselves, the availability of other resources such as funding to be used in the evaluation, how quickly decisions can be made in a participatory manner given the cultural context and inter-relationships and group dynamics in the community, the length of the actual implementation cycle of the policy structure (activity) in question, and to what extent the participants will choose to include supporting activities within the evaluation study. Depending on these factors the timeframe allowed for the evaluation might be anywhere from just a couple of weeks to months or even an entire year. Also, participatory evaluations are often conducted in phases. The UNDP gives the norm of about 10-person weeks of consultant's services in phases spread over three to six months (UNDP 1997).

3.7 Knowledge Sharing for Participatory Evaluation

According to Jackson and Kassam, participatory evaluation is a 'process of self-assessment, collective knowledge production, and cooperative action' (1998, p.3). This section is primarily concerned with the 'collective knowledge production' aspect of this process. It is through sharing knowledge among community members that collective knowledge production can occur. In order to facilitate this, much consideration is needed regarding the types of data, methods of data collection, and how to present results to the wider community. These decisions should be made whilst taking into account the purpose and objectives of the participatory evaluation, the nature of the evaluation questions posed, as well as the social and political context of the community.

Community participation can also extend to the data collection activities of the participatory evaluation study. Merits of having community members participate in data collection include the higher probability that information will be more quickly and easily accessed and opinions expressed by other community members or interviewees may be more direct and honest than if the researcher were an outsider (Gittell 1980). [10] Also this approach would allow for a higher level of intra-community interaction and higher quantity and quality of knowledge sharing.

The possibility of community members participating as researchers will inevitably depend on the education and abilities of the individual members and the evaluation facilitator will need to monitor and be responsible for ensuring that such data collection activities are conducted in such a way as to facilitate meaningful knowledge sharing. These community researchers might need special training and/or supervision from the external evaluation facilitator in collecting data and sharing knowledge for the study. Furthermore, even if community participation is pursued in the participatory evaluation data collection, the results of participant observation by the external evaluation facilitator during his/her time engaged in the community will also be invaluable and should be utilized.

4. Exploring the Potential of a Theory-Driven Participatory Evaluation for Community Capacity Development in 'Imori-dani'

Due to the nature of participatory evaluation and the requirement for stakeholder involvement and input from the planning and design stages onwards it is not a worthwhile effort to provide a detailed evaluation design without this involvement. Nonetheless, in order to illustrate further and apply the concepts outlined above, this section outlines a preliminary design for a participatory evaluation designed to both assess and influence community capacity in a case community, 'Imori-dani'. [11] The mock evaluation design below outlines ideas for pre-planning and preparation, the evaluation questions, data gathering and analysis, and reflection and action, allowing for clarification of the intention of the participatory evaluation prior to its implementation.

4.1 Introduction to Imori-dani

Imori-dani (Newt Valley) is the self-appointed nickname of Matsumoto, an agricultural-based hamlet in the administrative area of Ajimu Town in Oita Prefecture, Japan. The relatively small community consists of just 56 households. Many activities are held in Imori-dani throughout the year including lotus, flower, wine, *Bon* (ancestor worshipping) and music festivals, rice planting, soybean sowing and harvesting, lake clearing, mushroom cultivation and harvesting, buckwheat noodle making, and traditional charcoal burning. These activities are spread out over the year so that there is a frequency of about one activity or so per month. They are planned and held by one or more of the community's groups[12] and all Imori-dani residents as well as visitors from outside of the village (for example people from large neighboring cities such as Oita City or Fukuoka City seeking a rural experience) are invited and encouraged to participate.

    Community development in Imori-dani is conducted in a participatory style with input from all members including children and the elderly (for example in creating the community vision and future objectives) and has been quite successful to date. [13] Due to the participatory mechanisms already in place and the active nature of the community, it would be practicable to conduct a participatory evaluation and the current progress of the community in terms of development (i.e. quite recent with a certain level of momentum) imply that Imori-dani would benefit by reviewing their community's capacity at this particular point in time.

    During community visits by the authors and consultation with Mr. Nimiya, the head of Imori-dani's farmer's cooperative, the subject of participatory evaluation and how Imori-dani might benefit from conducting such an evaluation surfaced. Thus, there is the possibility that Imori-dani may decide to try implementing a participatory evaluation for community capacity development in the near future.

4.2 Pre-planning and Preparation Phase

First and foremost the evaluation will require support and input from community members. Thus one of the most crucial activities during the pre-planning stage is to garner such support at least from a few key community members. This could be done through existing communication mechanisms such as an announcement or discussion at the farmer's cooperative or *Imori-dani Kurabu* meetings. A special planning meeting could also be held at the community centre in order to openly discuss and decide on the objectives and purpose of the evaluation and to select a specific community policy structure (subject activity) to be evaluated, an evaluation timeframe and to identify evaluation facilitators and stakeholders/participants and clarify their roles in the evaluation study.

    For the purpose of this paper it will be assumed that the community members decided to create a temporary special participatory evaluation committee within the structure of the *Imori-dani Kurabu* to jointly facilitate a participatory evaluation (along with an outside evaluation facilitator) of the *hotaru no sato no ongakukai* (Firefly

Hometown Music Festival) held annually in June.[14] The purposes of the evaluation should be identified by the community participants with the outside facilitator providing guidance and support to ensure that the process takes a direction that will benefit community capacity. The stakeholders might decide that the purpose of the evaluation is to assess the community's capacity to hold the festival in order to recognize community strengths and identify areas that could be improved and to stimulate community capacity development in the process.

Training in participatory evaluation facilitation and tools/techniques should be provided to at least two of the evaluation committee members nominated to be counterpart facilitators to ensure that the community counterpart facilitators are equipped with the skills to facilitate a participatory evaluation. Stakeholders/participants that may be identified include: festival organizers and performers; young people, elderly people, women, and newcomers[15] in the community; audience/festival-goers from both within and outside the community; farmers; stall-holders and so on. A description of the policy structure of the activity should also be constructed and presented at this stage in order to clarify processes and to assist with question formulation.

An appropriate time-frame should also be decided on at this point jointly by the evaluation committee and the external evaluation facilitator. For the purpose of this article it is assumed that the time-frame of 10 person-weeks over a period of three months was selected. This would allow time for the introduction of the concepts and framework, training of community evaluation facilitators and community researchers, evaluation question formulation sessions, data collection (over the festival preparation period, the festival itself, and during the time directly after the festival), and a reflection and action phase.

4.3 Evaluation Question Formulation Phase

Creating the evaluation questions is the next step after the initial planning and preparation phase. Asking the right kinds of questions is key to a successful and capacity-influencing participatory evaluation and, in order to ensure that the evaluation causes community members to reflect on their capacity to implement the festival, it is important that the characteristics of community capacity are kept in mind when formulating the questions. In order to link community capacity issues with the specific policy structure of the music festival, the evaluation questions should also be based on the description of the policy structure (based on a logical framework) that was produced in the evaluation planning stages. The evaluation committee could host a question-creating session with a representative group of stakeholders or a series of sessions with different groups. Workshop-style methods such as using sticky-notes might work especially well. Table 1 illustrates the possible kinds of questions related to the community capacity characteristics as well as to the festival policy structure that might be produced. They include both community capacity and community policy structure related questions.

4.4 Data Collection Phase

The evaluation committee could be charged with forming a small team of community data gatherers. Due to the relatively small size of the community and the focused nature of the evaluation, it would be a more worthwhile endeavor to focus on qualitative rather than quantitative data. Potential methods include in-depth interviews, workshops or focus-groups, and observation (for example at the festival itself). One option for data collection involves a select number of (preferably diverse) stakeholders each producing a photo-essay or video documentary of activities related to the festival from their point of view. [16] These could be shown publicly to elicit wider community response and discussion. Other specific techniques that may be considered include: social mapping, testimonials, and participant observation (by the external evaluator). [17]

Table 1 Imori-dani Example Evaluation Questions

Policy Structure Components	Community Capacity Related Questions	Community Policy Structure Related Questions
End Outcome	Sense of community: - Are there common values for holding the music festival in Imori-dani community? - How strong are the common values for holding the music festival in Imori-dani community? - Do the community members feel that effects on the community due to holding the festival are important? - Is there a shared awareness about the condition of the community to hold the festival by Imori-dani community members?	- What changes happened in Imori-dani by conducting the festival? - Does the change of Imori-dani due to the festival fulfill priority needs of the community of Imori-dani? - Is the change of Imori-dani community due to the festival being realized as expected? - Will the festival be able to continue for the future of community?
Intermediate Outcomes	Commitment: - As members of Imori-dani community do they strongly expect the effects of festival to the changes of the Imori-dani community? - As members of Imori-dani community do they recognize their relation actively to implementation of the festival? - Do they believe firmly in the existence of appropriate mechanisms for their participation in the festival? - Do they have the intention and willingness to participate actively in the festival? Ability to set and achieve objectives: - Do members of Imori-dani community define the expected effects of the festival clearly? - Do they clearly understand the necessity of the festival for the Imori-dani community?	- What groups were most affected by the conducting of the festival? - Are effects of the festival enough for those groups? - Are effects of the festival what the participants in the festival expected? - What kinds of effects were expected by the participants in the festival? - Are the effects to those groups due to the products or services which were produced by the festival activities?
Outputs	Sense of community: - What kind of responsibility do individuals, groups and organizations of the Imori-dani community take for what happens in their community during the conduct of the festival? Commitment: - Do members of the Imori-dani community clearly understand what products and services are preferable as the result in terms of their participation in the festival?	- What kind of products and services were produced by the festival activities? - Were they what the participants in the festival activities expected? - Are they achieved using as little input of resources as possible?
Activities	Sense of community: - Do community members have the clear intention to work together for the success of the festival? Commitment: - Do Imori-dani community members actively participate in the festival (plan, set-up, perform, as audience)? - Are community members willing to actively participate in festival activities? Ability to set and achieve objectives: - Do they think existing mechanisms for producing the expected effect of the festival in terms of the change of the surrounding environment are sufficient? - Do they think that those mechanisms function properly?	- What kinds of activities are conducted to produce products and services of the festival? - Were those activities properly arranged? - Were those activities systematically arranged? - Do those festival activities have enough momentum to be sustainable and continue into the future?
Inputs	Recognition and access to resources: - Do the members of Imori-dani community show awareness of the existence of human resources, leadership, organizations, and networks for the conduct of festival inside and outside the community? - Do they know properly the utilization of economic, human, physical, and political resources available within and outside of the community?	- What kinds of resources are utilized for festival activities? - Are resources for festival activities sufficient for conducting the festival successfully? - Are such resources obtainable for future implementation of the festival? - What groups of Imori-dani community participated actively in the festival? - Who participated from outside of Imori-dani?

Source: The authors

4.5 Reflection and Action Phases

The results of the evaluation should be transparent and made available to the community members. The more traditional method of this is the evaluation report. However, in order to create an opportunity to reflect on the process and plan action based on the outcomes of the evaluation, a stakeholder meeting in the form of a presentation and discussion or workshop might be best. This would be especially pertinent if some of the data collection methods involved visual media. Participants in the evaluation should also be allowed a chance to express what they felt about the participatory evaluation process itself and whether they felt that they gained by being involved. The results of the

evaluation should cover not only the areas of community capacity that could be improved but those which were highlighted by the evaluation as being particularly strong in Imori-dani. Action could also be planned based on these positive outcomes.

5. Conclusion

This article has presented a response to the need for an alternative development that benefits rural communities. By conducting a participatory evaluation of a particular community activity it is possible to bring about changes in the thinking and behavior of the evaluation participants and stimulate an endogenous community capacity development. This is contingent, however, on the ability of the participants to formulate appropriate evaluation questions that specifically address community capacity related issues and processes. In order to do this, appropriate guiding theoretical framework and concepts are required. Thus the authors have presented a theory-driven participatory evaluation based on a community capacity development and community policy structure model. This model is useful for conceptualizing development in an alternative manner to the conventional focus on the economic sphere. It will prove a useful framework for both development planning and evaluation, although further elaboration of the model remains as a future task.

Notes

1　Ritsumeikan Asia Pacific University (APU) conducts Japan International Cooeration Agency (JICA)'s group training programs of technical cooperation for rural promotion and development. The programs include: Training Course in Community Capacity and Rural Development for ASEAN countries; Training Program on the "One Village One Product" Movement in Tunisia; and Training Course in Seminar for Municipal Mayors of Clustered LGUs: The Philippines. APU is also preparing to conduct a training program for African countries.

2　See community psychology literature such as Sarason (1974) or McMillan & Chavis (1986).

3　For resources on appreciative inquiry consult: http://appreciativeinquiry.case.edu/

4　On social capital see Putnam (1993).

5　For more on logical frameworks refer to Funnell (1997) and JICA (2004).

6　See Stenning & Miyoshi (2007).

7　See, for example, Miyoshi & Tanaka 2001.

8　The participation of stakeholders, however, should not be confused with mere providers of ideas or information. By active participation we refer to involvement in the design, formulation of questions, implementation, reporting and discussion of the evaluation. In other words, stakeholders should take an active role in the entire evaluation process, rather than being passive subjects or mere information-givers of the evaluation.

9　Naturally, on the other hand, there may also be effects that were not intended or expected by the evaluator. These occur when the influence of the evaluation reaches people and organizations through unconventional processes not predicted or expected by the evaluator.

10　See comments on a participatory research approach by Marilyn Gittell (1980). Gittell also notes that including and training community researchers in the study allows for a transfer of research skills that will prove valuable to the individuals, their organizations and the community later on after the study has been completed.

11　The authors visited Imori-Dani several times for research as well as study tours for the JICA training programs on community capacity and rural development. The evaluation has not been implemented. It is intended as an illustrative example.

12 The Matsumoto Farming Cooperative and the Ajimu Matsumoto Imoridani Kurabu play main roles in the community. Subordinate organizations of these include the Volunteer Fire Brigade, Senior People's Club, Children's Club and agricultural groups such as the Hyotan Pond Management Association and the Common Forest Management Association.

13 Matsumoto (Imori-Dani) was awarded the Emperor's Cup in the town building division of the 2004 Agriculture, Forestry and Fisheries Festival.

14 This festival was first organized in 1996 and provides an opportunity for people in Imori-dani to reflect on their situation in the context of community development.

15 'Newcomer' in Imori-dani refers to those individuals/families that have recently moved to Imori-dani (usually from the city). People who left the community and then returned are called 'U-turners' and 'I-turners' are those who moved there for the first time. The community has an open and friendly attitude (unlike many small rural communities) and welcomes new residents as this has been identified as a rejuvenation strategy for the community, which, like much of Japan is faced with an aging and declining population. Imori-dani proudly boasts that most of the 12 elementary school age children in the community belong to newcomer households.

16 This is a method used in the photo-voice community development projects developed by Caroline Wang, refer to: http://www.photovoice.com/

17 For a range of participatory methods and tools please refer to UNDP (1997) or World Bank (1996)

References

Chaskin, Robert J. Prudence Brown, Sudhir Venkatesh and Avis Vidal. (2001). *Building Community Capacity.* New York: Aldine De Gruyter.

Chen, H.T. (1990). *Theory-Driven Evaluations.* Newbury Park: Sage.

Chen, H.T. (2005). *Practical Program Evaluation,* Newbury Park: Sage

Friedmann, John (1992). *Empowerment: The Politics of Alternative Development.* Oxford: Blackwell.

Funnell, S. (1997). Program logic: an adaptive tool for designing and evaluating programs. *Evaluation news and comment,* 6(1): 5-17.

Gittell, M. (1980). *Limits to Citizen Participation: The decline of community organizations.* Beverly Hills: Sage.

Gittell, Ross and Avis Vidal. (1998). *Community Organizing: Building Social Capital as a Development Strategy.* Thousand Oaks: Sage.

Jackson, Edward T., Yusuf Kassam. (1998). *Knowledge Shared: Participatory Evaluation in Development Cooperation.* Connecticut: Kumarian Press.

JICA, Office of Evaluation and Post Project Monitoring, Planning and Evaluation Department (2004). *JICA Evaluation Handbook: Practical Methods for Evaluation.* Tokyo: JICA.

Johnson, R.B. (1998). Toward a Theoretical Model of Evaluation Utilization, *Evaluation and Program Planning,* 21(1), 93-110.

Kirkhart, Karen E. (2000). Reconceptualizing Evaluation Use: An Integrated Theory of Influence. In Caracelli, Vallerie J. & Preskill, Hallie (Eds.), *The Expanding Scope of Evaluation Use, New Direction for Evaluation.* San Francisco: American Evaluation Association, JOSSEY-BASS, 5-23.

McMillan, D. and Chavis, D. (1986). Sense of Community: A definition and theory, *Journal of Community Psychology,* 14 (January 1986), 6-23.

Minamoto, Yuriko (2008). Sankagata Hyouka no Rironn to Jissen (Theory and Practice of Participatory Evaluation) in Miyoshi Koichi (Eds.) *Hyoukaron wo Manabu Hito-no Tamameni (For People Studying Evaluation Theory),* Sekai-shisosha, 95-112, (Japanese Translation by authors)

Miyoshi, Koichi (2001). Feedback of Evaluation: Influencing Stakeholders through Evaluation, *Journal of International*

Development Studies, 9(2): 71-86, (in Japanese)

Miyoshi, Koichi (2005). Kokusai Kyouryoku no Hyoka (Evaluation of International Cooperation), in Utsumi, Seiji (Eds.), *Kokusai kyouryoku wo Manabu Hitono-tameni (For People Studying International Cooperation Theory)*, Sekai-shisosha, 256-277 (Japanese Translation by authors).

Miyoshi, Koichi, Yayoi Tanaka (2001). Future of Participatory Evaluation: Concept and Utilization of Participatory Evaluation, *The Japanese Journal of Evaluation Studies*, 1(1): 65-78.

Miyoshi, Koichi, Satoshi Morita, Yoshio Aizawa (2003). Toward Constructing More Suitable Program Theory for Japan's Evaluation: Focusing on International Cooperation Evaluation and Policy Evaluation, *The Japanese Journal of Evaluation Studies*, 3(2): 40-56.

Putnam, Robert D. (1993). *Making Democracy Work: Civic Tradition in Modern Italy.* New Jersey: Princeton University Press.

Rogers, Patricia, Hacsi, Timothy, Petrosino, Anthony, and Huebner, Tracy (2000). Program Theory in Evaluation: Challenges and Opportunities. *New Directions for Evaluation*, 87: 5-13.

Sarason, S.B. (1974). *The Psychological Sense of Community: Prospects for a Community Psychology.* San Francisco: Jossey-Bass.

Shulha, L.M., and Cousins, J.B. (1997). Evaluation Use: Theory, Research and Practice since 1986, *Evaluation Practice*, 18(3), 195-208.

Stenning, Naomi. (2007). Decentralization and Community Capacity: A Case Study of Community Capacity in the Context of Decentralization in Indonesia, (Master Thesis: Ritsumeikan Asia Pacific University) unpublished paper.

Stenning, Naomi and Koichi Miyohi (2007). Evaluting Community Capacity Development: The Case of Oyama-cho, *Proceedings: 8th Annual Conference of the Japan Evaluation Society*, 239-244

UNDP (1997). Who are the Question-Makers? A Participatory Evaluation Handbook, available online at: http://www.undp.org/eo/documents/who.htm

World Bank (1996). The World Bank Participation Sourcebook, available online at: http://www.worldbank.org/wbi/sourcebook/sbpdf.htm

Research Note　Invited Paper

Evaluation Systems and Changes
in Local Government Personnel Consciousness

Makiko Yokoyama

The University of Kitakyushu
myokoyama@ kitakyu-u.ac.jp

Abstract

This paper examines the introduction and effectiveness of evaluation systems by local governments in Japan. We particularly focus on whether such systems alter personnel commitment and motivation (referred to here collectively as" consciousness"). An analysis is provided of data obtained in a large questionnaire, conducted by the author in 2004. Our findings suggest that the introduction of evaluation systems for the purpose of instituting or improving policies, programs, and projects induces a change in consciousness among local government personnel. Furthermore, findings indicate evaluation systems generate long term benefits. In particular, we suggest that making the results of evaluations open to public scrutiny leads to changes in consciousness among personnel.

Keywords

1. Introduction

Since Mie Prefecture introduced a project performance evaluation system in 1995, the number of local government areas in Japan that have introduced similar evaluation methods has increased substantially. According to an investigation[1] undertaken by the Ministry of Internal Affairs and Communications in October 2006 (released March 2007), evaluation systems have been introduced in 96% of all prefectures and ordinance-designated cities. Evaluation systems have been introduced in 90% of heartland cities, 89% of specially designated cities, 48% of other cities and special wards, and 16% of towns and villages. It is expected that evaluation methods will also be introduced in those areas not currently using such methods[2].

As evaluation systems come to be used widely in local government, it becomes increasingly important for their efficacy to be researched. At present little is known about the effects of these evaluation systems and there are limited

Table 1 Purpose of Using Evaluation Systems (%)

	Multiple Answers	Most Important
Improvements in Citizen Satisfaction	57.9	18.1
Improvements in Quality of Administration	62.2	15.7
Increased Administrative Efficiency	83.5	21.1
Consciousness Reform in Personnel	86.0	13.9
Improvements to the Organization	22.6	0.6
Improvements in Resource Allocation	48.8	6.0
Increased Accountability	79.3	11.4
Improvements in the Policy Formation Capability of Personnel	30.5	1.2
Others	13.4	7.2
Number of Observations	166	166

Note: Questionnaire results 2004.
Source: The Author

data on which to conduct such analyses.

In 2004 we conducted a questionnaire in these local government areas in Japan. In this survey we requested responses from the supervisor division, which is responsible for carrying out the evaluation activities, of the local government concerned. Table 1 shows respondents' answers on the purpose of the evaluation systems in local government. Multiple answers were allowed. It shows that 86.0% of respondents selected "consciousness reform in personnel" as a reason, while 13.9% rated consciousness reform as the most important reason for using the evaluation system[3]. These responses highlight the importance that supervisors attach to using evaluation systems for changing personnel consciousness. Other perceived benefits included improvements in administrative functions and citizen satisfaction.

Given the perceived benefits of implementing evaluation systems in local government in relation to changing personnel consciousness, it is important to ask what factors actually lead to changes in personnel consciousness. We assume that changes to personnel consciousness through the application of evaluation systems affect their policy-making capabilities and administrative abilities.

In this paper, we use quantitative methods to examine, the influence of evaluation systems on personnel consciousness. The presence or absence of consciousness change among personnel becomes the dependent variable and is subsequently analyzed by independent variables via the practical method of evaluation system results, such as opening them to the public, and specific factors that induce consciousness change among personnel.

The structure of this paper is as follows. The following section sets out the hypotheses for examining the factors used to change consciousness among local government personnel, and describes the data used for analysis. In Section 3, we show the variables and estimation method in this paper. We verify the relationship between consciousness changes in personnel and evaluation systems and analyze our results in Section 4. We conclude our paper in Section 5.

In this paper, the term "evaluation" is used as a comprehensive concept embracing both the evaluation system and its associated administration.

2. Hypotheses and Data

2.1 Hypotheses

It is important to understand the value of the practical application of evaluation results. Relating staff evaluations to actual job requirements (for example, the creation of policies, programs, or projects) aids employee motivation. Moreover, using evaluation results for personnel and organizational management (for example, for promotion) and making employees aware of the purpose of these evaluation systems induces change in personnel consciousness. Therefore, our first two hypotheses are: "Hypothesis 1: If evaluation results are utilized for personnel' job specifications, personnel consciousness will change", and "Hypothesis 2: If evaluation results are used in organizational management, personnel consciousness will change".

Following Yokoyama (2006) who suggests that the longer term evaluation systems improve administrative services, we propose "Hypothesis 3: The longer the evaluation systems are in place, the more personnel consciousness will change". This hypothesis is based on the assumption that it will take time for changes in consciousness to become apparent.

Hatry, a scholar of evaluation research in the United States, discusses the importance of ongoing staff training when measuring performance (Hatry 1999)[4]. Furukawa explained that "shortage of specialization, lack and/or neglect of training for evaluating" was one of "the seven traps" into which local government can fall, and he described the need for implementing changes in local government management planning (Furukawa and Kitaōji 2004). From this we have developed: "Hypothesis 4: If an information session is held before/after conducting an evaluation, personnel consciousness will change".

Case studies were undertaken following the introduction of evaluation systems in Nagahama City and Suginami City (Ishihara 1999; Ueyama, Tamamura, and Iseki 2000). A survey of the local governments of three municipal districts in the Greater Tokyo metropolitan area shows changes made to the evaluation systems subsequent to this review enable personnel to see that management was prepared to listen (Furukawa 2003; Yokoyama 2003). This understanding enables us to develop "Hypothesis 5: If leadership is exercised in introducing and establishing an evaluation system, personnel consciousness will change".

The handbook developed for the introduction of evaluation systems was altered, on the advice of the then Study Group of the Ministry of Home Affairs, to suggest that evaluation results should be made public (Sinseiki Jichi Kenkyu Kai 2000). The reasoning for this advice was that making such results public enhances the motivation of personnel'. In Mie Prefecture, which has an advanced evaluation system, Umeda, the official who implemented the administrative evaluation, has supported this idea, arguing that publishing evaluation results has had a good effect on the personnel (Umeda 2000). It follows then that we can formulate "Hypothesis 6: If the evaluation results are made publicly available, personnel consciousness will change".

2.2 Data

The following section provides an analysis of data taken from the 2004 questionnaire on "evaluation systems in local government". This survey was conducted in 304 local governments of the prefectures, cities, and special wards in Japan that were introducing or attempting to introduce evaluation systems. These local governments were chosen following the results of the investigation by the Nikkei Research Institute of Industry and Markets (2002). A response was requested from the person in charge who supervises evaluation. The questionnaire was distributed and returned by post. The investigation commenced in April 1, 2004 and the survey was implemented over approximately two months from June 2004. The response rate was high, with 215 local governments returning questionnaires (36 prefectures and 179 municipal districts). Only cities and special wards were analyzed, and data from all prefectures

Table 2 The Correspondence Relationship of the Contents of the Independent Variables
and Six Hypotheses

the Contents of the Independent Variables		Hypothesis
Formation of New Policies, Programs, and Projects	YES=1, NO=0	Hypothesis 1
Improvement in Past Policies, Programs, and Projects	YES=1, NO=0	
Formation and Revision of Plans	YES=1, NO=0	
Boosting Morale of Personnel	YES=1, NO=0	Hypothesis 2
Organization Management	YES=1, NO=0	
Personnel Management	YES=1, NO=0	
Budget Compilation	YES=1, NO=0	
Duration of Introducing Policy Evaluation	Elapsed Years in the Investigation Time	Hypothesis 3
Duration of Introducing Program Evaluation	Elapsed Years in the Investigation Time	
Duration of Introducing Evaluation for All Projects	Elapsed Years in the Investigation Time	
Duration of Introducing Evaluation for Some Projects	Elapsed Years in the Investigation Time	
An Information Session BEFORE Implementing Evaluation	The Synthetic Variable†	
Publicity Campaign on Evaluation System for Personnel	YES=1, NO=0	
Training and/or Workshop	YES=1, NO=0	
Briefing Session	YES=1, NO=0	Hypothesis 4
An Information Session AFTER Implementing Evaluation	The Synthetic Variable†	
Publicity Campaign on Evaluation System for Personnel	YES=1, NO=0	
Training and/or Workshop	YES=1, NO=0	
Briefing Session	YES=1, NO=0	
Leadership of the Head of the Local Government in Introducing the Evaluation System	Head's Initiative=1、Others=0	Hypothesis 5
Leading Officers' Initiative to Establish Evaluation System	Leading Officers' Initiative=1、Others=0	
Making Evaluation Results Open to the Public	The Synthetic Variable†	Hypothesis 6
Making the Complete Evaluation Sheet Open to thePublic	YES=1, NO=0	
Making the Summary or Report of Evaluation Results Open to the Public	YES=1, NO=0	
Publishing Evaluation Results on the Web Site	YES=1, NO=0	
Publishing Evaluation Results in the Public Relations Magazine	YES=1, NO=0	

Note1:The variable with† which added the three or four following variables together.
Note2: Questionnaire results 2004. † added the three or four following variables.
Source: The Author

are reported in this paper.

Table 2 shows six hypotheses and the corresponding independent variables used for analysis. The independent variables set up the variable regarding the evaluation system and the contents of activity according to the hypotheses described in 2.1.

For practical application of evaluation results, multiple answers were permitted to the question on " the information session before/after implementing evaluation" and " opening evaluation results to the public"[5]. To gain greater understanding of both " information session" and " opening evaluation results", we prepared a synthetic variable that combined all variables. We also allowed for a comparison of the influence at a macro level with the influence of the concrete contents at a micro level.

The variable " duration of introducing evaluation" was defined by the policy level used for evaluation. There were named " policy", " program", and " project". Additional introductory periods of " all" and " some" were prepared separately for " project"[6].

Moreover, the leading personnel in the variable called " establishing an evaluation system" refers to those in the chief class section in this paper. Each is dependent on the above-mentioned questionnaire result as a candidate for evaluation. Details of the dependent variable are given in 3.1.

3. Variables and Estimation Method

3.1 Dependent and Independent Variables

The purpose of this paper is to use statistical methods to investigate what induces change in the consciousness of

local government personnel. First, the dependent variable used in this paper needs to be identified. In our questionnaire the following question was asked regarding changes to local government personnel consciousness:

"Did you notice any change in personnel behavior or consciousness after the introduction of the evaluation system?"
1. YES   2. NO

When assessing the results of this question it is important to keep in mind that even if the person in charge of the evaluation system section answers "YES", we have no way of judging whether the perceived change has a negative or positive meaning. However, many comments by respondents in the free answer space added to this question indicate that this change had a positive meaning. Examples included, "increased awareness of costs'", "increased accountability'", "[he/she] having come to have a new appreciation of the purpose of their work", and "increased number of personnel who understand the relationship between their job and the purpose of the local government", In addition, we can also see in these answers the process of change: "having come to participate actively in office discussions about their job, objectives, projects, and so on." [7]. Of course, while not all personnel displayed a changed consciousness, it is important that of those who did there was a positive reaction. Using the qualitative feedback as a guide, we can regard the answer "YES" as a positive change in personnel consciousness. For that reason, this question is used as the measure of change in personnel consciousness. Since it is the reply by the person in charge of carrying out the evaluation activities, the study is limited by the validity and reliability of conceptual grasp and directivity in detecting consciousness change in personnel. However, we believe that this person has the most information about the evaluation system in their local government, and the relevant personnel.

Second, we consider the explanatory variables. As we discussed in the previous section, we use the effect of mechanisms on personnel consciousness as an explanatory variable. These variables are categorized into five groups. The first category relates to the purpose of introducing an evaluation system, which is related to Hypotheses 1 and 2. The second group indicates the evaluation duration effect. This group corresponds to Hypothesis 3. Variables in the third class are used to investigate whether information about the evaluation system is important or not. The fourth and fifth categories relate to the importance of leadership and openness of evaluation results. The last three categories correspond to Hypotheses 4, 5, and 6, respectively.

3.2 Estimation Method and Models

As the dependent variable based on this question is binary, we use the probit model to find the elements that induce consciousness change.[8] Local government either experiences changes in personnel consciousness ($Y = 1$) or does not ($Y = 0$). We believe that a set of factors (expressed as matrix X) explains this difference, so that:

$$P(Y = 1) = F(X, \beta)$$
$$P(Y = 0) = 1 - F(X, \beta)$$

1

In these equations, $P(A)$ stands for the probability that event "A" happens. The set of parameters β reflects the impact of changes in factors on the probability. The model that applies the normal distribution to the right-hand side of the first part of equation (1) is named as the probit model. By increasing the number of explanatory variables included in the estimation model, we can check the robustness of the estimation results. Therefore, we have estimated seven models, each with a different number of independent variables.

4. Interpretations of Estimation Results

4.1 Estimation Results: Overview

The estimation results are summarized in Table 3. On the left side of this table, we list the hypothesis number; the related contents of the corresponding hypothesis are itemized to the right of the hypothesis number.

Our first question is: "what kind of application of the evaluation results promotes change in government personnel consciousness?" The estimated results are summarized at the top of Table 3. Applying the results to the formation and improvement of policies, programs, and projects, and to organizational management raises personnel consciousness, although personnel management itself does not have any influence over personnel consciousness. Although applying the results to broad work such as "the formation and revision of plans" and "budget compilation" has no influence on consciousness, applying results to routine work may be effective. Hypothesis 1 and Hypothesis 2 are, substantially supported.

Next, to test Hypothesis 3, we introduce as explanatory variables the duration of the evaluation system for policies, programs, and all (or some) projects. We can also use these variables as an indicator of the influence of task level on personnel consciousness. The estimated coefficient of the duration of the evaluation system for all projects is statistically significant although the other variables were not influential. Applying the evaluation system for the entire project is useful in improving personnel consciousness. That is, for many public servants it is better to apply the evaluation system to more familiar aspects of projects. This result is consistent with the result for Hypothesis 1 and Hypothesis 2. In order that we might consider the effect of time, the independent variable of "square of duration" was supplied to model 7. Even if not all results were significant, a linear relationship was shown. We have checked the tendency for continuation of the effect by continuation. We add further analysis of this effect in 4.2.

Our question related to Hypothesis 4 is whether information sessions before and after the introduction of an evaluation system are effective or not. Although there are models of the influence of publicity campaign to personnel, the significant result is indicated to be before/after the introduction of evaluation. The synthetic variable which was a composite of all variables was not significant. This result has not confirmed robustness. We found that these activities do little to improve employee motivation. Furthermore, it is not significant who initiates the introduction and administration of evaluation systems. The results, therefore, are not consistent with Hypotheses 4 and 5.

From the estimated coefficients of variables concerning ways to bring the evaluation results into the public domain, bringing the evaluation sheet *as is* into the public domain is efficacious in bringing about changes in the consciousness of public servants. By doing so, employees are placed under pressure as they feel they are being scrutinized by the public. Other ways of publishing evaluations are not effective. That is, the best way to promote change in personnel consciousness is to expose *all results as they are*.

4.2 Plotting the Marginal Effect

Using the estimated coefficients of Model 6 in Table 3, we have the probabilities "Formation of New Policy, Program, and Project: Yes and No" and "Making Complete Evaluation on Sheet Public: Yes and No" as a function of "Duration of Evaluation System for All Projects", at the mean of other variables. Figure 1 shows the former two functions: "Settlement of New Policy, Program, and Project: Yes and No" over the range of "Duration of Evaluation System for All Projects" observed across five years. The effect of applying the evaluation results to the settlement of new policy, program, and project is substantial. Figure 2 shows the latter two functions: "Making Complete Evaluation on Sheet Public: Yes and No", This figure also indicates that making complete evaluation sheets open to public is efficacious in bringing about change in personnel consciousness.

In addition to the effect of application and of making evaluation results public, these figures also show that a

Table 3 Estimation Results

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Hypothesis 1 and Hypothesis 2	Formation of New Policies, Programs, and Projects	0.346	0.421	0.543 *	0.551 *	0.647 **	0.643 **	0.633 **
		0.246	*0.263*	*0.277*	*0.283*	*0.282*	*0.280*	*0.290*
	Improvement in Past Policies, Programs, and Projects	0.700 **	0.775 **	0.880 ***	0.864 ***	0.926 ***	0.962 ***	0.874 **
		0.295	*0.310*	*0.322*	*0.331*	*0.330*	*0.334*	*0.345*
	Formation and Revision of Plans	0.031	-0.002	-0.126	-0.168	-0.141	-0.133	-0.128
		0.274	*0.279*	*0.306*	*0.303*	*0.308*	*0.312*	*0.311*
	Boosting Morale of Personnel	1.375 **	1.304 **	1.471 ***	1.487 ***	1.391 **	1.409 **	1.408 **
		0.536	*0.540*	*0.520*	*0.556*	*0.552*	*0.558*	*0.563*
	Organization Management	1.169 **	1.065 **	1.298 **	1.452 ***	1.483 ***	1.569 ***	1.572 ***
		0.519	*0.489*	*0.526*	*0.546*	*0.554*	*0.568*	*0.574*
	Personnel Management	-0.411	-0.505	-0.616	-0.634	-0.541	-0.541	-0.571
		0.580	*0.546*	*0.554*	*0.571*	*0.551*	*0.532*	*0.519*
	Budget Compilation	0.106	0.064	0.054	0.082	0.093	0.092	0.115
		0.240	*0.246*	*0.257*	*0.266*	*0.271*	*0.271*	*0.279*
Hypothesis 3	Duration of Introducing Policy Evaluation		-0.007	0.017	0.013	-0.036	-0.030	0.414
			0.118	*0.128*	*0.128*	*0.129*	*0.128*	*0.526*
	Duration of Introducing Program Evaluation		0.034	-0.035	-0.012	-0.002	-0.030	-0.129
			0.097	*0.111*	*0.113*	*0.111*	*0.112*	*0.381*
	Duration of Introducing Evaluation for All Projects		0.101 *	0.103 *	0.119 **	0.143 **	0.137 **	0.253 *
			0.059	*0.057*	*0.058*	*0.060*	*0.059*	*0.131*
	Duration of Introducing Evaluation for Some Projects		-0.064	-0.075	-0.062	-0.072	-0.073	-0.064
			0.061	*0.062*	*0.064*	*0.066*	*0.066*	*0.179*
	Square of Duration of Introducing Policy Evaluation							-0.125
								0.134
	Square of Duration of Introducing Program Evaluation							0.032
								0.098
	Square of Duration of Introducing Evaluation for All Projects							-0.017
								0.015
	Square of Duration of Introducing Evaluation for Some Projects							0.001
								0.033
Hypothesis 4	An Information Session BEFORE Implementing Evaluation			0.102				
				0.140				
	Publicity Campaign on Evaluation System for Personnel				0.358	0.452 *	0.440 *	0.429
					0.255	*0.263*	*0.260*	*0.263*
	Training and/or Workshop				-0.429	-0.390	-0.380	-0.389
					0.282	*0.294*	*0.297*	*0.295*
	Briefing Session				0.309	0.324	0.289	0.213
					0.296	*0.301*	*0.307*	*0.313*
	An Information Session AFTER Implementing Evaluation			-0.075				
				0.139				
	Publicity Campaign on Evaluation System to Personnel				-0.606 **	-0.696 **	-0.649 **	-0.698 **
					0.280	*0.294*	*0.298*	*0.313*
	Training and/or Workshop				0.119	0.053	0.032	0.043
					0.244	*0.249*	*0.250*	*0.262*
	Briefing Session				0.181	0.128	0.140	0.170
					0.226	*0.230*	*0.234*	*0.240*
Hypothesis 5	Leadership of the Head of the Local Government in Introducing the Evaluation System			0.095	0.235	0.227	0.247	0.203
				0.245	*0.261*	*0.272*	*0.277*	*0.282*
	Leading Officers' Initiative to Establish Evaluation System			0.031	-0.022	0.001	0.002	-0.015
				0.287	*0.295*	*0.292*	*0.290*	*0.292*
Hypothesis 6	Making Evaluation Results Open to the Public			0.210 **	0.268 ***			
				0.084	*0.090*			
	Making the Complete Evaluation Sheet Open to the Public					0.870 ***	0.848 ***	0.829 ***
						0.263	*0.276*	*0.285*
	Making the Summary or Report of Evaluation Results Open to the Public					0.194	0.135	0.132
						0.264	*0.287*	*0.288*
	Publishing Evaluation Results on the Web Site						0.251	0.158
							0.278	*0.289*
	Publishing Evaluation Results in the Public Relations Magazine						-0.160	-0.145
							0.299	*0.299*
	Pseudo R-square	0.113	0.142	0.191	0.227	0.249	0.254	0.261
	Number of Observations	168	168	163	163	163	163	163

Note: The values in italic face are standard errors. *p<0.10, **p<0.05, ***p<0.01
Source: The author

Figure 1 Marginal Effects of Applying Evaluation Results to Policy Formation
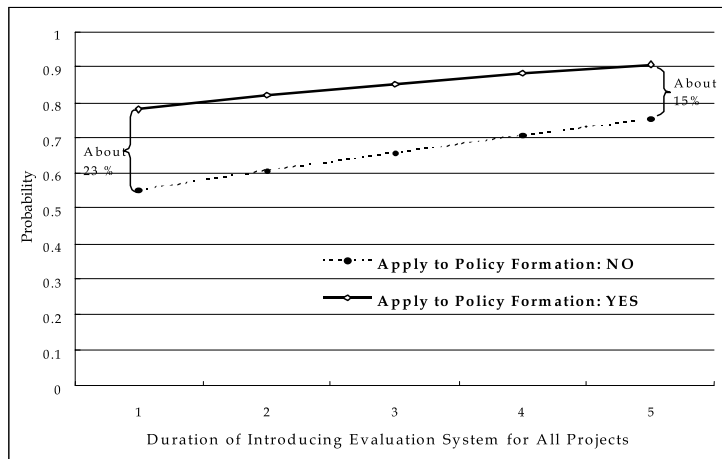
Note: Questionnaire results 2004.
Source: The Author

Figure 2 Marginal Effects of Making the Complete Evaluation Sheet Open to the Public
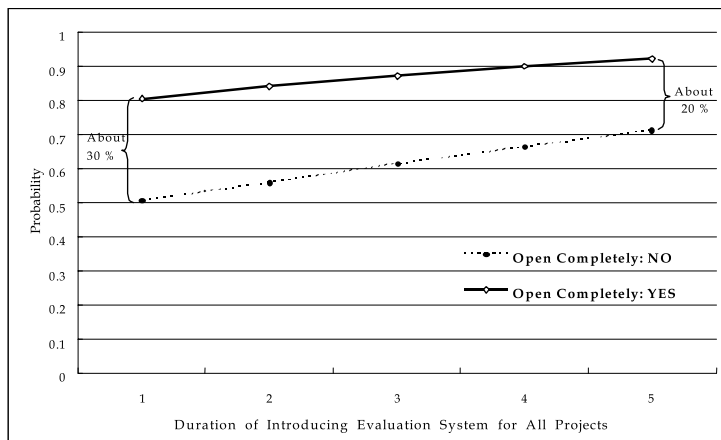
Note: Questionnaire results 2004.
Source: The Author

longer duration of introducing evaluation systems for all projects improves personnel motivation. For example, in Figure 1, the probability difference between"Yes"and"No"at the end of five years is less than at the end of one year. A difference in the probability should be observable at an early stage. However, as described in 4.1, the effect of continuation is sustained. That is, continuation is one of the most important factors for improving personnel consciousness.

5. Conclusion

Our purpose in this paper was to use a quantitative framework to identify factors that promote changes in personnel consciousness. We found that applying evaluation results for various purposes, for example, to form and improve policies, programs, and projects, to enhance personnel motivation to work, and to manage the organization, increases the professionalism of local government personnel. Allowing the evaluation results to be made public provides additional pressures. That is, the increased pressure associated with the introduction of evaluation systems serves to motivate employees. We have also found that adopting evaluation systems has both immediate and long-term benefits, and the long-run cycle of evaluation, application, and transparency provides the necessary pressures to motivate local government personnel.

However, there are some issues still to be examined. In this paper, actual change in personnel consciousness did not tauten the scope of analysis and actual change in behavior was not necessarily observed. As described in 3.1, we believe that benefits of the application of evaluation systems include the fulfilling of personnel capability through heightened cost-consciousness and accountability, and increased awareness of policies, programs, and projects among employees. It remains necessary to examine in more detail the influence that evaluation systems have on administrative activities. We prepared and verified the hypotheses with reference to a review of previous research and case studies. The same may be said of the contents of the evaluation systems. While we obtained results that suggest that making evaluation results open to public scrutiny helps to alter personnel consciousness, there may be variations that occur within this process (for example, by allowing the public to know the name of the person who conducted the evaluation). Questions such as these form part of our future research.

Notes

1 "The measure situation survey of evaluation in local government (*Chihōkōkyōdanntai ni okeru gyōsei hyōka no torikumi joukyou shirabe*)".

2 The ratio for the past fiscal year shows that 5.5% of cities and special wards and 3.8% of towns and villages introduced evaluation systems in municipal districts (Ministry of Internal Affairs and Communications 2007).

3 In this paper, we refer to personnel commitment and motivation as "consciousness".

4 As mentioned above, the term "evaluation" in this paper is understood in the larger sense; we would therefore include performance measurement.

5 We provided a free answer space in the variable "boosting morale of personnel" and received responses such as "for training of a key person", "in order to raise awareness of costs", "to draw up a plan" and so on.

6 This paper typically asks much at every policy level in the investigation of evaluation in local government. Moreover, surveys such as that of the Ministry of Internal Affairs and Communications show variation in the introductory phase for every policy level (Ministry of Internal Affairs and Communications 2006). In this paper, "policy evaluation" and "program evaluation" mean "evaluation for policy level" and "evaluation for program level".

7 There are also local governments that extracted a part of the result based on the personnel questionnaire on completion of evaluation and have expressed objective views about the descriptive content of the answer.

8 We also used the logit model in analyzing our results. As the results of the logit model were very similar to those gained via the probit model we have not reported them in order to save space.

References

Furukawa, Shun'ichi (2003). NPM rejime ni okeru jititaishokuin no isiki kōzō: hyōka no juyōtaido to seido sekkei . *Public Administration Review Quarterly*, (101), 3-18.

Furukawa, Shun'ichi, and Kitaōji, Nobusato (2004). *Kōkyō bumon hyōka no riron to jissai: seifu kara hieiri soshiki made*. Tokyo: Nihon kajo shuppan.

Hatry, Harry P. (1999). Performance Measurement: Getting Results. Washington, DC: The Urban Institute.

Ishihara, Toshihiko (1999). Chihōjichitai no jigyōhyōka to hasseishugikaikei. Tokyo: Chuokeizai-sha.

Ministry of Internal Affairs and Communications (2007). Chihōkōkyōdanntai ni okeru gyōsei hyōka no torikumi joukyou shirabe, March.

Nikkei Research Institute of Industry and Markets (2002). Jichitai no minkan kanri shuhou: sinka suru gyōsei hyōka. *Nikkei regional economic report*, (393), 1-24.

Ueyama, Shinichi, Tamamura Msatoshi, and Iseki Tomonobu(2000). *Jissen Gyōseihyōka: Jirei, Kaisetsu, sosite Q&A*. Tokyo:Tokyo Horei Shuppankai.

Umeda, Jiro (2000). Experiences in the Introduction of Policy Evaluation: Systematic Resistance within the Mie Prefectural Government to the Introduction of the Project Performance Evaluation. *Public Policy Studies Association Japan* (CD-ROM).

Yokoyama, Makiko (2003). A Note on the Consciousness of Personnel toward Evaluation in Local government: Cases of Suginami, Soka and Mitaka Cities. *The Tsukuba University Journal of Law and Political Science*, (35), 193-204.

Yokoyama, Makiko (2006). Improvements of the Administrative Services by the Evaluation in Local Government. *The Japanese Journal of Evaluation Studies*, 6(2), 59-71.

Article

Evaluation of Social Capacity for Urban Air Quality Management

Kazuma Murakami

Mitsubishi UFJ Research & Consulting

kazuma@ murc.jp

Shunji Matsuoka

Graduate School of Asia-Pacific Studies,
Waseda University
smatsu@ waseda.jp

Abstract

Since the early 1990s, Capacity Development and Capacity Development in Environment (CDE) are discussed in the filed of aid assistance. The necessity of capacity development is agreed, however, there are no concrete answers as to what capacity is and how to measure it. This article proposes a new evaluation method for capacity development to show evaluation criteria and evaluation form. First, we examined the local government's capacity for air quality management in Japanese cities. We showed four elements of capacity that contributes to air quality improvement by using factor analysis. Then we also showed an evaluation form called "Actors-Factors Matrix". The method we proposed is superior to UNDP's method in systematic classification of evaluation criteria, efficient and effective evaluation, and appropriate policy design for capacity development.

Keywords

Capacity Development, Evaluation, Indicator, Factor Analysis,
Air Quality Management

1. Background and Purpose

Since the beginning of the 1990s, the concept of capacity development in the field of development assistance has come to be used frequently when considering improving assistance quality, and the UNDP, CIDA, JICA and other assistance organizations have been advancing the debate about putting it into practice (UNDP 1998, Lavergne and Saxby 2001, JICA 2004). In the face of increasingly serious global environmental problems and considering the need to increase ODA to confront environmental destruction, which is one cause of poverty, attention has inevitably turned to the capacity to deal with environmental issues.

The report of the 1987 World Commission on Environment and Development advocated sustainable development, and Agenda 21, adopted at the 1992 UN Conference on Environment and Development, expressed the need for the capacity to cope with environmental problems[1]. In response to this, discussions have also occurred at the OECD/DAC about the concepts and definitions of capacity in environment and capacity development in environment

(CDE) (OECD/DAC 1999)[2].

However, even though the OECD/DAC (1999) and other organizations share a focus on capacity and a recognition of the importance of capacity development, no unified consensus has been reached regarding how to define capacity. Abstract, conceptual definitions are all that have been realized thus far. Regarding this, Matsuoka and Honda (2002) state," since the concept is broad and abstract, assistance organizations have taken it as an idea, but not enough consideration of the definition or concrete details has occurred." In addition, even when capacity formation elements have been expressed concretely, as in Boesen *et al.* (1998), Janicke *et al.* (1997), and UNDP/GEF (2003) and other organizations, they have not been expressed logically and systematically.

UNEP/WHO (1996) has tried to clarify internal structures related to urban air quality management capacity (Figure1). In this case, measurements, data assessments and verifications, emissions source investigations, and environmental management are set as 4 capacity formation elements related to air quality management from the point of view of experience, and the air quality management capacities of 20 cities have been assessed with scores. However, the relative importance and relevance of each capacity element have not been verified, and the degrees of their contributions to air quality improvement have not been shown. In addition, each capacity element is given the same maximum potential score of 25 points regardless of its number of component assessment items or its relative importance to air quality improvement. Considering this, doubts exist about the validity of these capacity elements. In addition to this arbitrary assessment evaluation, the clarity of the capacity concepts, the measurement methods and other issues suggest that more research is necessary.

Assessments of capacity are necessary both before and after capacity development efforts. However, effective assessment cannot be achieved if the capacity in question is not made clear and suitable assessment standards established. In order to establish assessment methods for the capacity to handle environmental problems, first, systematic delineation from a logical and empirical perspective of the capacity that is to be assessed is necessary. In this case, the capacity that should be assessed is the capacity to contribute to policy results, and the development of a capacity assessment frame that sets this as the assessment standard is necessary. This investigation process should

Figure 1 Internal Structures Related to Urban Air Quality Management Capacity
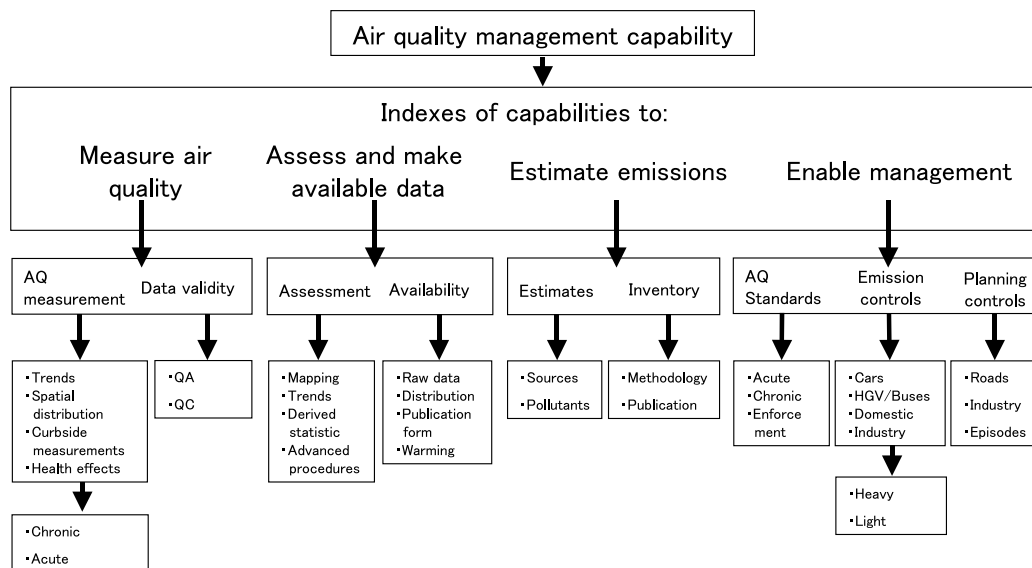
Source: UNDP/WHO (1996)

lead to the proposal of efficient, effective capacity assessment methods.

In this article, by empirically expressing capacity elements that contribute to environmental quality improvement, we seek to develop a practical capacity assessment framework that makes these into assessment measures. Specifically, by identifying capacity elements related to environmental countermeasures, and empirically showing their relationships to policy results that lead to environmental quality improvements, we will be able to suggest capacity elements that can contribute to environmental quality improvement.

In this article, Section 2 details analysis subjects, analytical methods and data, while Section 3 describes the results of the analysis. Then, in Section 4, we propose capacity assessment standards based on the clear results from section 3. Finally, in Section 5, we present a conclusion.

2. Subjects and Methods of Analysis

2.1 Setting Analysis Subjects

Our specific focus in the environmental countermeasure field is urban air pollution countermeasures, and we use UNEP/WHO (1996) efforts as a reference for comparison. Among air pollution substances, sulfur dioxide (SO_2) is a representative industrial air pollutant from factories, workplaces and other fixed emissions sources.

Our analysis period was from 1970 to 2000. According to Harashima and Morita (1995), this was the nascent period for environmental policies (1965　), and, according to Matsuoka and Kuchiki (2003), this was the period when the social capacity for environmental management began to function significantly (1970　). Environmental policies started to be implemented seriously, and their results began to become apparent during these decades. Around the start of this period, according to data from long-term continuous monitoring stations around Japan, annual average SO_2 concentration peaked in 1967, and NO_2 concentration peaked in 1971 before starting to decline.

The subjects of our analysis were the distinct capacities of governments, firms and citizens. This is in keeping with the definition of social capacity for environmental management (SCEM), which is the total capacity of society as a whole, links to the efficient and effective implementation of environmental policies and international environmental cooperation offered by Matsuoka *et al.* (2004). According to Ueta (1996), SO_2, as an industrial pollutant, is the result of external diseconomy．This means the failure of the market, and environmental policies based on government intervention, particularly command and control regulations, to have principal roles as countermeasures. Specifically, governments have a certain authority over firms that are polluters and can advance air quality improvements by promoting fuel conversion, operation management, new technology development and introduction, and other pollution reducing behavior. In addition, the so-called " Pollution Diet (Kogai- Kokkai)," the Japanese national parliamentarians in office around 1970, passed and revised 14 laws related to pollution (1970), established the Japanese Ministry of the Environment (1971) and strengthened other government environmental functions (Harashima and Morita 1995). For these reasons, we will focus on the capacity of governments as the subject of our analysist.

Our subject cities are Kitakyushu and Osaka, two ordinance-designated cities with different urban structures[3]. As Figure 2 shows, the urban structures of these two cities are not very similar, with differences in the number of manufacturing workplaces and the proportion of large-scale workplaces (as shown by the ratio of the value of manufactured products shipped from workplaces with over 300 employees). In a comparison of all the ordinance-designated cities in 1970, Kitakyushu had the fewest manufacturing workplaces, but its proportion of large-scale workplaces was highest. In contrast, Osaka had the most manufacturing workplaces, but its proportion of large-scale workplaces was lowest. In addition, Osaka had the largest government budget and the greatest number of public

Figure 2 The Numbers of Manufacturing Workplaces (left), The Proportion of Large-scale Workplaces (right)
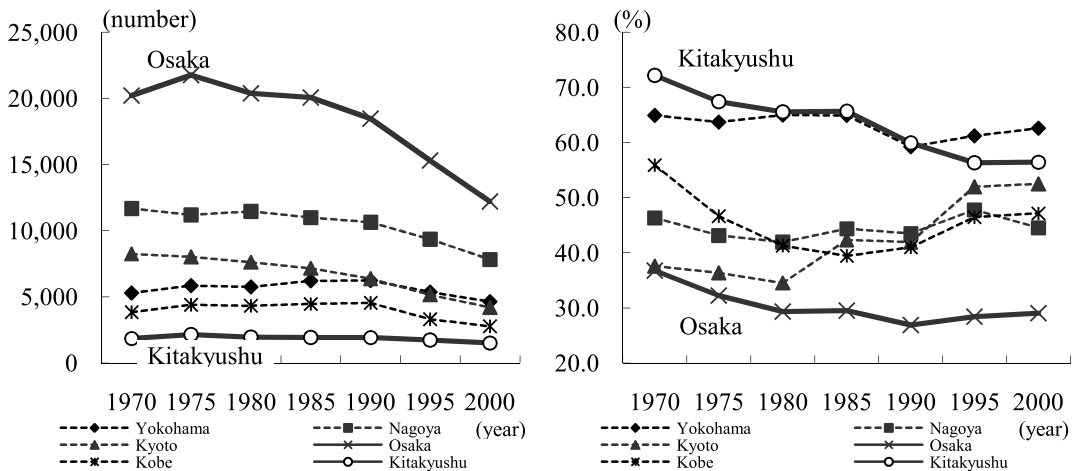
Source: Census of Manufactures / Ministry of Economy, Trade and Industry

employees, while Kitakyushu and Kyoto had the smallest budgets and staffs. Furthermore, Osaka had the highest population and population density and Kitakyushu had the lowest[4].

These differences in urban structures, especially in industrial structures, resulted in different pollution emission situations and countermeasures for the emissions from factories, workplaces and other fixed emissions sources, particularly for the main pollutant SO_2. The main subjects of regulations in Kitakyushu have been a small number of large firms with extensive technological and financial resources to apply to environmental measures, so transaction costs for communication with these firms have been small. The implementation and efficacy of policies have been achieved through information exchanges, preliminary meetings and other regular meetings (Kitakyushu City 1998, Katsuhara 2001, Fujikura 2002). In contrast, the numerous medium and small firms in Osaka have had limited technological and financial resources to apply to environmental measures, and transaction costs for communication have been large, so there has been no choice but to seek the implementation and efficacy of policies through the accumulation of government guidance efforts for individual firms (Osaka City 1994, Fujikura 2002). Given this, the extent of the transaction costs needed for promoting countermeasures by firms has been one primary factor in the formation of the SO_2 policies of both cities. This resulted in the selection of different policy measures with Kitakyushu adopting pollution control agreements and Osaka choosing administrative guidance[5].

For these reasons, we selected Kitakyushu and Osaka as representative examples of applying rational policy measures based on urban structure. We clarified their capacity formation elements related to air quality management and contributions to air quality improvement in these two cities with different urban structures and policy measures. Then, through this comparative investigation, we developed conclusions that are common to both cities and that are specific to each.

2.2 Analytical Methods

We conducted a two-stage empirical analysis in which we first identified capacity elements related to air quality management and then confirmed that these capacity elements actually contribute to air quality improvement.

2.2.1 Analysis 1: Identification of Capacity Elements

We conducted exploratory factor analysis using data related to government air pollution policies to identify capacity elements. Factor analysis is a method to identify multiple common, latent variables that are behind actual phenomena. In this case, several air pollution policies were classified by common factors. The latent commonalities of those classified air pollution policies were derived as the capacity of those policies to be effective. Thus, those factors were shown to be capacity elements for government air pollution policies. In short, factor analysis makes the latent, abstract concept of capacity visible. This analytical method suited the goals of our research because it also allows the quantitative determination of each factor's contribution to total capacity.

2.2.2 Analysis 2: Verification of Capacity Elements for Air Quality Improvement

We verified whether the individual capacity elements determined in Analysis 1, as well as the total capacity of air quality management formed by the combination of weighted capacity elements, contributed to air quality improvement. In other words, we verified the contribution of the derived capacity elements, individually and as a whole, to policy effectiveness. This also had the result of verifying the existence of the individual capacity elements and the validity of their weightings.

To verify each capacity element for air quality improvement, we set the factor scores for the capacity elements derived from the preceding factor analysis as the explanatory variables and conducted regression analysis with SO_2 concentration as the dependent variable. This clarified the relationship between each capacity element and air quality.

To verify the total combined capacity of the capacity elements, first, we used the contribution rates of the factor loads as weights to make a weighted average of the factor scores of all the capacity elements obtained through the factor analysis. Thus, we derived the total capacity for air quality management[6]. Then, we conducted regression analysis using this as the explanatory variable with SO_2 concentration as the dependent variable to clarify the total capacity for air quality management and its relationship with air quality. The contribution rates of the factor loads show the degree that each factor explains the total capacity for air quality management, so they can be taken as the weights of the capacity elements. Thus, by confirming the fitness of these relationships through regression analysis of the capacity derived from weighted capacity elements and air quality, which is the result of the policies, the validity of the capacity for air quality management, and, thus, the weightings of the capacity elements are verified.

2.3 Data

For our analysis of government subjects, in addition to city governments, we also examined the municipal environmental science laboratories involved in pollution countermeasures that provide scientific data and information for policy development and execution[7]. Specifically, we choose to investigate the Kitakyushu City Institute of Environmental Sciences and the Osaka City Institute of Public Health and Environmental Sciences.

We were able to obtain continuous quantitative data for both cities for the period from 1970 to 2000 regarding each organization's "input" (personnel, equipment and facilities, funds, information) and "output" directly related to air pollution policies (Table 1). Our subject was data that shows the degree of government activity related to air pollution policies for comparison with SO_2 concentration improvement, which is the "outcome" of those policies.

For these municipal institutes involved in pollution countermeasures, we were able to gather data on their budgets and personnel (number of people employed × average number of years employed), the number of research presentations at academic conferences, the number of articles in academic journals, and the number of specimens analyzed for air pollution by the institutes in response to requests from government agencies and residents.

For the cities themselves, we gathered data on sanitation expenses[8], the number of ambient air pollution monitoring stations, environmental division personnel (number of people employed × average number of years

Table 1 Data

Data	Source
Municipal environmental science laboratories budgets personnel the number of research presentations at academic conferences the number of articles in academic journals the number of specimens analyzed for air pollution	Report of Kitakyushu City Institute of Environmental Sciences Annual Report of Osaka City Institute of Public Health and Environmental Sciences The Survey on Wages of Local Government Employees / Ministry of Internal Affairs and Communications
The cities sanitation expenses environmental division personnel the number of ambient air pollution monitoring stations the number of onsite inspections of factories and workplaces related to air pollution the number of cases of air pollution prevention financing funds provided to firms the amount of air pollution prevention financing funds provided to firms	Environment in Kitakyushu City White Paper of Environment Pollution in Osaka City The Survey on Wages of Local Government Employees / Ministry of Internal Affairs and Communications The Annual Report of Local Finances / Ministry of Home Affairs Air Pollution in Japan / Ministry of Environmen

Source: The Authors

employed), the number of on-site inspections of factories and workplaces related to air pollution due to air pollution control laws and pollution prevention regulations, citizen complaints and related incidences, the number of cases of air pollution prevention financing funds provided to firms according to economic support plans, as well as the amount of these funds.

When considering personnel, not only the number of people, but also the quality of their work is important (Honadle 2001). For empirical analysis of human capital in economic growth, Barro (1991) and Mulligan *et al.* (1997) mainly use the number of years of education, laborer income and other representative variables as measurements of the quality of labor. Considering this, we set the average number of years employed as an indicator of experience and skill to measure the quality of personnel involved in environmental policies, and multiplied this by the number of staff to determine the overall value of personnel.

3. Analysis Results

3.1 Analysis 1: Identification of Capacity Elements

Through exploratory factor analysis (principle factor method, promax rotation) using the scree test as the standard, 4 factors were derived for both of the cities (Table 2, Table 3). For cronbach's α coefficient, Osaka's 4th factor was slightly low at 0.684, but the others were between 0.782 and 0.976, so we were able to confirm the existence of each factor with some confidence. The data structures of the factors were, with one exception, the same for both cities, making the same interpretation of factors for both cities possible.

For Kitakyushu, we interpreted the 4 factors as follows. The 1st factor was the "environmental policy resource" management capacity provided by people, goods, money and other resources that support environmental policies. The 2nd factor was the "command and control" capacity indicated by on-site inspections of pollution emissions sources and specimen testing. The 3rd factor was the "financial support" capacity to economically assist countermeasures by pollution emitters. The 4th factor was the "scientific knowledge" provision capacity for creating a basis for policy development and execution. For Osaka, on the other hand, the 1st factor was the "environmental policy resources" management capacity, the 2nd factor was the "financial support" capacity, the 3rd factor was the "scientific knowledge" provision capacity, and the 4th factor was "command and control" capacity.

If we compare the data of the two cities for "city/division environmental personnel," the data is "command and control" capacity in Kitakyushu and "environmental policy resource" management capacity in Osaka. The factor ranks differ, but otherwise they have exactly the same factor structures. The differences between "city/division environmental personnel" results can be understood by explaining the air pollution policies of the two cities. In

Table 2 Factor Load and 4 Factors(Kitakyushu)

Data	1st factor	2nd factor	3rd factor	4th factor
budgets/Municipal laboratory	0.933	-0.182	0.000	0.058
sanitation expenses/city	0.819	-0.080	0.380	0.342
personnel/Municipal laboratories	0.733	0.310	0.411	0.347
the number of ambient air pollution monitoring stations	0.692	0.172	0.502	0.408
environmental division personnel/city	0.096	0.915	-0.076	0.216
the number of onsite inspections of factories and workplaces related to air pollution/city	-0.229	0.855	0.167	0.024
the number of specimens analyzed for air pollution/Municipal laboratories	0.133	0.707	-0.033	0.450
the amount of air pollution prevention financing funds provided to firms/city	-0.198	0.073	-0.818	-0.100
the number of cases of air pollution prevention financing funds provided to firms/city	-0.571	-0.372	-0.603	-0.286
the number of research presentations at academic conferences/Municipal laboratories	0.394	0.253	0.170	0.864
the number of articles in academic journals/Municipal laboratories	0.193	0.420	0.271	0.526
eigenvalue	3.363	2.508	1.821	1.594
contribution(%)	52.0	21.4	6.9	4.2
cumulative contribution(%)	52.0	73.4	80.2	84.4

Source: The Authors

Table 3 Factor Load and 4 Factors(Osaka)

Data	1st factor	2nd factor	3rd factor	4th factor
the number of ambient air pollution monitoring stations	0.971	0.189	-0.034	0.002
environmental division personnel/city	0.832	0.443	0.220	0.216
personnel/Municipal laboratories	0.687	0.629	0.225	0.213
budgets/Municipal laboratory	0.665	0.604	0.257	0.317
sanitation expenses/city	0.613	0.542	0.381	0.409
the amount of air pollution prevention financing funds provided to firms/city	-0.225	-0.952	-0.046	-0.049
the number of cases of air pollution prevention financing funds provided to firms/city	-0.492	-0.827	-0.052	-0.108
the number of articles in academic journals/Municipal laboratories	0.068	-0.029	0.992	-0.091
the number of research presentations at academic conferences/Municipal laboratories	0.389	0.496	0.580	0.230
the number of onsite inspections of factories and workplaces related to air pollution/city	0.020	-0.383	-0.503	-0.568
the number of specimens analyzed for air pollution/Municipal laboratories	-0.212	-0.103	0.043	-0.489
eigenvalue	3.538	3.479	1.995	0.891
contribution(%)	42.6	30.3	10.7	4.9
cumulative contribution(%)	42.6	72.9	83.5	88.4

Source: The Authors

Kitakyushu, a main function performed by personnel is on-site inspection of a small number of firms based on pollution control agreements, while in Osaka a large number of firms are the subjects of administrative guidance. Therefore, the functions of government personnel in Osaka are wide-ranging and include the implementation of numerous explanatory meetings and technology investigations.

In addition, considering the contribution rate for each factor, the 1st factor, "environmental policy resource" management capacity, accounts for 52.0 percent for Kitakyushu, and 42.6 percent for Osaka City, showing that this accounts for roughly half of their capacities for air quality management. If we add the 2nd factor for each city, we can explain a little over 70 percent of the capacity. The 2nd factor was "command and control" capacity for Kitakyushu and "financial support" capacity for Osaka. This can be explained by their different policy measures based on their distinct urban structures. In Kitakyushu, where the subjects were a small number of large firms with significant technological and financial capabilities for environmental measures, the main activity of personnel is after-the-fact monitoring of firm countermeasures. In Osaka, where the subjects were a large number of medium and small firms without significant technological and financial capabilities for environmental measures, the main activity of city

personnel is guidance and support of the firms' measures. This shows that a large part of the capacities of both cities can be explained by capacity elements for the execution of policies suited to each city's characteristics, and by the capacity elements for management of environmental policy resources that make policy development possible.

Regardless of their distinct urban structures and their use of differing policy measures, if we consider their cumulative contribution rates, this shows empirically that government capacities for air quality management in both cities can be explained by these 4 elements. Moreover, given the large contribution rates of the 1st and 2nd factors to capacity elements, we confirmed the practicality and suitability of the air pollution policies of both cities.

3.2 Analysis 2: Verification of Capacity Elements for Air Quality Improvement

Making SO_2 concentration the dependent variable, and the factor scores of the 4 capacity elements the explanatory variables, we conducted regression analysis using OLS. All coefficients for both cities had negative signs and were statistically significant, making clear the contribution of all 4 capacity elements to SO_2 concentration decrease (Table 4). In addition, the standard partial regression coefficient shows that the " environmental policy resource" management capacity of both cities was the highest contributor. Next in order for both cities were " financial support" capacity, " command and control" capacity, and " scientific knowledge" provision capacity.

Next, for the factor scores of each of the 4 capacity elements, we set the contribution rate of their factor loads as their weights, calculated their weighted average values, and set these as their capacities for air quality management. Then, we conducted regression analysis using OLS with these as explanatory variables and SO_2 concentration as the dependent variable (Table 5). As a result, the coefficients for both cities were negative and statistically significant, confirming that capacities for air quality management contribute to air quality improvement.

In addition, we graphed the development process for capacities for air quality management from 1970 to 2000 (Figure 3, Figure 4). The development of capacities for air quality management for both cities can be interpreted as growing rapidly from the early 1970s[9]. Harashima and Morita (1995) classify the stages of Japan's environmental policy development as a progress period when environmental policies were enhanced and strengthened and their results were recognized (1965-1974), and a modification, stagnation and global environmental period when the environmental problems that are the subjects of environmental policies have changed qualitatively (1975-present). In comparison, Teranishi (1994) classifies the historical transformation of Japan's environmental policies as the formation and development period (1965-1974), the rapid change and retreat period (1975-1988), and the muddling and reshuffling period (1989-present).

Table 4 Regression Analysis of 4 Capacity Elements

	Kitakyushu	Osaka
scientific knowledge	-0.366 **	-0.113 *
	(-6.49)	(-2.51)
environmental policy resource	-0.567 **	-0.881 **
	(-10.01)	(-19.55)
command and control	-0.368 **	-0.222 **
	(-6.54)	(-4.94)
financial support	-0.501 **	-0.323 **
	(-8.85)	(-7.21)
constant	0.010 **	0.016 **
	(18.06)	(26.79)
F value	72.587 **	118.367 **
Adj.R2	0.905	0.940

** $p < 0.01$, * $p < 0.05$

Source: The Authors

Table 5 Regression Analysis of Capacities for Air Quality Management

	Kitakyushu		Osaka	
capacities for air quality management	-0.807	**	-0.935	**
	(-7.36)		(-14.19)	
constant	0.010	**	0.016	**
	(9.26)		(18.19)	
F value	54.155	**	201.247	**
Adj.R2	0.639		0.870	

** p< 0.01

Source: The Authors

Figure 3 Trends of Government Capacities for Air Quality Management (Kitakyushu)
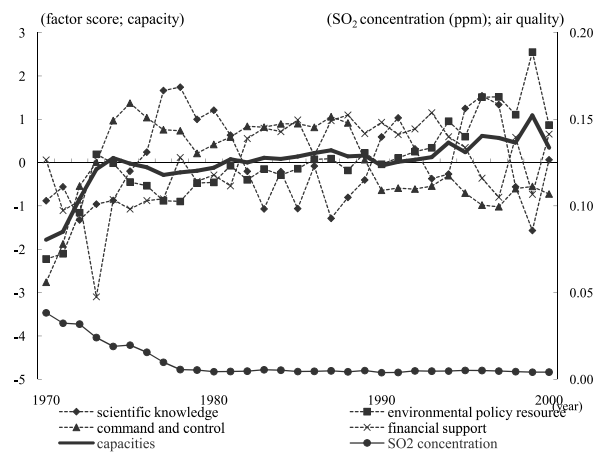
Source: Authors

Figure 4 Trends of Government Capacities for Air Quality Management (Osaka)
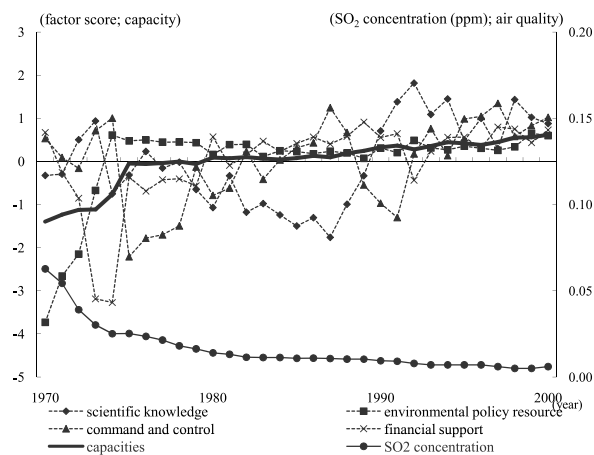
Source: Authors

The formation processes of capacities for air quality management in this analysis are consistent with these environmental policy period classifications, and they also contribute to the policy results of air quality improvement. This shows that capacity, which is responsible for the efficacy of environmental policies, also developed in the same sequence and raised policy results. By confiming the relationship between air quality, which is the result of policy, and capacity, which is responsible for the efficacy of policy, we verified the validity of capacities for air quality management. In other words, we confirmed the existence of the 4 capacity elements and their weights.

4. Setting Assessment Standards for Capacities Related to Air Pollution Countermeasures

In section 3, through exploratory factor analysis, we identified 4 capacity elements and empirically confirmed that all of them have the function of contributing to the policy result of air pollution improvement. We also confirmed that the weight of each capacity element was an appropriate reflection of its total capacity contribution. This result indicates that these 4 capacity elements are suitable as the capacity elements that should be assessed.

Given this, considering that capacity levels are responsible for policy efficacy and the extent of policy results, by showing the roles of the 4 capacity elements during policy cycles, we created assessment standards that link capacity elements with policy cycles. By presenting the relationships between capacity elements according to their roles in the policy cycles, we can show capacity assessment standards in all stages of the policy cycle. This results in a capacity assessment framework that allows a diversified perspective during assessment.

In addition, policy cycles can be viewed from the perspective of knowledge management (Yamauchi *et al*. 2000), and by expressing policy cycles as knowledge cycles, we made this a guide to setting the roles of capacity elements in policy cycles. The correspondence of capacity elements with policy cycles and knowledge cycles not only clarifies the relationships between capacity elements, it also helps embody the capacity concept and has the practical merit of making data collection and other tasks easier during capacity assessment. Of course, since capacity is not something that has only specific purposes, the policy cycle stages when capacity elements are effective cannot always be clearly divided and some overlap is possible.

4.1 Proposal of a Framework for Capacity Assessment of Government Air Pollution Countermeasures

The 4 capacity elements fill different roles in policy cycles, as shown in Figure 5. "Scientific knowledge" provision capacity fills roles of knowledge creation and accumulation in policy issue setting. "Command and control" and "financial support" capacities fill roles in putting knowledge into practice in policy execution. "Environmental policy resource" management capacity fills both a knowledge use role in policy development and a knowledge reproduction role in policy assessment. In addition, as we will describe later, "environmental policy resource" management capacity also influences the effectiveness of policy issue setting and policy execution.

"Scientific knowledge" fills the role of identifying issues that should be handled by policies with a scientific basis created through analysis of air pollution conditions, assessment and consideration of pollution elimination technology, investigative research, test inspections, and information gathering. According to Katsuhara (2001), the Kitakyushu City Institute of Environmental Sciences provides scientific data and information and makes the implementation of persuasive environmental governance possible. Fujikura (2002) states that governments are able to show the validity of scientific data when imposing pollution countermeasures on firms, and highly evaluates the early establishment of laboratories by both cities[10].

"Command and control" and "financial support" capacities also have roles in policy execution[11]. "Command and control" is the implementation of on-site inspections related to air pollution at factories and workplaces based on

Figure 5 The Correspondence of Capacity Elements with Policy and Knowledge Cycles
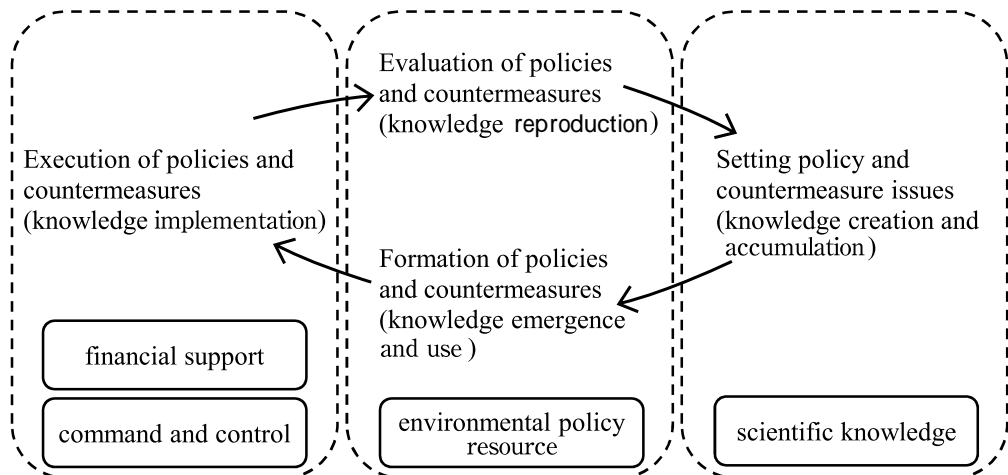
Source: The Authors

air pollution control laws, pollution prevention regulations, citizen complaints and other reasons. In addition to inspections of fuel and raw material samples, verifications of report details on equipment efficiency evaluations, maintenance and management conditions, standards compliant conditions, and inspections of the status of pollution prevention systems and the activities of pollution prevention managers and similar items are also conducted. They are also important in the capacity to promote emission source countermeasures in response to violations, and include the issuing of improvement orders, improvement recommendations, improvement instructions and other administrative guidance.

"Financial support" is the provision of economic support for the pollution countermeasures of mainly medium and small firms. This includes the contribution of funds and loan assistance for the purchase of necessary machinery, the installation of equipment and the transfer of offices and facilities with the goal of pollution prevention[12]. These systems are set in consideration of the difficulty faced by medium and small firms in meeting regulation standards and are rooted in an interdependent relationship with command and control. As one aspect of air pollution policies, they fill the role of "carrot and stick" of command and control. The capacity to implement "financial support" is not simply the assurance of a financial framework, but it also includes the implementation of explanatory meetings, examinations, management oversight, and other process capacities[13].

"Environmental policy resource" management capacity includes obtaining facilities, equipment, personnel, budget and other resources, their appropriate distribution, and their efficient use at every stage of the policy cycle. Facilities and equipment include ambient air pollution monitoring stations that provide information about both the need for environmental policies and their effectiveness. In short, they are resources for understanding the air pollution situation that are particularly important for policy development and policy assessment.

In addition, personnel are responsible for all actions related to environmental policies, and budgets are the frameworks that determine the extents of those environmental policies. Both personnel and budgets, as the means for carrying out political decision-making regarding what types of policies will be implemented and to what extent, fill important roles in policy development and policy assessment functions in particular. The reason for this is that at the stages of policy issue setting and policy execution the purposes and the extent of application have been made clear to

some degree, so the need for political decision-making can be said to be relatively low.

Furthermore, according to policy evaluation theory, we can explain the placement of policy development and policy assessment in the same types of roles. According to recent policy evaluation theory, policy development should occur after policy outcome goals are determined and policy assessment standards are set (Ueyama 1998, Mayama 2001). The reason for this is that, in policy cycles that spiral up, the results of developed policies are assessed and then they are reinvestigated. Thus, policy development and policy assessment are two aspects of the same policy process, and the capacity that supports their efficacy can be taken as the same. Therefore, "environmental policy resource" management capacity that allows the realization of political decision-making can be explained to fill both policy development and policy assessment roles.

Thus, by showing the policy cycle roles of capacity elements that we identified through empirical analysis, we are able to use these elements as capacity assessment standards. Given this, we can show the assessment framework for government air pollution countermeasure capacity as in Figure 5.

4.2 Advantage of the Actor-Factor Matrix

The government capacity element classifications shown in Figure 5 for the policy-countermeasure and knowledge cycles are, to some extent, classifications based on ordinary standards, so they are also applicable in determining the capacity formation of firms and citizens, which are also relevant social actors. Thus, we apply the government capacity assessment framework in Figure 5 to show the capacity formation related to air pollution countermeasures of firms and citizens to clarify their roles in the formation of the social capacity for environmental management (SCEM) that contributes to air quality improvement.

The Actor-Factor Matrix in Table 6 shows the core of the assessment details that we determined quantitatively with statistical data. We classified the factors according to their 3 roles as capacity elements in the policy-countermeasure and knowledge cycles. Combined with the 3 actor classifications, a 9-cell matrix is formed. In actual assessments, this 9 cells matrix can be used to understand the capacities for air pollution countermeasures based on qualitative information from interviews and questionnaires in addition to statistical data.

This Actor-Factor Matrix is a capacity assessment method that was developed after the elucidation of systemic capacity elements, the confirmation of the contribution of these capacity elements to air quality improvement, and the verification and setting of the assessment frame based on the policy-countermeasure and knowledge cycles. Compared to the capacity assessment methods developed by the UNDP (1998) and other organizations with 3 classifications individuals, organizations, social systems this method can be said to be superior for the following three points.

4.2.1 Systematic Structure of Assessment Standards

For social system capacity elements, the UNDP (1998) method gives policies, laws and regulations, management accountability, resources (people, budgets, information), and processes. Then, for organizations, it gives mission and strategy, organizational culture and structure, capacity, process, resources (people, budgets, information), and infrastructure. However, these do not clarify the necessary and sufficient conditions of the capacity elements or their relationships in internal capacity structures, and they are not systematic assessment standards. For this reason, the UNDP framework is insufficient for capacity assessment data gathering and selection. In the Actor-Factor Matrix, systemic capacity is derived through empirical analysis, and capacity elements are set in a form that corresponds to policy-countermeasure and knowledge cycles. By making guidelines from coordinated classifications of policy-countermeasure cycles, knowledge cycles and capacity elements, systematic gathering and selection of data related to capacity elements becomes easy.

Table 6 The Actor-Factor Matrix (Air quality management)

		Factor		
Process		Execution of policies and countermeasures (knowledge implementation)	Evaluation of policies and countermeasures (knowledge reproduction) / Formation of policies and countermeasures (knowledge emergence and use)	Setting policy and countermeasure issues (knowledge creation and accumulation)
Capacity		P: Policy and countermeasure execution capacities (policy & measure)	R: Environmental policy and countermeasure resource management capacities (resource management)	K: Knowledge, information and technology provision capacities (knowledge & technology)
Actor	G	○Regulatory methods • Creation and application of air pollution countermeasure legal regulations • Setting and enforcement of environmental and emission standards • Creation and application of air pollution countermeasure ordinances and fundamental plans • Observation through onsite inspection, etc. ○Market based methods • Formation and application of environmental taxes, charges and subsidy payment systems ○Autonomous methods • Formation of pollution prevention agreements (without legal basis) • Promotion of environmental study and education	○Funds, Budgets • Expansion of air pollution countermeasure budgets ○Personnel and organization • Establishment of environment-related divisions • Establishment of air pollution countermeasure organizations (committees, councils, business and citizen assemblies) • Increase of air pollution countermeasure division staff ○Facilities and equipment, etc. • Arrangement of air pollution monitoring systems • Installation of air pollution warning equipment and arrangement of information systems	○Investigation and research • Research on air pollution causes, mechanisms etc. • Development of air pollution countermeasure technology and know-how accumulation • Policy research on air pollution countermeasures ○Information disclosure and sharing • Disclosure of air quality conditions and pollution countermeasure information • Implementation of staff education and training
	F	○Regulatory methods • Observance of air pollution countermeasure laws and regulations • Observance of environmental and emission standards • Compliance with air pollution countermeasure ordinances and fundamental plans ○Market based methods • Countermeasures through the use of subsidy payment systems, etc. ○Autonomous methods • Formation of pollution prevention agreements (without legal basis) • Environmental load reduction through the entire business process including financial service production • Acquisition of ISO14001 certifications, introduction of ESCO projects, etc.	○Funds, Budgets • Expansion of air pollution countermeasure budgets ○Personnel and organization • Establishment of environment-related divisions • Increase of environment-related division staff • Increase numbers of environmental managers and pollution prevention managers ○Facilities and equipment, etc. • Arrangement of in-house monitoring systems • Arrangement of warning equipment and information systems	○Investigation and research • In-house monitoring of factories and workplaces • Development of air pollution countermeasure technology and accumulation of expertise ○Information disclosure and sharing • Creation and disclosure of environmental reports and environmental accounting • Implementation of staff education and training
	C	○Regulatory methods • Observance of air pollution countermeasure laws and regulations (field burning, etc.) ○Market based methods • Countermeasures that make use of subsidy payment systems, etc. (NGO, NPO) ○Autonomous methods • Complaints, requests and lobbying • Transition to lifestyles that conserve energy and resources, including more ecological automobile use, public transportation use, etc. • Green purchasing and environmental funds	○Funds, Budgets • Expansion of environmental countermeasure budgets ○Personnel and organization • Increase of the number of environmental countermeasure personnel (NGO, NPO) • Participation in NGO and NPO activities • Participation in environmental events, etc. ○Facilities and equipment, etc. • Acquisition of facilities and equipment related to environmental countermeasures (NGO, NPO) • Introduction of environmentally friendly products (energy conservation, new energy devices)	○Investigation and research • Investigation and research (NGO, NPO) • Observation and monitoring (NGO, NPO) ○Information disclosure and sharing • Understanding air pollution status • Implementation of environmental education and training

Source: The Authors

4.2.2 Efficient and Effective Assessment Implementation

The ambiguity and malleability of the UNDP (1998) assessment framework result in the creation of large checklists for each individual environmental problem. Furthermore, the relationships and roles in policy cycles of the capacity elements that are given to be checked are unclear, its uniqueness makes relationships with other assessment standards difficult to see, and as assessment standards they are hindered by repetitiveness. As a result, it is also impossible to confirm whether all checklist items are appropriate assessment standards for the contribution to policy results, or even if there is a need to check them all. The Actor-Factor Matrix makes assessment standards from the capacity elements that should be assessed because they have verified contributions to policy results and it gives precedence to the efforts and effects related to the assessments. The Actor-Factor Matrix is also suitable because it is a simple assessment method that developing countries can implement themselves, which is valuable for developing ownership and self-assessment capacity in those countries. Table 6 is an illustration of an air pollution countermeasure matrix that classifies capacity by the ordinary standards of policy-countermeasure cycles, knowledge cycles and capacity

elements. Since the lower assessment items are also standard item classifications, this could also be used as a shared assessment frame for other environmental fields.

4.2.3 Appropriate Policy Proposals for Capacity Development

Since the UNDP (1998) methods treat relationships among individuals, organizations and social systems as a comprehensive layered relationship structure, it is impossible to determine where resources should be applied from the results of capacity assessments for individual layers. In contrast, the Actor-Factor Matrix classifies 3 actors - governments, firms and citizens- and capacity elements (and their places in the policy and knowledge cycles) to create 9 cells with no overlap, making the division of roles clear. Thus, based on capacity assessment results, considering the possibilities of substitutions between actors, it is clear where resources should be applied in the matrix, making possible more effective policy proposals for capacity development.

However, social capacity for environmental management is dependent on the interrelationships between socioeconomic conditions and environmental quality, so in actual assessments, SCEM should not be the only focus, but rather total systems based on these interrelationships must also be considered.

5. Conclusion

In this article, by empirically expressing capacity elements that contribute to environmental quality improvement, we developed a practical capacity assessment frame that makes these into assessment measures. Specifically, by identifying capacity elements related to environmental countermeasures, and empirically showing their relationships to policy results that lead to environmental quality improvements, we suggested capacity elements that can contribute to environmental quality improvement. Then we also showed evaluation form that's called "Actors-Factors Matrix". We showed the following results.

(1) We showed four elements of capacity (i.e., scientific knowledge, command and control, financial support, environmental policy resource) that contribute to air quality improvement by using exploratory factor analysis.
(2) By confirming the relationships of the factors leading to air quality, which include policy, and capacity, capacity being responsible for the efficacy of policy, we verified the validity of capacities for air quality management. In other words, we confirmed the existence of the four capacity elements and their weights. This result indicates that these four capacity elements are suitable as the capacity elements that should be assessed.
(3) By showing the policy cycle roles of capacity elements, we are able to arrange the four elements into three factors as capacity assessment standards. Given this, we have the assessment framework for capacity development.
(4) The new evaluation method, "Actors-Factors Matrix", is superior to UNDP's method in the systematic classify of evaluation criteria, efficient and effectives evaluation, and appropriate policy design for capacity development.

Notes

1　" The ability of a country to follow sustainable development paths is determined to a large extent by the capacity of its people and its institutions as well as by its ecological and geographical conditions" (Agenda 21: Chapter 37-National mechanisms and international cooperation for capacity-building in developing countries)
2　Capacity in Environment (CDE):" the ability of individuals, groups, organizations and institutions in a given context to address environmental issues as part of a range of efforts to achieve sustainable development" (OECD 1995)

3 Ordinance designated cities have received the transfer of some authority from their prefectures, allowing them to conduct a certain amount of independent policy creation and implementation, so bias is low in decision-making. In Osaka, pollution prevention ordinances were significantly revised in 1965, and most of the regulatory authority was given to the mayor of Osaka with some of the enforcement regulations for these ordinances partially revised in 1967. In Kitakyushu City, the prefectural governor authority related to the Air Pollution Control Law was transferred in 1970. Then, the Kitakyushu City Air Pollution Control Coordination Committee was established and measures were advanced.

4 Yearbook of Metropolitan Statistics (Metropolitan statistical conference)

5 Regarding this, Fujikura (1998) states, " In [Kitakyushu], if they could convince the 30 companies and 32 factories that are members of the Kitakyushu City Air Pollution Control Coordination Committee, we could expect great effects from the measures. In contrast, the conditions in Osaka City differed and the Nishi Yodogawa Area Pollution Countermeasure Special Task Force, had no choice but to follow a grassroots approach by thoroughly canvassing the affected medium, small and tiny factories, which numbered over a thousand."

6 Nagashima and Shindo (2001) made the welfare-index with the same method.

7 Japan Society for Atmospheric Environment's Committee for Research on Historical Materials (2000) has praised the large role that regional pollution testing and research agencies have filled in promoting air pollution research. In addition, although not referring only to municipal research laboratories involved in pollution countermeasures, Japan Society of Air Pollution (1993) has said, " The engineers of local public organizations played a great role in the development of measurement technologies and countermeasure technologies, for example, in the early days of the pollution problem when business efforts were insufficient."

8 Sanitation expenses, in addition to environmental countermeasure costs, include costs related to hygiene and health. However, since no data exists that only show expenses for environmental countermeasures throughout the period from 1970-2000, we use this instead.

9 There is the opinion that capacity should be shown by stock rather than flow. Matsuoka and Honda (2002), based on Sen (1985) and the UNDP Human Development Report, designate the ability of performance bases as Capacity, and potential ability related to the freedom to select and expand human activities as Capability. Furthermore, the relationship between the two is established as Capability being the potential ability that is the foundation for Capacity, which is the execution ability and results achieved within given conditions. In this research, for abilities related to air pollution policies, we measured Capacity as the performance resulting from actual efforts, and determined ability from the performance flow by fiscal year. Of course, Capacity as shown by flow is the demonstration of the level of Capability formed as stock through development over time.

10 According to Japan Society for Atmospheric Environment's Committee for Research on Historical Materials (2000), air pollution research should only be evaluated according to whether it influences politicians and administrations and contributes to the achievement of pollution reduction.

11 As in Ueta (1996), environmental policy methods can be classified as direct methods, indirect methods and infrastructural methods, and economic support falls into the category of indirect methods. However, policy methods that use environmental taxes and other market mechanisms, which are ultimately economic incentive measures, are of a different nature.

12 Kitakyushu established the Kitakyushu City Pollution Control Fund Investment System Plan in 1968, while Osaka City established the Pollution Control Equipment Fund Investment Plan in 1967.

13 At the World Bank, investment activities are perceived as filling capacity building functions in recipient countries (from an interview). In this case, in addition to government abilities, it can be said to be policy systems that contribute to the formation of business abilities.

References

Barro, R. J. (1991). Economic growth in a gross section of countries. *Quarterly Journal of Economics*, 106, 407-443.

Boesen, J., and Lafontaine, A. (1998). *The planning and monitoring of capacity development in environment (CDE) initiatives*. CIDA.

Fujikura, R. (2002). SOx control measures of local government in Japan. In Terao, T. and Otsuka, K. (Eds.), *Dynamism of the policy process in development and the environment* (in Japanese). Chiba: Institute of Developing Economies, 37-78.

Harashima, Y., and Morita, T. (1995). A comparative analysis of environmental policy development process in East Asian countries (in Japanese). *Planning Administration*, 18 (3), 73-85.

Honadle, B. W. (2001). Theoretical and practical issues of local government capacity in an era of devolution. *Journal of Regional Analysis and Policy*, 31 (1), 77-90.

Janicke, M., and Weidner, H. (Eds.) (1997). *National environmental policies: a comparative study of capacity-building*. Berlin: Springer.

Japan Society for Atmospheric Environment's Committee for Research on Historical Materials (2000). History of air pollution of Japan (in Japanese). Tokyo: LATTICE.

Japan Society of Air Pollution (1993). *Global air pollution and measures* (in Japanese). Tokyo: Ohmsha.

JICA (2004). *Handbook of Capacity Development: improve validity and sustainability of JICA activity* (in Japanese). Tokyo: JICA.

Katsuhara, T. (2001). *Development and environmental problem in East Asia* (in Japanese). Tokyo: Keiso Shobo.

Kitakyushu City (1998). *Environmental pollution control in Kitakyushu city- analysis* (in Japanese). Kitakyushu: Kitakyushu City.

Lavergne, R., and Saxby, J. (2001). Capacity development: vision and implications. *Capacity Development Occasional Paper Series*, 3, CIDA.

Matsuoka, S., and Honda, N. (2002). Environmental cooperation and capacity development (in Japanese). *Journal of International Development Studies*, 11 (2), 149-172.

Matsuoka, S., and Kuchiki A. (Eds.) (2003). *Social capacity development for environmental management in Asia* (in Japanese). Chiba: Institute of Developing Economies.

Matsuoka, S., Okada, S., Kido, K., and Honda, N. (2004). Development of social capacity for environmental management and institutional change (in Japanese). *Journal of International Development Studies*, 13 (2), 31-50.

Mayama, T. (2001). *Essence of policy design* (in Japanese). Tokyo: Seibundoh.

Mulligan, C. B., and Sala-i-Martin, X. (1997). A labor income-based measure of the value of human capital. *Japan and the World Economy*, 9 (2), 159-191.

Nagashima, N., and Shindo, S. (2001). Making a new welfare-index (in Japanese). *ESP*, 426, 68-71.

OECD (1995). *Developing environmental capacity: a framework for donor involvement*. Paris: OECD.

OECD/DAC (1999). Donor support for institutional capacity development in environment: lessons learned. *Evaluation and Effectiveness*, 3, OECD.

Osaka city (1994). *Environmental pollution control in Osaka city* (in Japanese). Osaka: Global Environment Centre Foundation.

Sen, A. (1985). *Commodities and capabilities*. Amsterdam: Elsevier Science.

Teranishi, S. (1994). A Critical review of pollution issues and environmental policy in Japan. In Kojima, R. and Fujisaki, S. (Eds.), *development and environment*. Chiba: Institute of Developing Economies, 203-227. (in Japanese)

Ueta, K. (1996). *Environmental economics* (in Japanese). Tokyo: Iwanami Shoten.

Ueyama, S. (1998). *Reinventing Japan: a review of government performance* (in Japanese). Tokyo: NTT Publishing.

UNDP (1998). *Capacity assessment and development in a systems and strategic management context*. UNDP/BDP/Management

Development and Governance Division.

UNDP/GEF (2003). Capacity development indicators. *UNDP/GEF Resource Kit,* 4, UNDP/GEF.

UNEP/WHO (1996). *Air quality management and assessment capabilities in 20 major cities.* London: MARC.

Yamauchi, Y., Suzuki, H., and Shibukawa, S. (2000). Policy design and knowledge management (in Japanese). *GLOCOM review,* 5 (5), 1-17.

Research Note

An Evaluation Study of Program-Based Approaches (PBAs) for Development Assistance in the Primary Education Sector: A Case Study of the Education for All-Fast Track Initiative (EFA-FTI) in Honduras and Vietnam

Satoko Miwa

Japan International Cooperation Agency (JICA)
miwa.satoko2@jica.go.jp

Abstract

Program-based approaches (PBAs) are increasingly being applied to development assistance in the primary education sector. To examine their current operations and challenges, the author conducted a case study of the Education for All-Fast Track Initiatives (EFA-FTI) in Honduras and Vietnam carried out under PBAs, focusing on progress towards development results. The study shows that, despite similar approaches and tools employed under the EFA-FTI, the program and progress between the countries differ due to factors such as aid-dependency, capacity and an enabling environment. This paper argues that differentiated approaches based upon specific country contexts are essential. The importance of capacity development of partner countries is also emphasized. The author raises the issue of evaluating aid under PBAs; she proposes that evaluation be based on the concept of contribution and points out the importance of assessing strategic positioning and the role of aid, given overall efforts towards the attainment of development results.

Keywords

program, development results, aid effectiveness, contribution, capacity development

1. Introduction

1.1 Background

A program is defined as' a set of interventions marshaled to achieve specific objectives' (DAC 2003). Within the field of development assistance, there are two types of program: development programs owned by development countries and aid programs (e.g. country and thematic cooperation programs) provided by donors. The program approach advanced at present under Japan's ODA is one whereby a particular donor formulates and implements aid programs by strategically combining individual projects to attain certain cooperation objectives. A program-based

approach (PBA), on the other hand, is an approach whereby donors cooperate with one another to support a development program owned by a developing country with the aim of achieving certain development goals.

The PBA has been developed in a way to synthesize international discussions and efforts toward the achievement of development results since the latter half of the 1990s. The approach includes the following four elements: 1) leadership by developing countries, 2) assistance based on the developing country's development program and budget framework, 3) aid coordination and harmonization, and 4) utilization of the developing country's own system (country system). One example of a PBA is the sector-wide approach (SWAp) (CIDA 2003, DAC 2006).

The PBA has been increasingly applied to the development assistance. Especially in the education sector, where international collaboration has been actively carried out to attain universal primary education (UPE) set as an international development goal, the PBA has been becoming a major aid approach. This study takes up the PBA because of its significance such as this.

Then, to what degree is the PBA effective in achieving development results? Figure 1 presents the logic of the PBA toward the achievement of development results. It can be summarized that the above-mentioned four elements, by enhancing commitment of the developing country, promoting systematic support by the donors, and facilitating the developing country's more efficient management of aid and development, will lead to an adequate implementation of the agreed development program and finally the achievement of development results.

There are not several previous studies on PBAs[1]. These studies show: 1) PBAs have been implemented, applying similar measures in all countries, such as mid-term expenditure framework, budget support, institutionalization of aid coordination meetings, and joint monitoring and evaluation, 2) the approach has been contributing to promote alignment and aid coordination, and, in this sense, considered effective especially in highly aid-dependent countries, 3) transaction costs are not necessarily decreasing in the short-term due to increased work for aid coordination and so on, and 4) progress of the programs varies depending on the country, and influencing factors include the country's economic management, level of commitment, and implementation capacity.

Figure 1 Concept of the PBA
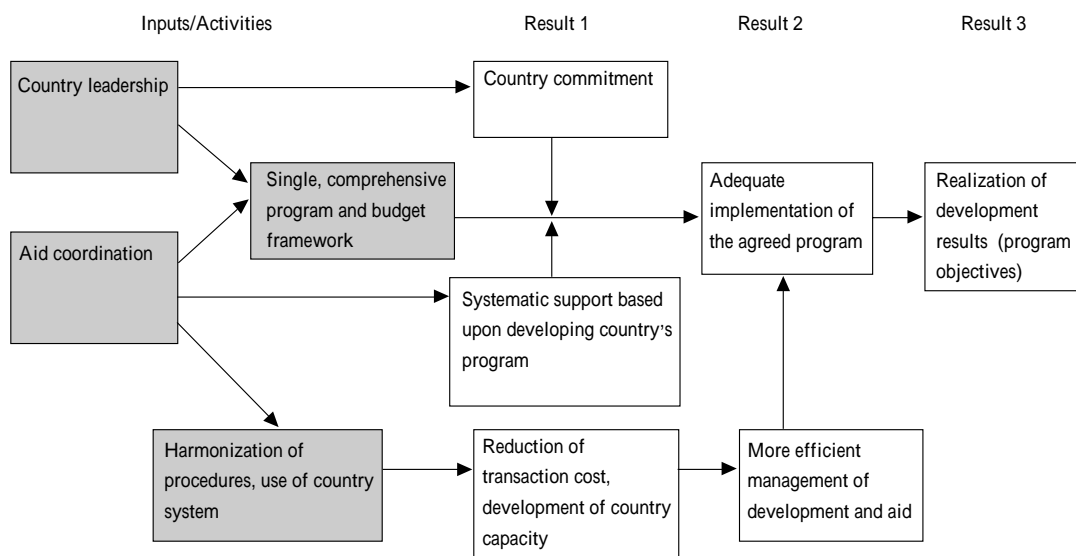
Source: Created by author based on Bartholomew et al 2006, CIDA 2003, OECD/DAC 2005, and Walford 2003

However, in the previous studies, it has not been established whether the PBA is effective for the achievement of development results. Given the fact that most of the cases are still on-going, it may be too early to verify it based upon the empirical evidence; but even the possibility is hardly examined. For the most part, the existing research only covers the progress and influencing factors up to the 'Inputs/Activities' or as far as 'Result 1' in Figure 1. The achievability of the development results is hardly mentioned. Research that attempts to analyze the design of the program under the PBA does not exist nor do the details of the contents of donor assistance to the program from the viewpoint of achievability of such an outcome.

1.2 Objectives and methodology of the research

Given the above gap in the previous research, this study intends to examine the PBA in terms of the achievement of development results and to provide suggestions for the future direction of the PBA. To do so, the study takes up the 'Education for All Fast Track Initiative' (EFA-FTI), which is implemented widely under the PBA, and evaluates the progress in 'Result 2' and 'Result 3'.

The EFA-FTI is an international initiative launched in 2002 with the objective of accelerating progress towards universal EFA. Target countries are selected, and international assistance is provided under the PBA to support the implementation of EFA-FTI programs in those countries. This paper takes up as case studies EFA-FTI programs in Honduras and Vietnam. Honduras and Vietnam are two of the starting target countries of EFA-FTI, being approved as target countries in 2002 and 2003 respectively. The paper evaluates their EFA-FTI programs based upon the progress up until the time of evaluation studies (2005 for Honduras and 2006 for Vietnam). The evaluation questions are as listed below:

Evaluation Questions

1. Is the EFA-FTI in both countries proceeding with appropriate implementation and heading towards the achievement of development results (program objectives)?

    Is the program design adequate to realize the objectives?

    Do the developing country's commitment and aid coordination under the PBA lead to systematic program interventions required for the realization of the objectives?

    Is the implementation process of the program smooth?

2. What lessons learned can be obtained from the two case studies for future PBAs?

This research is based upon information and data collected for evaluation studies on JICA programs in the primary education sector in both countries (field surveys were conducted in Honduras in May 2005 and in Vietnam in June 2006; the author participated in both surveys as the leader of the survey teams). This paper widens the scope from evaluation of the aid programs of a particular donor (JICA) to that of the development programs (EFA-FTI) of developing countries and the activities carried out by both donors and the countries themselves under the programs to assess the progress towards achievement of the EFA-FTI objectives. Sections 2 and 3 below discuss each of the two cases and respond to the first of the evaluation questions outlined above. Section 4 compares the two cases and addresses the second evaluation question.

2. Case 1: EFA-FTI in Honduras

The EFA-FTI in Honduras is an example of the PBA in a highly aid-dependent developing country. It presents many issues relating to the operation of the PBA for achieving development results, including: 1) the relationship between

the level of aid dependency and the country's autonomous program decision as well as systematic program interventions, 2) the implementation capacity of the developing country and its relationship with country leadership and the use of the country system, and 3) the implementation environment, particularly including the political and social context.

2.1 Program design

In Honduras, education sector development has been prioritized since the beginning of the 1990s. As a result, the gross enrolment rate of primary education[2] reaches over 97%. However, the primary completion rate is still low at only 69%. Also, while gender disparities are low, there are large income and regional disparities. For example, the percentage of school-age children not attending school in rural areas was 10% compared with 4% in urban areas. The average school years also differs with an average of 6.7 years in urban areas and only 3 years in rural areas. Further, as for indigenous populations, 36% of school-age children are not attending school.

Factors for these include : 1) the low level of pre-basic education (39%), 2) a high repetition rate (only 32% of pupils who enter first grade graduate from 6th grade without repetition), 3) a high dropout rate (3%), 4) low number of schooldays (110 actual schooldays per year, whereas official school calendar establishes a minimum of 200 schooldays), 5) the low quality of teachers, 6) a large portion of multi-grade class (44% of schools are with a single teacher), 7) problems of access to schools in rural areas, 8) regional economic disparities and low education demands from poor families   (Honduran Ministry of Education 2002).

In order to address these issues, EFA-FTI sets general goals by the year 2015 to: 1)   graduate 100% of the students of all ages from 6th grade, 2) graduate 85% of the students aged 12 from 6th grade in 6 years without repeating, and 3) increase academic achievement in mathematics and Spanish. The EFA-FTI program is comprised of five components. Each component's respective objectives and contents are as shown in Table 1 below.

Let us take a look at the characteristics of the program. First of all, the program does not set the increase of enrolment rate (an indicator for the access improvement) as an explicit objective at the primary level though referred to at the pre-school level. Also, the focus of the program is placed on the improvement of education service at the school level; and activities to address outside-of-school factors such as economic gaps and lack of education demand are hardly covered. Furthermore, activities to address both quality and access are included under the improvement of education service, but an expansion of school facilities is not incorporated. Taking all these into account, it is said that the Honduran EFA-FTI program is a more narrowed-down program focused mainly on the improvement of quality in education service at the school level, rather than a comprehensive program widely tackling the issues influencing the achievement of the UPE.

When the EFA-FTI was formulated in Honduras, there existed four development strategies related to primary education: the PRSP (2001), the Government Development Plan, the Education Strategy, and the Education Action Plan (all 2002). The EFA-FTI was based on these. However, in the other strategies, increase of the enrolment rate was set as a major target. Those strategies also incorporated the issues not included in the EFA-FTI such as improvement of school facilities and strengthening of education management capacity. When compared with these other strategies, a narrow scope of the EFA-FTI is clear.

This is turning out to be a limitation of the EFA-FTI. For example, with the start of the EFA-FTI, it has become evident that education management capacity is a bottleneck in implementing the program and that a more comprehensive approach, including the strengthening of the said capacity is necessary. As such, to complement the EFA-FTI, the formulation of a sector program with a wider scope is being discussed[3].

The EFA-FTI and the sector program are plans based on international agreements, while Government Development Plan, Education Strategy, and Education Action Plan are domestic plans not drawn from international

Table 1 Honduras EFA-FTI Program Summary and Government/Donor Interventions

Component	Indicator by 2015	Major content	Donors	Assistance content
1. Efficiency of basic education To guaratntee admissionof the 6 year-old population into 1st grade and to achieve its graduation from the 6th grade in 6 years	Graduate 85% of the students aged 12 from 6th grade in 6 years without repeating Graduate 100% of students of all age from 6th grade Increase academic achievement average to 70% in math and Spanish for 6th grade students	Admission to 1st grade	World Bank	Alternative education for rural/poor regions (Comunitaria)
		Academic standards and teaching materials	Government Government JICA USA USA USA Canada Sweden UNICEF	New curriculum development Participatory math learning by radio Development of math teaching materials Development of learning standard, development of common exam (MIDEH) Alternative education by radio (EDUCATODOS) Improvement of math ability by radio (APREMAT) Printing of math teaching materials Printing of math and Spanish teaching materials Literacy improvement in lower grades
		Efficient promotion		
		Levelling for pupils over-aged		
		Monitoring and assessment of internal efficiency		
		Rescuring dropout students	Government	Radio education for school-leavers
2. Teacher quality and efficiency To improve the quality and efficiency of teacher training and performance)	Achieve effective compliance of the 200 schooldays calender as well as the average of 1000 hours of school Graduate 3000 teachers at a university level university Create at least 1500 pedagogical support centers Count with an institutionalized system for school and teacher human resources management	Support to the teahcer pre-in service training	Germany Japan World Bank	Financial support for pre-service training PRODES) Construction of teacher colleges Supply of equipments to teacher college, and planning support (Communitaria)
		Admission into the teaching practice		
		In-service teacher training	Government/World Bank JICA Spain Germany Germany	Training in teaching methods In-service training for math teachers (PROMETAM) Support to wider use of math teaching materials (Luis Landa) Materials dev./training for Spanish/Science(FEBLI) Pres-service training (PRODES)
		Teacher performance and incentive		
		School and teaching human resources mangement		
		Monitoring and evaluation	USA USA	Superintendent training (Salvemos) Learning standards/development of common exam (MIDEH)
3.Strengthening of the pre-basic education (Universality of the pres-basic education for boys and girls 5 year-old	A coverage of 100% of the 5 year-old population	Coverage expansion	World Bank Government UNICEF	Alternative education for rural/poor regions (Comunitaria) Institutionalization of nonformal pre-basic education Literacy improvement and support for community participation (Esquela Amiga)
		Didactic resources	USA Sweden Government/World Bank	Development of teaching materials for radio learning Procurement of teaching materials Procurement of teaching equipments
		Training of teachers, volunteers, and tutors	Government JICA	Volunteer leader training Kindergarten teacher training (Model Project
		Monitoring and evaluation		
4. Equity/access to intercultural bilingual basic education (To guarantee the equity and access of the ethnic populations to the intercultural bilingual education)	Provide education services for pre-basic education to 100% of the 5 year-old population Graduate 85% of the 12 year-old population from 6th grade Graduate 100% of the multiple-aged population from 6th grade Increase the academic achievement of pupils in the 3rd and 6th grades in math and Spanish to 70%	Database		
		Institutionalization of bintercultural bilingual education	Government World Bank UNICEF	Training of provincial technical support teams, incentives to children Supply of equipments to special education schools (Comunitaria) Support to intercultural bilingual education
		Adjustments to basic national curriculum		
		Teacher training and performance		
		Community involvement		
		Special education	Government Spain	Organization of core teacher groups Technical/financial support for special education sector
5. Rural education networks To create rural education networks in order to ensure access of the 5 to 15 year-old population to education services in the isolated rural areas)	Ensure the access to 100% of the 5 year-old population Graduate 85% of the 12 year-old population and 100% of the multiple-aged population from 6th grade Increase the academic achievement of pupils in the 3rd and the 6th grade in math and Spanish to 70% Integrate 7000 rural multi-grade schools into a maximum of 466 networks	Establishment of rural education networks	World Bank Germany	Distribution of teaching materials (Comunitaria) Technical support (two provinces)
		Network pedagogical model		
		Promotion bonuses		
		Network management and supervision		
		School lunch, vouchers	WFP	School lunches
		Network evaluation/info system		
		Other than above	Sweden	Distribution of school goods to children in poor regions (pooled fund)
		Support to capacity development of the Ministry of Education	Germany Germany Canada Germany USA World Bank	Development of education management capacity of the Ministry of Education(ASED) EFA support/educational reform support Institution building of the Ministry of Education, supply of equipments Supply of equipments to Ministry of Education Technical support of teachers assigned to provincial offices Provision of equipments to and staff salaries of EFA-related units
		Budget support (common fund	Sweden Germany Canada World Bank Spain	Pooled fund (signatory) Pooled fund (signatory) Pooled fund (signatory) Pooled fund (signatory) Pooled fund (expected to participate)

Source: Created by author based on JICA 2006

agreements. The latter is renewed with regime changes every four years. Thus, in Honduras, those plans based on international agreements rather than domestic plans play the role as the stable basis of development. This reality forms the background to the consideration of the sector program while there are domestic plans covering wider issues than the EFA-FTI.

The EFA-FTI is to be formulated based upon the agreement of stakeholders, including donors. Why has the program in Honduras end up with one with such problems as limited coverage? One reason for this concerns a problem with the EFA-FTI formulation process. Honduras was announced to be a candidate for the EFA-FTI target countries in June 2002. Then, to go through the selection process, a program proposal was submitted to the FFA Secretariat in October; it was assessed by the country-level donor representatives and revised in November before getting approved as a target country within the month. This is an extremely short period of time for the program formulation process (JICA 2006). In Honduras, the education sector donor meeting (Mesa Redonda de Cooperantes en Educacion - MERECE) has been very active. However, due to such processes as outlined above, MERECE's involvement in the formulation of the EFA-FTI was confined to the assessment of the proposal in a short period. This indicates limited consultation with donors in the program formulation process. Similarly, it is also pointed out that involvement of civil society was lacking in the process. Insufficient consultation with stakeholders in the formulation process can not be dismissed as a factor for the observed problems in the program design.

Another influencing factor can be a consideration on practicability. The program seems to focus on those activities that are more feasible in implementation and likely to produce measurable results. For instance, in an aid-dependent country like Honduras, external assistance is indispensable in actually implementating the program. As is elaborated in section 2.2 below, the focus areas of the EFA-FTI correspond with the areas where donor assistance has been actively provided and continues to be expected. Not only development issues, but also these aspects seem to have influenced on the program design.

2.2 Progress of intervention

Financial resources for implementing the EFA-FTI come from the government budget and donor assistance. Of the former, in 2005 31% of the total budget was allocated to the education sector, with 30% of that directed to primary education. These show that primary education is a priority sector for the government spending[4]. The source of budget in education sector is comprised of government budget 85%, foreign loans 3%, grants 8%, HIPC debt relief 4%, making the government budget contribution appear relatively large. However, looking at the government budget breakdown, 97% is for current expenditures (the government budget average is 69%). Most of the current expenditures are used for personnel expenditures such as teachers' salaries, which are at six times per capita GDP and considered high even internationally[5]. Due to this, the budget for development activities relies on aid. In fact, in the same year's EFA budget, 85% came from grants, 11% from foreign loans; and the government budget constitutes only 4%.

As such, the financial plan of EFA-FTI is planned, expecting increases of aid (about 86 million dollars for each three-year cycle), particularly including budget support. To facilitate such increase of aid 'General Framework Memorandum of Understanding', which was signed by all governments and donors in October 2003, indicates the acceptance of wide aid modalities, such as:1) pooled funds, 2) projects, 3) support in-kind, and 4) complementary programs (activities not included under the EFA-FTI but have a complementary role).

As for budget support, Memorandum of Understanding concerning Pooled Fund Support was signed by three donors, and the Fund was established. However, disbursement to the Fund was delayed due to the concern about the fiduciary risk at the Ministry of Education. Actual disbursement until the end of 2005 was only 4.5 million dollars from Sweden.

Given such situation, assistance to the EFA-FTI has been mainly carried out under project modality. Table 1 outlines assistance to the EFA-FTI (assistance in the primary education sector from 2002 to May 2005). There are two years left over since the start of the EFA-FTI; however, the data reveals a divide between those areas which have been receiving various supports and have many activities and those which do not. Examples of the former include components concerning quality improvement such as teaching standards, materials, and teacher training. Those of the latter include rural education network and intercultural bilingual education. Further, even within the area of quality improvement, there are not many interventions addressing teacher deployment and school management. Moreover, among items with several interventions, there are cases where coherence and complementary of these interventions are clear and those where the interventions are simply made to the same item without such coherence. All these indicate that systematic interventions are not yet fully ensured.

The above-mentioned situation is related to the fact that much of on-going assistance had been already underway when the EFA-FTI was introduced. These interventions have various start dates. While most of them began after 2002, there are interventions such as radio education by USA or the rural network support by Germany, which are continuing from the late 1990s. The introduction of new program and start of the assistance usually have a time lag. In this case also, while there are some new activities started immediately after the introduction of the EFA-FTI (e.g. assistance to teaching standards by USA), most donors are taking a more gradual approach, reviewing their aid activities with the EFA-FTI introduction and starting a new cooperation upon the completion of existing ones in such a way as to align to the EFA-FTI. The progress of such alignment may be set forth as an added value of the PBA.

Another added value can be the progress of aid coordination. In Honduras, aid coordination has been actively promoted at the country-level, with the reconstruction assistance after Hurricane Mitch. In the education sector, the MERECE has been formed since 1998 as a mechanism for aid coordination. MERECE meets every month to share information on donors' respective activities, discuss issues related to education aid including the EFA-FTI, and carry out joint work. Based upon such existing mechanisms, donor collaboration for the EFA-FTI is progressing. Examples include the cases where the math teaching materials developed under a JICA project were printed and distributed country-wide with financial cooperation by Sweden and Canada. Likewise, in-service training for math teachers under a JICA project, implemented as a model in five provinces, was expanded to the country-wide under Spain's support.

However, further interventions are still wanting even to implement the present program components/areas. It is not easy to guarantee necessary systematic interventions by piling up of each donor's project support even if active aid coordination is undertaken. It is imperative to secure a system to ensure more systematic interventions in such a way that the government interventions form a base and donor assistance complements them. To do so, strengthening of the government's capacity and further progress in the pooled fund support are key to the implementation of the EFA-FTI program from here on.

2.3 Implementation process

Several problems are also observed in the implementation process of the EFA-FTI. All 11 in-country donor interviews revealed that the weak implementation capacity of the education ministry, problems of teachers union, as well as the political environment were factors impeding the achievement of EFA-FTI objectives. The need for strengthening implementation capacity of the education ministry was pointed out in the donor assessment at the EFA-FTI formulation stage, but this has not been reflected in the program. The problem has increasingly turned out to be evident with the start of the EFA-FTI and became a major issue for discussions at the 2004 and 2005 joint evaluation by the government and donors. Against the background that strengthening of government financial and procurement capacity has become urgent with the establishment of the Pooled Fund, the concerned donors started the assistance

for a capacity assessment and formulation of organizational capacity development plan of the education ministry. Also, the formulation of a sector plan which incorporates capacity development of the education ministry is under consideration.

A basic element of the PBA is the use of the country system. However, due to the fiduciary risk accompanying the weak capacity of the education ministry, instead of using the existing country system, the Pooled Fund was to be managed at the beginning by applying a system based upon the World Bank's. This is a big request for the Honduran side, given the level of their capacity. Such discrepancy between the concept of the PBA and the actual operation is also observed, related to country leadership; for example, the above-mentioned sector program discussions originate in the demands of the donors rather than coming from the Honduran side.

The Honduran side is expending considerable amounts of time and effort in order to respond to these kinds of demands and, despite the progress in aid coordination and harmonization, transaction costs borne by the Hondurans are not necessarily being reduced.

Regarding the problem of the ministry's implementation capacity, it needs to be taken into consideration that the EFA-FTI is the first PBA to be executed in Honduras. However, it can not be downplayed that the frequent personnel and organizational reshuffles accompanying the regime changes are obstacles to stable policy implementation and human resources development and that these result in actual weak implementation capacity. Because of the weak capacity of the ministry, external consultants have been assigned to the ministry and have taken charge over the cooperation activities. Challenges presented by this situation include the salary difference between the consultants and education ministry staff, a lack in the transfer of knowledge and skills from the consultants to staff and so on[6].

The teachers unions in Honduras are quite strong and often oppose the government, holding strikes over pay and working conditions. The union problem has become a large obstacle to carry out activities such as increase in class hours, reduction of salaries, teacher performance evaluations, and reshuffling of teachers. The political environment especially includes problem of policy continuity. An example is a total change of teacher salary policies with the regime change in 2006, whereby a new government adopted a policy to increase them while the previous government's policy was to reduce them.

In relation to the political environment, in 2005, textbooks, teaching materials as well as school supplies were distributed country-wide with the cooperation of the military. These activities were acknowledged as efforts under the EFA-FTI; however, there is also some indication that they were used to drum up popularity by the regime (at that time) which had its eye on the presidential election later in the year. Particularly the funding of school supplies from the Pooled Fund is questioned by several donors as it is not a priority activity under the EFA-FTI, though it was the government itself which made the decision for such funding.

The problems with unions and politics are generally regarded as 'external factors'. In fact, it is difficult to address these problems solely within the education sector. However, without addressing the union problem for instance, reforms like increase in class hours and teacher reshuffling remain difficult. In this sense, the union problem poses a significant risk for the implementation of the EFA-FTI. Therefore, in order to achieve the EFA-FTI objectives, it is essential to address the problems with unions as well as the policy continuity, by tackling these with efforts including levels above the education miistry (Office of the President for example).

3. Case Study 2: The EFA-FTI in Vietnam[7]

As an example of a low aid-dependent country, the EFA-FTI in Vietnam provides suggestions for: 1) relationship

between aid dependency and such issues as alignment and systematic interventions, 2) country leadership and implementation capacity, and 3) continuity of policies and program implementation. Also, it is an interesting case to analyze in which kind of countries the PBA is more effective.

3.1 Program design

In Vietnam, efforts towards the UPE through the 1990s resulted in an increase in the net enrolment rate from 86% to 95% and the completion rate improving from 57% to 74% in 1990 and 1999 respectively[8]. After this, issues of the 2000s include the increase of remaining 5% in net enrolment rate and improvement of the quality, not just quantity, of education. Regarding the former of these two foci, while the gender gap is small, there are income, ethnic, and regional disparities. For example, differences in enrolment rate between the richest quintile and poorest quintile counts for 11%; those between the 'Kin' ethnic group which accounts for 90% of the population and the ethnic minority groups, and between the urban and rural areas are 12%, and 6% respectively (2002).

As for quality, the following can be listed as factors of the problem: 1) insufficient class hours (the gross number of class hours in the year are approximately two-thirds of the international standard), 2) the large number of half-day schooling due to lack of classrooms and low salary of teaching staff (only 10% of students have full-day schooling), 3) inadequate school buildings and teaching equipment, and 4) quality of teachers and teaching methods. On the other hand, in Vietnam, the number of teachers is not necessarily lacking due to increasing numbers of teachers and a decreasing school-age population in recent years. Indeed in eight regions, the number of students per teacher is 27-33, which is about the level of the international standard (30 students per teacher). However, as for the quality of teachers, the percentage of teachers with national standard teaching qualifications was no more than 70% as of 1998[9].

It was against this background that the 'Socio-economic Development Strategy (SEDS) 2001-2010' and the ' Education Development Strategy (EDS) 2001-2010' were formulated in 2000 with the first five-year plan (2001-2005) of each. These strategies clearly stated as their goals the UPE (net enrolment rate of 99%), together with universalization of lower secondary education (increasing the net enrolment rate from 74% to 90%). To achieve such goals, the following strategies are presented: 1) strengthening of education for children from remote regions, ethnic minority, and low-income groups, 2) improvement of education quality (curriculum, quality of teachers, teaching methods and so on), and 4) reinforcement of education management. The UPE, improvement of education quality, and reduction of disparities among different ethnic groups are also incorporated into the 'Comprehensive Poverty Reduction and Growth Strategy (CPRGS)' formulated in 2002.

The goals and targets of the EFA-FTI are rooted in the above-mentioned plans. It also reflects the introduction of the new curriculum in 2002 and the results of additional situational analysis conducted in the formulation process. It is stated in the EFA-FTI document that the EFA-FTI does not replace existing plans and strategies such as SEDS or EDS. Rather, the EFA-FTI brings together in a coherent framework those plans and strategies which have been prepared at different times with different focuses. EFA-FTI is also placed as a program which gives a longer perspective until 2015 while the existing plans cover the period until 2010 (SRVN 2003). In Vietnam, the national development plans such as SED and EDS have remained a basis for development; and agreements with donors, such as CPRGS and EFA-FTI, are placed as complementary to these national plans.

Strategic goals of the EFA-FTI include: 1) moving from quantity to a focus on quality, 2) completing universal primary and lower secondary education, 3) providing life-long learning opportunities, 4) mobilizing full community participation, and 5) ensuring effective management and ever better resource utilization. The program covers early childhood care and education (ECCE), primary, lower secondary and non-formal education (NFE). For each of these, objectives are set in three domains–' access', ' quality and effectiveness' and management' – and specific targets

by the year 2015 and action plans to achieve such objectives and targets are indicated. Action plans in the primary education include: 1) increasing access, including the construction of school facilities, 2) improving the quality in both hard (facilities) and soft (curriculum, teachers, teaching materials) aspects and evaluating improvement, 3) enlarging opportunities of disadvantaged children for education with quality and rectify education gaps for disadvantaged, and 4) strengthening of education management at the central and local governments. (See Table 2)

As these present, the EFA-FTI is a quite systematic plan, covering a wide levels, objectives, and activities to address education problems in Vietnam. This is much to do with the fact that EFA-FTI was formulated in a way to integrate various existing plans. It can be also attributed to its formulation process that the EFA-FTI was prepared with participation by not only central government but also local governments, through consultation with wider donors, and with technical cooperation by such donors like UNESCO and World Bank. 2.5 years have been spent to prepare for the plan.

Because of the decentralization in Vietnam, it is local governments who play the central role in the EFA-FTI execution. Each local government is to formulate the provincial EFA plan based on the national level EFA-FTI and implement it. Of 64 provinces, 10 have developed the provincial EFA plans; and these are currently being piloted. As such, achievement of the EFA-FTI objectives will depend on how adequately provincial EFA plans will be formulated interms of responding to the specific needs of the province as well as ensureing consistency with the national EFA-FTI and, then, how appropriately such provincial plans will be implemented.

3.2 Progress of the intervention

In Vietnam, the education sector's proportion of the government budget[10] is 18% (in 2004) [11]. Of the education sector budget, 32% is for primary and 21% for lower secondary. Financial resources for the education sector budget are from the government budget (73%), and non-governmental budget (27%). Apart from aid (ODA), the non-governmental budget includes education bonds and contributions from parents, communities and companies through the 'Socialization of Education' campaign[12]. ODA accounts for no more than 1% of the education sector's entire budget. Of the budget expenditure, 75% is allocated to current expenditures whereas the remaining 25% is capital expenditures. Personnel expenses account for 71% of current expenditures; however, teachers' salaries were only 1.45 times per capita GDP in 2001, well below the international indicator (3.5 times) of the EFA-FTI[13]. In Vietnam, the level of aid dependency in the education sector budget is low. In EFA-FTI, it is also no more than 10% of the total financial estimates from 2003-2015 and only 1.6% in 2015. The Vietname side plays a central role in funding the implementation of the EFA-FTI.

The major interventions of the EFA-FTI up until June 2006 are presented in Table 2. There are already various activities underway under the action plans of each component. With regard to the interventions by the government, under the EDS, it has established an earmarked special account in the education budget and is implementing the National Target Program (NTP). The NTP consists of the following seven projects: 1) promotion of UPE and literacy, 2) revision of education programs and teaching materials, 3) training of teachers and improvement of teaching colleges, 4) education support for children in remote areas and from the ethnic minority in the mountain areas, 5) improvement of facilities in core schools, universities, and vocational training centers, 6) introduction of IT and promotion of foreign language learning in schools, and 7) strengthening capacity for vocational training. Of these, 1) through 6) are EFA-FTI related activities[14]. In addition, curriculum revision, which includes a shift to 'active teaching and learning (ATL)', was implemented in 2002 and is being advanced in order to improve the quality of education. School construction and improvement of facilities are also being conducted using government funds as well as non-government funds such as contributions from parents and communities.

Because the EFA-FTI was introduced in the middle of the five-year EDS (2001-2005) the implementation of

Table 2   Vietnam EFA-FTI Program Summary and Government/Donor Interventions

Objective	Target by 2015	Action plan	Donor	Contents of support

Access
1. To provide access to affordable and quality primary education for all children, pariticularly from ethnic minorities, disadvantaged groups and girls.
2. To ensure all children complete the full cycle of 5 grades of primary education
	School construction in remote areas at least 250 new schools built every year up to 2005 and 100 schools built every year up to 2010
	  Reduce repetition rate (2.5% by 2010, 1% by 2015), and dropout rate 2% by 2010, 1% by 2015)
	  Reintegration rate of dropouts by 2010 70% of the previous year's dropouts, up to 95 % by 2015)
		Provision of affordable school for all children in primary school age
		  Program to ensure that all children complete a full 5 year-grade primary cycle
		  Societal program for disadvantaged/excluded children (school maintenance, teacher dispatch, textbooks, bilingual education, scholarships, enlightenment of parents)
		  Provision of full primary education to out-of-school youth
			Government, World Bank, UK, EC, Canada,Belgium, New Zealand	Targeted budget support(TBS) for the National Target Program(NTP . [Related programs: mandatory primary education/literacy, support of education for mountainous areas, ethnic minorities, disadvantaged children]

			Government, World Bank, UK, Norway, Canada, Australia	Education support for disadvantaged children, improvement of school facility , teaching materials, development of education managament capacity and community capacity (PEDC)
			Government	School construction(gov't budget, community donations
			Japan	School constrution in northern mountain region
			Japan (UNESCO assoc.)	Community learning centres-northern mountain region
			USA CRS)	Inclusive education
			Save the Children	Inclusive education

Quality and relevance
3. To ensure the transition from quantitative development to quality primary education of a high level of learning achievement, starting with a fundamental school quality level in all primary schools
	All teachers receive 30 days training per year (from 2003) and meet national standards (by 2010)
	  Yearly distribution of teaching guide
	  Scholarships and IT training for teachers
	  Salary/allowances in line with socio-economic growth and government policy
	  Continuous improvement of curriculum, methods, textbooks and continuous evaluation system
	  Free textbooks(by 2006 in all disadvantaged areas, all children by 2015)
	  Increase in education expenditure(US$15 per child and US$400 per school by 2015)
	  Replace temporary classrooms by solid structure (by 2010 with priority to disaster-prone areas)
	  Full-day schooling in all schools (by 2015
	  Increase in class-contact hours (international standard 900hrs/year by 2015
		Implementation of the curriculum reform (2002-2007) (distribution of textbook/teaching materials, teacher training, IT)
		
		  Primary teacher development and training (teacher/principal training, performance standards, class support, establish quality assurance system etc)
		  Assessment of learning achievement (new exams, implementation/evaluation)
		
		  Improvement of learning environment and learning outcomes (class hours, classroom renovation, introduce facility standards, textbook distribution)
		  Continuous curriculum improvement(2008-15) (Curriculum evaluation/reform, 9-year education cycle preparation
			Government, World Bank, UK, EC, Canada, Belgium, New Zealand	TBS for NTP [Curriculum reform/teaching material overhaul core school facility improvement, improvement of teacher's college facility]

			Government	NTP [IT staff training/introduction of IT into schools]
			Gov't, World Bank UK, EC, Canada,	Establishment of teacher standard, teacher training, support for education management capacity (PTDP)
			Japan Belgium	Development of teaching model under new curriculum, training of teachers, principals, provincial officials Teacher training in northern mountain region
			UNICEF	Teaching methods, curriculum/materials development, teacher training, supply of equipment/materials
			Government	Improvement of school facilities education bonds)

Management
4 To strengthen management at central, provincial, district, and school level to improve the day-to-day functioning of primary education
5. To ensure comprehensive sector development and reform, especially the decentralization of state management, the creation of continuous 9 year basic education cycle and transition from quantity to quality
	Functioning education management information system(EMIS), government information system(GIS), and decentralization(by 2010)
	  Quality private schools operating throughout the country
		Development of national level info. Base policy formulation mechanism, modernizing policy implementation monitoring system, decentralization management, introduction of 9-year cycle, private school development
		  Capacity development for planning and decentralized management development at province, district and school levels (training support program, financial and human resources mgt system, establishment of quality management unit)
		  Mechanism and capacity development for efficient resource utilization and affordable cost-sharing
		  Mechanism and capacity development for information-based decision-making at all administrative levels
			Government, EC	EMIS development, provincial-level application, information base policy formulation

			Canada, UNESCO	Support for forumulation of the national and provincial level EFA action plan
			Japan	Support for the formulation of national and provincial level education sector development pland(PEDP)
			World Bank	Support for the formulation of mid-term expenditure framework (MTEF)
			Government	Establishment and management of private schools
Other			Norway	Support for entire NTP program 7 projects) not TBS
			World Bank, ADB, Canada, Denmark, England, EC, Japan, Ireland	Education sector share of poverty reduction support credit (PRSC) IV up to 10% of PRSC

the EFA-FTI until 2005 has been carried out in a way to align pipeline activities under the current EDS with the EFA-FTI action programs Then, on the basis of these results, the action programs as well as a follow-up of the previous activities have been added to the next five-year EDS starting from 2006 in order to ensure consistency between EFA-FTI and EDS.

Donor assistance has been aligned with these Vietnamese side's activities. A notable example of this is the targeted budget support (TBS), which aims at supporting five projects   1) to 5)   of the above-mentioned NTP. Since 2005, $128 million has been provided under the TBS. The provision of a large amount of budget support in a relatively short period after the introduction of the EFA-FTI in 2004 can be attributed to the fact that budget support modality has been already introduced and implemented in Vietnam as the Poverty Reduction Support Credit under the CPRGS. Other than the TBS, there is also much assistance being provided by various donors as projects in such areas as teacher training and formulation of teacher standards under a new curriculum, improvement of school facilities, and reinforcement of education management in mountainous areas. Projects such as the primary teacher development project (PTDP) and the project targeting disadvantaged children (PEDC) are carried out under co-financing by the World Bank and a number of donors; Vietnam is also contributing around 16-18% of the costs for these projects.

Many of the on-going projects are new ones that have been started in around 2001-2003. There are some which started with the introduction of the EDS. In Vietnam, before 2000, assistance to the primary education sector was very small except limited assistance from a few UN agencies and bilateral donors/NGOs. It is said that onset of the EFA-FTI formulation started in 2001 played a certain role in sparking the increase in assistance to the sector. The fact that the EFA-FTI was formulated immediately after the introduction of the EDS in a way to ensure consistency with the latter also made it possible to have many interventions for the EFA-FTI from its early stages.

On the other hand, there is little direct cooperation among the existing projects. For example, the World Bank, Japan and Belgium are each conducting projects in the area of teacher training. However, sharing of information on activities, experiences, and results among these projects is very limited (JICA 2007). As for aid coordination, with the introduction of the EFA-FTI, an education sector group has been established since 2004 and holding periodic meetings; but the activities there are at most exchange of information and views. Regarding the current levels of communication among donors, the problems such as insufficient information sharing and coordination of activities are pointed out (UNICEF 2004). This may be due to the fact that Vietnam is playing the central role in advancing the EFA-FTI, including aid coordination.

3.3 Implementation process

With a high level of ownership and low aid dependency, the Vietnamese side has taken the lead in implementing the EFA-FTI in Vietnam. Further, with generally good implementation capacity of the government as well as a stable implementation environment, there are no major problems observed so far that might seriously impede the implementation. However, the capacity of the local governments, which are charged with the actual execution of EFA-FTI at each province, is still weak compared to the central government; and capacity differences among regions are also large. From the perspective of achieving development objectives, the following two points are particularly important: 1) how to ensure appropriate implementation at the provincial level under the decentralized system, and 2) how to distribute the funds and provide technical support, essential for the implementation.

For example, in Vietnam, government revenues are first collected centrally before being distributed to the local governments. While the contributions from parents and communities are directly allocated to the provinces concerned, other funds are distributed, considering population, socio-economic conditions and so on. The distribution of the funds within the province as well as their management are entrusted to each local government,

(Seki *et al* 2006). Under such system, the following challenges have been pointed out: 1) financial management system differs among local governments and there is a lack of simultaneity in financial management system between central and local governments, 2) fund distribution at the provincial level is not necessarily consistent with the education development objectives and general direction at the national level, and 3) it is difficult to comprehend and put together the information on the financial situation at local, central, and sector levels. Similarly, with regard to the NTP which is implemented at the province level, while many achievements have been made, necessity of further clarification of targets and strengthening of periodical reporting are pointed out (SRVN 2005c). These indicate that, when EFA-FTI is implemented at the provincial level from here on, it is important to address such issues as a certain degree of standardization of planning/implementation methods, consistency between the national and provincial level EFA-FTI, and the establishment of progress monitoring and reporting systems.

4. Comparison and Analysis

In both Honduras and Vietnam, the governments and donors have committed and are joining efforts towards the achievement of UPE under a single EFA-FTI program. Both governments have formulated mid-term expenditure framework and been allocating the budget to the program, while donors are supporting the program through coordinated assistance. Further, in both cases, similar approaches such as harmonization (including budget support) are being taken. On the other hand, even though both programs are undertaken under the same EFA-FTI framework, program structures and implementation differ resulting in differences in the progress towards the achievement of development results.

4.1 Program formulation

Let us first examine the differences between the programs in Honduras and Vietnam. The Vietnamese EFA-FTI consists of a wide range of activities to achieve UPE and improve quality. In contrast to this, the Honduran EFA-FTI has had its focus narrowed and thus is quite a restricted program. Efforts required for realizing the objectives of the focus areas are also not completely covered by the program. Furthermore, whereas Vietnam's EFA-FTI is consistent with the existing education development strategies, it is evident that the EFA-FTI in Honduras was not well integrated with the existing strategies.

Why have such differences occurred? Firstly, if one is to look at the EFA-FTI formulation process, while in Vietnam the program was introduced after sufficient consultation with stakeholders, in the case of Honduras, the program was formulated in a short time with very minimal stakeholder participation. Secondly, in Honduras, there was a fair amount of time pressure; and in order to meet the deadline of the selection process of the FTI-FTI target countries, the proposal was submitted in haste. This means that there were also problems in the way done by the EFA-FTI Secretariat, who gave such pressure in order to start the FTI in 2002 by approving initial target countries. Thirdly, Honduras has a high level of aid dependency and faces constraints to carrying out development activities autonomously. Therefore, in designing development programs, they cannot help but consider existing donor interventions and the likelihood of receiving further assistance. These points were particularly notable in the case of the EFA-FTI, in which feasibility of achieving results is a requirement of the program. Due to this, rather than formulating the program based on ends/outcomes, importance was placed on means/inputs. As a result, compared to if the program were to be formulated based on ends/outcomes, there were more constraints on ensuring the systematic program design. This tendency could be especially significant because the volume of aid in Honduras was also limited.

In order for program efforts to contribute to the realization of development results, adequate program design is indispensable. The difference between Honduran and Vietnamese EFA-FTI programs as mentioned above indicate the following two issues to be taken into account regarding the adequate program designing.

The first one is the importance of consideration of the developing country's level of aid dependency. Under PBA, the developing country's own programs are taken as the base. However country contexts differ between those less aid-dependent countries where aid constitutes only a complementary intervention in the program and those with a high aid dependency where the contents of the program itself will be influenced one way or another by the availability of aid. In the latter case, aid and donors assume a larger responsibility for the realization of development results. In order to formulate an appropriate program, which systematically addresses developing country's needs, it is essential to ensure sufficient dialogue between the developing country and donors from the formulation stage. PBA tends to be discussed and implemented uniformly; however it is important to consider an appropriate way of partnership for each partner country, taking into consideration different country contexts.

The second point concerns the importance of careful consideration about consistency with developing country's existing strategies when applying international initiatives and formulate localized program In Honduras, consistency with existing strategies has not been sufficient, and introduction of EFA-FTI has resulted in co-existence of similar but different parallel strategies and in a confusion in policy system. Further, with the tendency that development strategies are more heavily based upon international agreements, the situation is observed that more narrowly focused EFA-FTI has superseded government strategies with wider scope.

On the other hand, in Vietnam, the consistency between the EFA-FTI and government development strategies has been secured by having the country's own policy system as the basis. Also, to respond to the introduction of the EFA-FTI in the middle of the five-year period, the content of the existing five-year plan was integrated into the EFA-FTI, and the latter's content was further reflected in the next five-year plan . One might speculate that this was only possible, given strong sense of ownership, low aid dependency, and sufficient implementation capacity of the Vietnamese government. However, even in the case of Vietnam, substantial efforts were needed to ensure such consistency since the EFA-FTI exceeds not only the five-year plan but also the higher level ten-year plan.

These indicate that, when new programs are started under PBA based on international agreements, compatibility with the developing country's existing development strategies, consistency with the policy system, and timing of program formulation/introduction all need to be sufficiently taken into consideration. Particularly when localizing international initiatives, the difference in negotiation power which results from aid dependency needs to be paid attention; donor discretion is required to prevent pushiness and policy duplication.

4.2 Program implementation

Let us now take a look at the progress and implementation of the program interventions. Two cases show that such factors as systematic interventions, the implementation capacity of developing countries, a stable implementation environment, and stakeholder relationship influence program implementation and in turn influence the probability of achieving development results.

In Vietnam, donor assistance has been provided in such a way as to align with the country's own efforts. Because of high implementation capacity on the Vietnamese side, a large quantity of budget support has been give from the early stages of the program. This, together with the increase of new assistance, has enabled in a short period the advancement of systematic interventions, which widely cover the action plans of the program. Further, under the ' Socialization of Education' campaign, cooperative relationships have been constructed with an extensive range of stakeholders including parents, communities, and companies, and these are also contributing to the program interventions. Continuity in policy and organization has also made it possible to reflect the content and results of the

EFA-FTI in the next five-year EDS through the latter's rolling plan process.

On the other hand, in Honduras, donor assistance has constituted the major part of program interventions. The increase of new assistance has been yet limited; and alignment has been made mainly through scrap-and-build of existing assistance. Due to this, the advancement of systematic interventions has been rather slow. Further, the assistance has been provided mostly through projects being implemented directly by donors. Therefore, despite PBA, donors are the major actors of interventions and also playing the central role in coordinating them.

Due to the problem of weak implementation capacity of the Honduran side, it has taken time for budget support to actually be provided; and the amount provided is limited. In addition, because of the concern about the fiduciary risk, instead of using country system, donors' systems have been applied; and standards higher than the ability of Honduras government are being demanded. The policy changes and organizational, and personnel reshuffles accompanying regime change as well as the problem of the teachers' unions are also slowing down the implementation process.

These differences between Vietnam and Honduras show the following points as issues for consideration, regarding the implementation of the program towards the achievement of development results.

The first of these is the importance of a close examination of the implementation environment. Political, economic, and social conditions influence projects, too. However, influence at the program level is larger since the program is more directly linked to the higher, sector or macro level development policies and since its targets are more extensive. Further, because program activities span over a long period of time, an enabling environment that promotes policy continuity and stability becomes more important. The implementation environment such as policy continuity tends to be regarded as 'external factors' but, when it comes to the program, they are better considered to be 'prerequisites'. Developing country's commitment, an element of the PBA, should be examined, taking into account of such aspects as political will to ensure an enabling environment.

Similarly, there are some aspects which, even if regarded as 'external factors' at the project level, should be treated as internal elements in the case of the program. An example of this is the teachers' union problem in Honduras. Given that the teachers unions are major stakeholders in the education sector, solving the problem and building a cooperative relationship with the unions can be considered as a part of stakeholder relations. In order to ensure smooth implementation of the program towards the achievement of development results, it is important to grasp major stakeholders from a wide range of perspectives and to build cooperative relationships with them.

The second is the necessity to take into consideration the developing country's implementation capacity. Implementation capacity is presented also in previous studies as a factor influencing the progress of the PBA. However, the real problem is not that the PBA is not being advanced due to the constraints in the implementation capacity but that the way PBA is carried out without sufficiently considering implementation capacity. An example of this is observed in Honduras, where strengthening of implementation capacity has been started only after the Pooled Fund was established and a weak management capacity of the Honduran side turned out to be a constraint. Further, in the case of Honduras, the donor system is being applied to address fiduciary risk. However, in this kind of situation, it is considered to be more appropriate to take a step-by-step approach to first strengthen implementation capacity of the Honduran before setting up the Fund. PBA tends to be carried out uniformly, applying the same measures to all the countries. However, the PBA is just the means and not the ends. It is necessary to take into account each country's specific country context, to examine what is necessary for supporting country-led efforts, and to devise appropriate entry points and routes for carrying out the program under the PBA.

4.3 Evaluation of the contribution of aid

The sections above looked at how the PBA should be elaborated from the perspective of achievement of development

results. How should we, then, go about evaluating the contribution of donor assistance, especially that of individual donors, to such achievement?

The international development community has been promoting the PBA, including SWAp, based upon the recognition that achievement of high level development results (such as improvement in enrolment rate or completion rate of specific countries) requires a comprehensive approach and joint efforts by both developing countries and donors. Along with this, a concept of "contribution" rather than "attribution" has come to be applied more and more in evaluating the effects of aid. While the latter attempts to define a strict causal relationship between a specific intervention of a donor and the results, the latter examines a plausible association between them, attaching importance to the contribution of such intervention to the overall efforts towards achievement of a development results. There is not yet an established methodology for the evaluation based upon the concept of "contribution"; but it generally involves the examination of strategic positioning of a specific intervention within the entire efforts, cooperation with other interventions, and the degree of achievement of planned outcomes of the evaluated intervention (Miwa 2005). This type of evaluation has been already introduced in many donors, including JICA (JICA 2006, 2007).

Based upon such an evaluation approach, this study expanded the target of the evaluation from a particular donor's assistance to the developing country's program as well as to all the interventions into the program and tried to evaluate the probability of achieving development results. The study results indicate the importance of a comprehensive approach to solve education problems and attain the EFA goals. For example, even to achieve improvement of education quality, various issues need to be tackled, including teacher training, curriculum, teaching materials, school facilities, education management, and monitoring and evaluation of education achievement. These present the necessity of examining the relationship between a specific donors' intervention and the development results from the concept of "contribution".

Furthermore, by taking a broader view and examining the entire program of developing countries not only the assistance of specific donors as well as by comparing case studies from different countries, the study revealed several points of importance as below.

The first point involves which type of intervention contributes more to the realization of development results. Donor assistance is normally provided, given their experience and comparative advantages. This in itself provides a certain legitimacy; and as far as the program design is appropriate, interventions in the program components can be assumed to be contributing to a certain extent to the realization of development results. However, if one considers the results of the program as a whole, the impact of interventions are not equal. For example, in a case where weakness in the policy and institutional base poses constrains on the functioning of the sector including service delivery, support for the strengthening of the policy and institutional base will have a larger impact than assistance targeted at specific activities related to service delivery such as distribution of teaching materials, teacher training, and school construction. In this sense, there is an instance that technical assistance through a policy advisor, though the scale of input is small, contributes to the program as a whole, more than large-scale assistance for school construction.

The second point is that the same intervention may have a different meaning depending on the country. For example when the developing country possesses sufficient implementation capacity but faces such problems as shortage of funds or facilities, what is required is the supply of funds or construction/maintenance of school facilities rather than technical assistance.

To give another example, both in Honduras and Vietnam, JICA has been supporting teacher training under the EFA-FTI. While JICA's interventions in both countries are more or less similar, aiming at improvement of in-service teacher training and involving pilot projects in selected provinces, the meaning of the interventions differ because of their different positioning in the overall program. In Honduras, JICA has supported the development of teaching materials for and teacher training in mathematics. There, JICA is the lead donor in the area of mathematics; and the

teaching materials developed with JICA support are distributed to and being used by all teachers in the country. Therefore, the teacher training introduced by JICA assistance is in a position to go beyond pilot activities in several provinces and to be expanded country-wide. In Vietnam, on the other hand, JICA has assisted a teacher training project to pilot the 'student-centered teaching method' in view of the introduction of a new curriculum including the ATL. However, there, World Bank and other donors have also supported a teacher training project to operationalize ATL. The latter has been carried out in a larger number of provinces, combining other activities such as formulation of teacher standards. Furthermore, in the Northern mountainous regions, Belgium has been implementing a project assisting teacher training. If any one of these three initiatives is to become the core for the nationwide teacher training, it would most probably be the one supported by the World Bank, which incorporates more comprehensive activities and is implemented in wider provinces. JICA project sets as its overall goal the expansion of training activities to other provinces. However, given the above positioning of JICA's intervention in the overall program, the meaning of its cooperation is, rather than a pilot of training activities for me widespread implementation, a pilot of an innovative and complementary approach (the 'student-centered teaching method') to be incorporated into the future teacher training, which is the most likely to be introduced based upon the World Bank–supported project. Japan has supported in-service teacher training in many countries. Most of these are defined as pilots with the aim of implementing on a wide scale in the future. However, depending on the situation in each individual country and the relationship with similar assistance by other donors, the meaning and effects that should be aimed for will inevitably differ. It is necessary to clarify the positioning and meaning of pilot activities when conducting assistance.

5. Conclusion

PBA tends to be discussed uniformly, focusing on its process such as the progress of aid coordination and harmonization. However, by analyzing from the perspective of how the PBA leads to the realization of development results, this paper has clarified the importance of elaborating the approach such as steps and selection of measures, taking into consideration each country's specific context. Elements of the PBA include: 1) country leadership, 2) assistance based on the developing country's single, comprehensive program and budget framework, 3) aid coordination and harmonization, and 4) use of country system. However, in order to have these elements lead to development results, it is important to carefully examine the following three points: 1) the adequacy of program design, 2) the progress of the systematic interventions, and 3) the implementation process. Particularly in order to deepen the discussion regarding effective implementation of the PBA, the study emphasizes the importance of fully considering the degree of aid dependency, implementation capacity, and the implementation environment of developing countries.

In this relation, it is indicated in previous studies that the PBA is particularly effective in highly aid-dependent countries. In terms of improving aid delivery, this can be true since aid proliferation and duplication are serious in those highly aid-dependent countries and the PBA helps the progress of aid coordination and harmonization. However, if one examines the effectiveness of the PBA in terms of achieving development results as in this study, it cannot be asserted that the PBA is especially effective in highly aid-dependent countries. In fact, the case studies of Vietnam and Honduras indicate that country leadership and the progress towards the development results are more likely to be ensured when the developing country possesses more implementation capacity and is at the level of aid dependency that allows the country's autonomy in program interventions. As such, it remains to be a subject of further study which kind of country the PBA is more effective.

Finally, based upon the above analysis, I would like to make some suggestions to the evaluation of Japan's

ODA. Japan's ODA evaluation has a tendency to be conducted by examining consistency of aid with the developing country's strategy and assessing aid performance mainly based upon inputs and outputs of individual projects. Moreover, attention has been paid mainly to the effects of Japanese assistance rather than achievement of the development results; and the evaluation has been made from the viewpoint of "attribution" rather than "contribution". As a result, the overall picture of the efforts towards the development results, of which Japanese assistance is just one part, as well as the positioning of Japanese assistance within such efforts have not been sufficiently examined. This tendency has continued under the program approach, currently being advanced. This is because the said program approach is intended for strengthening programmatic approach of Japan's ODA to make it more effective and efficient rather than promoting the alignment with the developing country's program towards development results. However, as has already been mentioned, it is difficult to achieve development results by 'Japan's program' alone. In order to conduct more strategic and effective ODA that will lead to the development results, it is important to plan Japanese assistance and evaluate its effects, by examining the developing country's entire program, overall interventions into the program including those by the developing country and other donors, and the positioning of Japanese assistance in it.

Acknowledgement

Notes

1   Major studies containing case studies from several countries and their synthesis include the Netherlands Ministry of Foreign Affairs 2003 for the primary education field, and such publications as Brown 2001 and Foster 2000 for research on the PBA involving the education sector in general.

2   Honduras adopts the 9-3-4 system in education. Basic education covers the first nine years, which consists of three, three-year cycles. EFA-FTI targets the first two cycles (six years) and it is this period that is referred to as primary education in this paper.

3   According to a proposal presented at the 2005 donor meeting, the sector program has three main pillars: efficiency and quality of education, maintenance of infrastructure, and organizational strengthening (organizational development and decentralization, human resource development). However, despite the fact that it is called a sector program, it does not mention detailed plans for all levels (including lower and upper secondary school levels); and it is more like a complement to the EFA-FTI.

4   The following government budget data is based on the documents of the Ministry of Finance (SEFIN 2006) and the paper presented by the Ministry of Education at the time of the field survey (Presentacion del Prespuesto de la Secretaria de Educacion 2005).

5   According to the targets set under the EFA-FTI indicative framework, teachers' salaries shall be around 3.5 times per capita GDP, and that of education recurring expenditures other than salaries shall be 33%. In Honduras the former is 6 times, and the latter is 52%.

6   The EFA Directorate, which was established in the Ministry of Education in 2003 as an unit unitarily overseeing implementation of the EFA-FTI (underneath the politically appointed head), relied heavily on external consultants to carry out the major tasks. However, the integration of the EFA-FTI with the original organization and duties of the Ministry of

Education as well as the costs for many consultants became issues; and the unit was abolished in the end of 2005, based upon the results of a joint evaluation by the Ministry and donors.

7 The Vietnamese EFA-FTI targets four areas: early childhood care and education, primary school education, lower secondary education, and non-formal education; however, the paper will take primary education as the subject. The Vietnamese education system is structured 5-4-4.

8 Education statistics based on SRVN 2003, SRVN 2005b, and UNESCO 2000.

9 Previously, the national standard teaching qualification was graduation from lower secondary school and a three-year course at teacher training college (9+3). This was up-graded to graduation from upper secondary school and a two-year course at teacher training college (12+2). However, there are many teachers who do not hold this qualification and who only have the previous qualification.

10 The following budget data is based on SVRN 2005b.

11 The EFA-FTI implementation indicator is 20% by 2015. After just one year 18% has been achieved.

12 For the expansion and diversification of education financial resources, education bonds have been issued since 2003. Also the ‘Socialization of Education, through which the entire society supports education, has continued since 1990, collecting donations and funds from parents, communities, and companies for the maintenance/construction of schools and other activities to promote education.

13 With regards to teachers' salaries, there is a policy to increase the minimum salaries from the perspective of education quality improvement . Increase of teachers' salaries is also included in the EFA-FTI as an indicator; but, even after reaching this indicator in 2015 the standard does not go above 2.19 times per capita GDP (SRVN 2003).

14 3-4% of each year's education expenditure was spent for the NTP from 2001 to 2005 (SRVN 2005a).

References

Bartholomew, Ann, Leurs, Robert and McCarty, Adam (2006). *Joint Evaluation of General Budget Support 1994-2004: Vietnam Country Report,* OECD/DAC.

Brown, Adrienne, Foster, Mick, Norton, Andy and Naschold, Felix (2001). *The Status of Sector Wide Approaches,Working Paper 142,* ODI.

CIDA (2003). *CIDA Primer on Program-Based Approaches.*

CIDA (2004). *Honduras Project Approval Package: Project No. A-32111-001.*

European Commission (EC) (2003). *Guidelines for EC Support to Sector Programmes.*

Foster, Mick (2000) *New Approaches to Development Co-operation: What can we learn from experience with implementing Sector Wide Approaches? Working Paper 140,* ODI.

FTI Development Committee (2006). *Progress Report for the Education for All-First-Track Initiative.*

Honduran Ministry of Education(SE)(2002). *Fast Truck Initiative Education for All Honduras 2003-2015 FTI-EFA.*

JICA (2006). *Thematic Evaluation - Program Evaluation (Honduras Basic Education Sector).*

JICA (2007). *Thematic Evaluation - Program Evaluation (Malawi, Vietnam Education Sector)*

JICA (2002). *Effective Approach to Development Problems: Basic Education.*

Mesa Redonda de Cooperantes Externos en Educaciõn (MERECE) (2002). *Evaluatiõn de la Propuesta del Gobierno de Honduras para participar en EFA-FTI.*

Miwa, Satoko (2005). Beyond Difficulties of Country Programme Evaluation - A Proposal of Practical Methodology. Japanese *Journal of Evaluation Studies,* 5 (1) 27-44.

Netherlands Ministry of Foreign Affairs (2003). *Local Solutions to Global Challenges: Towards Effective Partnership in Basic Education.*

OECD/DAC (2003). *Glossary of Key Terms in Evaluation and Results Based Management.*

OECD/DAC (2005). *DAC Guideline and Reference Series Harmonizing Donor Practices for Effective Aid Delivery, 2.*

Secretaría de Educación, Rupublica de Honduras (SE) (2004a). *Memoria de la Pevisitun Conjunta del Plan EFA.*

Secretaría de Educación, Rupublica de Honduras (SE) (2004b). *Informe Pevisitun Conjunta del Plan EFA. 2002-2004.*

Secretaría de Educación, Rupublica de Honduras (SE) (2005a). *Memoria de la Pevisitun Conjunta del Plan EFA.*

Secretaría de Educación, Rupublica de Honduras (SE) (2005b). *Plan Estrategico Sectoral de Educación Periodo 2005-2015.*

Secretaría de Finanzas, Rupublica de Honduras (2006). Informe de la Liquidación del *Presupuesto General de Ingresos y Egresos de la Rúpublica: Ejercicio Fiscal 2005 Gobierno Central.*

Seki. Moriichi et al. (2006) *Case Study Research on Strategies for Universalization of Primary Education.*

Socialist Republic of Vietnam (SRVN) (2001). The Education Development Strategic Plan for 2001-2010.

Socialist Republic of Vietnam(SRVN) (2003). National Education for All (EFA) Action Plan 2003-2015.

Socialist Republic of Vietnam (SRVN) (2005a). *5 Year Strategic Education Development Plan 2006-2010.*

Socialist Republic of Vietnam(SRVN)(2005b). *Vietnam Managing Public Expenditure for Poverty Reduction and Growth: Public Expenditure Review and Integrated Fiduciary Assessment.*

Socialist Republic of Vietnam(SRVN)(2005c). *ODA Strategic Framework for Period 2006-2010.*

UNESCO (2000). *Education for All 2000 Assessment.*

UNESCO(2005). *Education for All: The Quality Imperative, EFA Global Monitoring Report 2005.*

UNICEF (2004). *Potential for Education Sector Wide Approaches (SWAps) in East Asia, Proceeding Document.*

Walford, Veronica (2003). *Defining and Evaluating SWAps, A Paper for the Inter-Agency Group on SWAps and Development Cooperation, Institute for Health Sector Development.*

World Bank ed. (2003). *Vietnam: Partnership for Development, An Informal report prepared for the CG Meeting for Vietnam, Hanoi, December 2003.*

World Bank ed. (2004). *Moving Towards 2010: Vietnam Partnership Report 2004, An Informal report prepared for the CG Meeting for Vietnam, Hanoi, December 2004.*

Research Note　Invited Paper

The Role of Evaluation in Planning and Learning from Short Training Programs

Pamela St. Leger

Visiting Professor Department of Learning Science
Graduate School of Education Hiroshima University
and
Centre for Program Evaluation Melbourne
Graduate School of Education The University of Melbourne Australia
pksl@unimelb.edu.au

Abstract

This paper reviews the role of short course training programs within broader organisational learning frameworks and the part that evaluation plays in informing decision making about future development.

Four dimensions of program planning and evaluation are discussed. These are: (a) the kinds of learning that can reasonably be expected from short course training programs; (b) approaches and methods of evaluation that may be used to gain information about the value of short course training interventions; (c) other aspects of training and workplace learning that might be added to short course interventions to increase participants' use of knowledge and skills gained; and given this broader concept of workplace learning, (d) evaluation approaches that might be used to provide useful information to program managers to aid decision making about future training needs and program provision.

While short training courses may be easier to organize than comprehensive and systematic training and development programs, they may have little effect on changing workplace practices. In an increasing organisational climate of accountability for use of resources for which there are competing demands, coupled with organisational goals to achieve strategic advantage over competitors, short courses need to be located within broader strategic plans to develop organisational learning capacity.

Keywords

Training and Development, Training Program Design, Training Evaluation,
Transfer of Training, Capacity Building

1. Introduction

Short seminars and courses have long been a prominent feature of training programs in a wide range of sectors that includes business, education, health and community development. Usually, program managers attach an evaluation activity of some kind to "get feedback" about the value of the session or course to the participants. This information may then be used for a range of purposes: for example, to justify spending and allocation of other resources and/or the continuation of the course, or to inform the design of subsequent programs. However, stand alone short course programs and limiting evaluation to participant feedback present a narrow view of training, learning and development.

This paper discusses: (a) the kinds of learning that can reasonably be expected from short course training programs; (b) approaches and methods of evaluation that may be used to gain information about the value of short course training interventions; (c) other aspects of training and workplace learning that might be added to short course interventions to increase participants' use of knowledge and skills gained; and given this broader concept of workplace learning, (d) the evaluation approaches that might be used to provide useful information to program managers to aid decision making about future training needs and program provision.

2. What Kinds of Learning can be Reasonably Expected from Short Course Training Programs?

Before embarking on a discussion of short course training programs, I want to make clear that within the context of current thinking and trends in training and development, such a narrow focus would be regarded as somewhat outdated. This is because knowledge of what does and what does not constitute effective training has shifted in the last decade or so from a focus on training individuals to a focus on developing organisational learning through systemic planning to achieve business and organisational goals (Rohmetra and Easter-by Smith 2004). Thus, short courses are now regarded as one intervention in the training and development process that complement, or are complemented by, other interventions such as ongoing mentoring and coaching, and management practices that promote and support organisational learning.

Therefore, due consideration needs to be given to the kinds of learning that result from various training activities, how and when they are scheduled, and the conditions that are needed to support learning. For example, Tannenbaum and Yukl's (1992) Analysis of studies of effectiveness of simulation games used in management education concluded that mechanisms such as debriefing should be built into simulation training activities. They also concluded that simulations alone are not likely to change management behaviour. Additional complementary activities such as "opportunities for coaching, feedback, reinforcement and practice of skills on the job" (p. 408), and a clearly articulated model of effective and appropriate management were necessary to produce positive training results. Organisational approaches to training and building organisational capacity are discussed in Sections 4 and 5 of this paper.

With respect to selecting training methods that are likely to produce different kinds of learning, Rothwell & Sredl (1992) offer a useful planning framework. They suggest the first step is planning an "organisational curriculum". This is "a long-term, strategic instructional plan for all formal learning events" (p. 341) which provides "long-term direction" for training to assist an organisation to:

adapt to change; and

maintain effective and efficient work practices.

Thus the "curriculum" is both a "policy" and a "tool" to guide learning. Within this framework, Rothwell and Sredl (1992) make explicit the links between particular theories of instruction and different types of organisational goals and purposes. These links are outlined below:
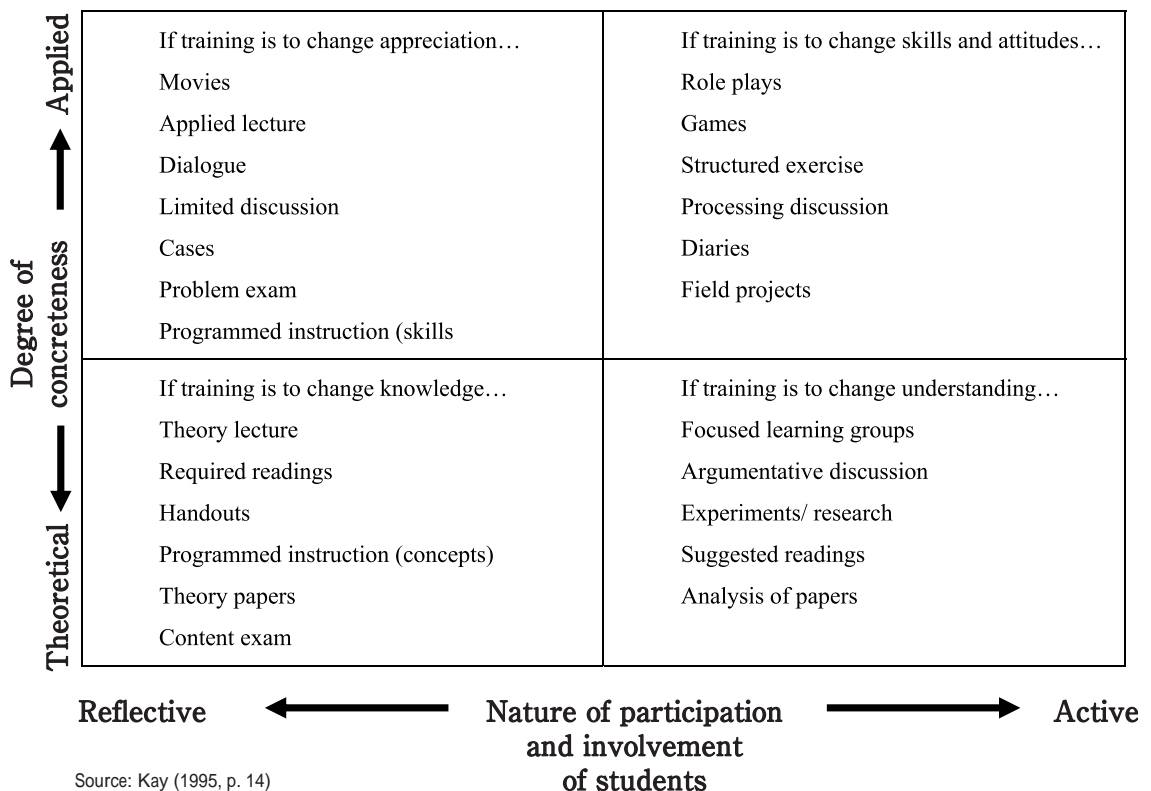
Figure 1 Theories of Instruction matched to Organisational Goals and Purposes

Purpose	Theory of Instruction
To disseminate new information	Subject centred curriculum
To develop technical skills	Objectives centred curriculum
To generate new ideas	Experience centred curriculum
To match learners' needs to relevant learning experiences?	Opportunity centred curriculum

Source: Rothwell and Sredl (1992)

Kay (1995) uses similar logic to identify particular training methods in relation to four categories of purpose. These are outlined in Figure 2 below:

Figure 2 Which Training Methods? For What Purposes?

Applied ↑ Degree of concreteness ↓ Theoretical	If training is to change appreciation… Movies Applied lecture Dialogue Limited discussion Cases Problem exam Programmed instruction (skills	If training is to change skills and attitudes… Role plays Games Structured exercise Processing discussion Diaries Field projects
	If training is to change knowledge… Theory lecture Required readings Handouts Programmed instruction (concepts) Theory papers Content exam	If training is to change understanding… Focused learning groups Argumentative discussion Experiments/ research Suggested readings Analysis of papers

Reflective ⟵   Nature of participation and involvement of students   ⟶ Active

Source: Kay (1995, p. 14)

These kinds of training logic frameworks help to make decision making about what, why and how to evaluate training activities easier for both program managers and evaluators because the clarity of purpose and plausibility of training program designs that are created from them delivering desired outcomes are likely to be clearer. Many evaluators now regard clarification of the links between purpose, design and intended outcomes as the first step in any evaluation (Owen, 2006). Without such coherence, it is difficult to attribute program effectiveness (or ineffectiveness) with confidence. Program clarification evaluation (program theory building) is discussed later in this paper.

3. Approaches and Methods of Evaluation that may be used to Gain Information about the Value of Short Course Training Interventions

Donald Kirkpatrick's (1998) Model of Evaluating Training is well known all over the world. It is a useful framework to apply to the logic of short course training activities because it helps program managers to recognise which dimensions of training outcomes can be realistically evaluated given the nature of the training activities and what it is realistic for participants to learn within those activities. Thus, they can direct scarce evaluation resources (ie. money and personnel time) to gathering useful data to inform decision making about ongoing program development rather than simply engaging in symbolic evaluation to demonstrate compliance.

Kirkpatrick (1998) identifies four levels of evaluation to measure training effectiveness. These are:

Level 1 Reaction: What participants liked and/or disliked about the training;

Level 2 Learning: Knowledge, skills and shifts in attitudes participants gained from the training;

Level 3 Behaviour: Knowledge, skills and attitudes participants apply routinely in their work; and

Level 4 Results: Impacts on the organisation, for example, reduced absenteeism, increased productivity.

The first two levels of Kirkpatrick's model have been widely used evaluate training effectiveness. The use of simple self-reporting instruments such as reaction surveys is common. These range in format from simple check box and Likert scale responses to item statements to questions that require participants to write more descriptive responses. The extent to which they are of value depends on the skill and experience of the person who designed the survey items and his or her skills in analysis and reporting findings, and the relevance of the information produced to the ultimate user/s. Kirkpatrick points out that the kind of information that they generate is of limited use in making decisions about future program provision. However, he does claim that a measure of participants' reaction to a training experience may be important to gauge participant levels of motivation to undertake a particular form of training. This may be an important consideration for e-learning training programs where participants' success depends so much on their comfort with the medium of delivery. For example, Strother (2002) found that measures of reaction provide useful evidence about levels of motivation and satisfaction with the learning format in e-learning training programs.

However, the extent to which participants liked the experience they had in a training course may imply that it was useful to them, but this level of evaluation but it does not claim to produce evidence that participants are likely to be able to apply what they learned in their work. Tannenbaum and Yukl (1992) argue that a better predictor of the worth of a training intervention may be to ask participants to focus their reactions on how useful and how applicable the training was to their work situations rather than whether or not they liked the training.

Kirkpatrick advocates the use of pre and post tests or control groups to evaluate Level 2: Learning. These kinds of measures, he claims, provide objective information as "proof" of increased learning. However, Kraiger,

McLinden and Casper (2004) caution that due consideration should be given to the extent to which training can be isolated from other factors that may affect work performance. For example, increased sales may be due in part to a sales training intervention, but they may also be due to parallel marketing activities. To counteract any devaluing of training contributions to business results, Kraiger et al. (2004) propose comparing sales results between a group which received training and a control group which did not. Evaluation designs of this kind, they argue, " can greatly enhance the evidence that changes in criteria are due to training and not alternative factors" (p. 345). Nevertheless, in employing such designs, evaluators also need to consider any potential ethical issues that might arise from withholding a training intervention to individuals who make up a control group.

However, Brinkerhoff (2003) argues that it is important to identify different contextual factors that may influence participants' take up of training. In order to do this, he advocates use of his Success Case Method which involves selecting most successful and least successful respondents from survey data, and interviewing them to develop best and worst case scenarios or stories of what worked, or did not work, and why. This evaluation approach relies on " evidence" of participants' experiences to seek understanding of a training initiative rather than " proof" that it did, or did not work.

Such an approach may also capture antecedents such as participants' existing expertise and intrinsic motivation to undertake training. For example, Green and Skinner 2006 found that their evaluation of a time management short course revealed that, in one instance, a group of senior management consultants showed " modest improvement" in comparison to a " younger, less experienced group", the difference being that the former group were already " highly skilled in most aspects of management" (p. 135).

4. Other Dimensions of Training and Workplace Learning

Kirkpatrick (1998) maintains that, for participants' learning from short course activities to be transferred into change in their routine work practice, Level 3: Behaviour, four conditions are necessary. The first two relate to motivation and knowledge and skill acquisition, that is, (1) participants must want to change their work practices and the " dose" (to use a health metaphor) and quality of the training intervention needs to be sufficient for (2) participants to know what do and how to do it. Green and Skinner (2006) have identified another layer of complexity when " behaviour" is applied to organisational or personal management concepts such as time management. Citing (Atkinson 1999), they point out that " soft skills" are based on:

> cognitive skills; thought processes which allow the individual to read situations and which can be used to understand and address issues or problems, the inherent assumption being that individual effectiveness is inextricably linked to organisational effectiveness (p. 135).

Thus, Kirkpatrick's remaining conditions for changing practice, that is, (3) a supportive work climate and (4) incentive for changing work practice need to be taken into consideration. Regardless of how good a short course training program is, or how much participants' liked it, decisions that affect the degree to which participants may change work practices may be outside the control of training personnel (Tovey and Lawlor 2004) or the participants. For example, managers need to be willing to provide opportunities for participants to put new knowledge and skills into practice immediately on their return, and possibly during their training program, along with coaching and mentoring support. Also, in some cases, incentives for undertaking training may have more to do with increases in salary rather than intrinsic motivational reasons. At the other end of the training process, lack of transfer may be

affected by other organisational factors such as increased workloads due to reduced staffing (Green and Skinner 2006).

Much of the recent literature on training and development, and evaluation of it, focuses on the transfer issue and how training policies, interventions and necessary supportive conditions and infrastructure can be built into organisational strategic planning (Caffarella 2002; Foley 2001; Haskell 1998; Kraiger et al. 2004; Salas and Cannon-Bowers 2001; Tennant 1999; Tovey and Lawlor 2004).

5. Evaluation for Decision Making about Future Training Needs and Program Provision

The shift in focus in the training and development literature from individual to organisational learning has also turned attention to human resource capacity building, accountability for time and money allocated to training and development activities, and results (Caffarella 2002; Foley 2001; Inman, Blumenfeld, and Ko 2005; Kraiger et al. 2004; Rohmetra and Easter-by Smith 2004; Salas and Cannon-Bowers 2001; Tovey and Lawlor 2004). Rohmetra and Easterby-Smith (2004), citing the banking industry, argue that, in addition to business strategies such as "repositioning" institutions through acquiring other companies, and developing more distinctive products to be more competitive, organisations also recognize the need to review operational functions such as training and development. Thus this function is now viewed through a more systematic lens.

As such, a more systematic approach to evaluation is required. Over a decade ago, Tannenbaum and Yukl (1992) concluded that the focus of research on training effectiveness required a shift from training methods to articulating the nature and purpose of training interventions, and what works in particular situations and contexts, why and for whom. Making explicit, firstly, the purpose of training and, secondly, the causal links between planned activities of the intervention and desired outcomes to determine its plausibility. This kind of evaluation approach (known as program theory building or program logic) is gaining increasing acceptance in sectors such as international development, health and community development as they strive to better understand how interventions designed to build capacity work in particular settings.

However, the literature indicates that the training sector has not adopted this kind of thinking to such an extent, for example, Strother (2002) cites a American Society of Training and Development (ASTD) study (2000) that "found that while 95 percent of surveyed organisations gauged trainees' reactions to courses (e.g. how well they liked the courses), … only three percent of respondents [in that study] make a real effort to measure the business results of [their net-based] training programs" (p. 3). Strother also noted that reviews of previous research (Perraton 2000; Saba 2000) indicated that few research and evaluation studies of distance education programs were based on sound methodological frameworks. They reported that "the researchers simply carried out their experiments in which they compared distance learning with traditional classroom content delivery and reported statistical results" (p. 8).

Such simple measures fail to take into account the teaching requirements of different learning formats such as e-learning, mentoring and coaching, and other personal and external variables that might influence participants' outcomes. Thus, there appear to be large leaps of faith in the internal coherence between the purposes, content, activities and outcomes of these kinds of programs, let alone the effects of local factors that might contribute to different levels of success (or failure) at particular sites.

In order to take these complexities into account, many evaluation theorists (Davidson 2005; Funnell 2000; Owen 2006; Patton 1996; Pawson and Tilley 1997; Rogers, Petrosino, Huebner, and Hacsi 2000) advocate utilization-focused approaches such as needs assessment (organisational and participant), program theory development (to identify the critical components of training programs), and planning and implementing evaluations

in close consultation with program managers and other key users of evaluation findings.

It is only when this critical clarificative evaluation work is done that it might be realistic to undertake an impact study of a training program or what Kirkpatrick refers to as Level 4: Results. However, it is difficult to prove that a training intervention has contributed to organisational effectiveness. What participants choose or choose not to embrace from a training experience is entangled with personal antecedents and the organisational climate that surrounds them. Also, from a practical perspective, Green and Skinner (2006) argue that:

> For the average practitioner the identification of direct causal links in terms of impact on the organisation's 'bottom line' is often perceived as problematic, complex and daunting, particularly in relation to the development of soft skills and can act as a significant barrier to evaluation beyond the reaction level (Skinner 2004).

Impact studies are based on the assumptions that there is good documentation that clearly articulates the program's objectives (not common according to Owen, 2006) and that training programs have been running long enough to show effects over time or that participants and their organizations are followed up some months later to investigate effects. The scale and therefore the expense to undertake such extensive evaluation work may not be warranted (Phillips and Phillips 2001).

6. A Framework for Evaluating Training Programs

One way to determine the kind of evaluation that may be best suited to a particular set of circumstances is from a program development viewpoint, that is, at what stage of maturity is the training program, who is likely to use the evaluation findings, and for what purpose? These fundamental questions underpin Owen's (2006) conceptualization of five forms of evaluation (proactive, clarificative, interactive, monitoring and impact). However, though working with education and training practitioners who are new to evaluation over many years, I have found a simplified adaptation of the model (see Figure 3) useful as a starting point in helping them to think evaluatively about their training programs (short course, mentoring or other learning and development activities) from the very beginning and within the context of broader organisational learning goals. Thus the evaluation framework can be used as a tool for informing decisions about both program design and tracking program effects at various stages of a program's development.

Focusing evaluation on the proactive and responsive phases of training program development may also help to identify the supportive conditions and infrastructure that are critical to achieving the desired training outcomes. Once these elements are identified, Owen (2006) argues that there is a sound conceptual basis for then considering how other forms of evaluation such as monitoring and impact evaluation might be used to measure program effectiveness. The framework therefore presents a conceptual and practical view of evaluation work (that is, the gathering and analysis of evidence) as an integral part of policy and program development and decision making, and suggests the kinds of data collection techniques that could be used to provide useful evidence (Owen, 2006).

7. Conclusion

While short training courses may be easier to organize than comprehensive and systematic training and development

Figure 3 Framework for Evaluating Training and Development

	Proactive Phase	Responsive Phase	Judgment Phase
Timing	Before program begins	During program implementation	After completion of program
Typical Questions for training and development programs, sub programs or projects	What specific goals do we want to address? What are the target group's needs? What do we already know about similar projects and programs ie. good practice? Where could we find useful information? Who might be able to provide ideas? What is the rationale for proposed project? What actions should we try? What will be our single focus?	How is the program going against anticipated milestones? What are we learning about developing capabilities of staff? What changes appear to be happening during the program? What seems to be working? What problems persist? How could we deal with these problems? Do we need to make any organisational changes? Do we need to make changes to the program plan?	To what extent has the program met intended objectives? To what extent has the program met organisational needs? What outcomes are there for the organisation? What outcomes are there for staff? What evidence do we have for these claims? What processes within the program led to these outcomes? Are there any unintended outcomes? What structures do we need to put in place to take these types of project further? What are the next steps?
Possible Data Collection Tools for training and development programs, sub programs or projects	Review of documents, books, journals, conference papers, websites, reports of similar projects. Audits of similar programs/projects. Focussed discussions with stakeholders. Visits to similar "light house" programs. Observations of similar programs in action Review of relevant organisational administrative data.	Observe training and work practice. Analyse data collected during implementation. Swap stories about new ways of working on similar programs or projects. Observe short term changes training and work practices in action. Interview appropriate stakeholders	Analyse any appropriate evidence eg (if appropriate) use of resources, direct/indirect economic and social benefits, especially marginal increases. Interview managers Tell/write stories to illustrate successes and challenges.

Source: Adapted from: Owen (2006) and (Schmuck 1998).

programs, they may have little effect on changing workplace practices. In an increasing organisational climate of accountability for use of resources for which there are competing demands, coupled with organisational goals to achieve strategic advantage over competitors, short courses need to be located within broader strategic plans to develop organisational learning capacity.

To counteract potential negative arguments about the relative value of training evaluation such as the cost, time required, and difficulty of attributing training initiatives to achieving business goals, training managers need to develop persuasive arguments to integrate evaluation into a framework that encapsulates all stages of program development. Used in this strategic way, evaluation resources may be used to greater effect to: review effective practice elsewhere; assess organisational and individual needs to develop training program theory; and refine program theory through action research projects generated by program staff. Once a clearly articulated and plausible program theory is developed, it should then be possible to design a monitoring system of evaluation that enables useful data to be collected to inform ongoing decisions about future training initiatives.

References

Atkinson, S. (1999). Reflection: Personal development for managers - Getting the process right. *Journal of Managerial Psychology, 4*(6), 502-510.

Brinkerhoff, R. O. (2003). *The success case method: Find out quickly what's working and what's not.* San Francisco, CA: Berrett-

Koehler.

Caffarella, R. S. (2002). *Planning Programs for Adult Learners: A practical guide for educators, trainers, and staff developers* (2nd ed.). San Francisco: Jossey-Bass.

Davidson, E. J. (2005). *Evaluation Methodology Basics: The nuts and bolts of sound evaluation.* Thousand Oaks, CA: Sage Publications.

Foley, G. (2001). *Strategic Learning: Understanding and facilitating organisational change.* Sydney: Centre for Popular Education, University of Technology.

Funnell, S. (2000). Developing and using a program theory matrix for program evaluation and performance monitoring. In P. J. Rogers, A. Petrosino, T. A. Huebner and T. A. Hacsi (Eds.), *Program Theory in Evaluation: Challenges and Opportunities,* 87, 103-112. San Francisco: Jossey-Bass.

Green, P., and Skinner, D. (2006). Does Time Management Work? An Evaluation. *International Journal of Training and Development,* 9(2), 124-139.

Haskell, R. E. (1998). *Reengineering Corporate Training: Intellectual capital and transfer of learning.* Westport, Conn.: Quorum.

Inman, B. R., Blumenfeld, D. E., and Ko, A. (2005). Cross-training hospital nurses to reduce staffing costs. *Health Care Management Review,* 30(2), 116-126.

Kay, D. (1995). *The Complete Small Business Training Guide.* Canberra: RIAL Publishing.

Kirkpatrick, D. L. (1998). *Evaluating Training Programs: The four levels* (2nd ed.). San Francisco, Calif.: Berrett-Koehler Publishers.

Kraiger, K., McLinden, D., and Casper, W. (2004). Collaborative planning for training impact. *Human Resource Management,* 43(4), 337-351.

Owen, J. M. (2006). *Program Evaluation: Forms and approaches* (3rd ed.). St Leonards, N.S.W.: Allen and Unwin.

Patton, M. Q. (1996). *Utilization-focused Evaluation* (3rd ed.). Thousand Oaks CA: Sage.

Pawson, R., and Tilley, N. (1997). *Realistic Evaluation.* Thousand Oaks CA: Sage.

Perraton, H. (2000). Rethinking the research agenda. *International Review of Research in Open and Distance Learning,* 1(1).

Phillips, P. P., and Phillips, J. (2001). Symposium on the Evaluation of Training. *Journal of Training and Development,* 5(4), 240-247.

Rogers, P. J., Petrosino, A., Huebner, T. A., and Hacsi, T. A. (2000). Program theory evaluation: Practice, promise, and problems. In P. J. Rogers, A. Petrosino, T. A. Huebner and T. A. Hacsi (Eds.), *Program Theory in Evaluation: Challenges and Opportunities* (Fall ed., 87, 5-13). San Francisco: Jossey-Bass.

Rohmetra, N., and Easter-by Smith, M. (2004). Training and development in the changing technological environment: The case of Barclays UK. *Journal of Management Research,* 4(2), 68-76.

Rothwell, W. J., and Sredl, H. J. (1992). *The ASTD Reference Guide to Professional Human Resource Development Roles and Competencies,* 1. (1st ed.). Amherst, Mass.: HRD Press.

Saba, F. (2000). Research in distance education: A status report. *International Review of Research in Open and Distance Learning,* 1(1).

Salas, E., and Cannon-Bowers, J. (2001). The science of training: A decade of progress. *Annual Review of Psychology,* 52, 471-499.

Schmuck, R. A. (1998). *Practical Action Research for Change.* Cheltenham, Vic.: Hawker-Brownlow Education.

Skinner, D. (2004). Primary and secondary barriers to the evaluation of change: Evidence from two public sector organizations. *Evaluation,* 10(2), 135-154.

Strother, J. (2002). An Assessment of the effectiveness of e-learning in corporate training programs. *The International Review of Research in Open and Distance Learning,* 3(1).

Tannenbaum, S. I., and Yukl, G. (1992). Training and development in work organizations. *Annual Review of Psychology,* 43,

607-632.

Tennant, M. (1999). Is learning transferable? In D. J. Boud and J. Garrick (Eds.), *Understanding Learning at Work*. New York: Routledge, 165-179.

Tovey, M. D., and Lawlor, D. R. (2004). *Training in Australia: Design, delivery, evaluation, management* (2nd ed.). Sydney: Pearson Education.

Publication Policy of the Japanese Journal of Evaluation Studies

Last revised on 15ᵗʰ February 2005

(The Purpose and the Name)
1 The Japan Evaluation Society (hereinafter referred to as "evaluation society") publishes "The Japan Journal of Evaluation Studies (hereinafter referred to as "evaluation study") in order to widely release evaluation studies and outputs of practical activities to domestic and international academic societies, interested individual and institutions, and contribute to the advancement and prevalence of evaluation practice.

(Editorial Board)
2 The editorial board administrates editing of evaluation study based on the editorial policy stated below.
3 The editorial board is formed with less than 20 members of the evaluation society who are assigned by the board of directors. Terms of editors are two years but can be extended.
4 The editorial board assigns one editor-in-chief, two vice-editors-in-chief, and a certain number of standing editors among the members.
5 The editorial board may hold at least one meeting to discuss the editing policy, plans of editorial board, and others.
6 The editorial board reports activities to the board of directors as needed and receives approval. Also it is required to report the progress of the past year and an activity plan for the following year at the annual conference.
7 The editor-in-chief, the vice-editors-in-chief and the standing editors organize the standing committee and administrate editing on a regular basis.

(Editorial Policy)
8 The evaluation study, as a principle, is published twice a year.
9 The evaluation study is printed on B5 paper, and either in Japanese or English.
10 Papers published in the evaluation study are categorized as five types;
  1 Review
  2 Article
  3 Study Note
  4 Report
  5 Others
11 The qualified contributors are members of the evaluation society (hereinafter referred to as "members") and persons whose contribution is requested by the standing editors. Joint submission of members and joint submission of non-members with a member as the first author are accepted. Submission by the editors is accepted.
12 Submitted manuscripts are treated as the above categories, however, the standing editors will decide based on the application of the contributors and the following guidelines;
  1 "Review" is a paper, which provides an overview of evaluation theory or practice. The editorial board will make the decision regarding publication.
  2 "Article" is considered as a significant academic contribution to the theoretical development of evaluation or understanding of evaluation practice. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.

    3 " Study note" is a discussion equivalent to the intermediate outputs of a theoretical or empirical enquiry. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.

    4 " Report" is the study report related to a practical evaluation project or evaluation. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.

    5 " Others" includes requested papers for special editions organized by the editorial board and announcements from editorial board to members regarding publication.

13 The editorial board selects two referee readers. For the " article" the editorial board makes adoption judgments referring to the results from referee readings and comments provided by one editor assigned by the editorial board. For " review" " study note" " report" and " others" the editorial board makes adoption judgments referring to the results from referee readings.

14 When editors submit a manuscript, the editors are not allowed to attend any of the standing editors committee meetings or editorial board meetings regarding the manuscript.

15 The standing editors have alternative of approval or not-approval for adoption judgment of manuscripts submitted to any categories. However exception is permitted if the editorial board approves the publication after minor rewrite. Even if the manuscripts are considered insufficient as an " article" standing editors can decide whether the manuscripts are published as a " study note" or " report" if the authors wish to publish.

(Formulation and Release of Submission Procedure)

16 The editorial board formulates the submission procedure based on the editorial policy described above and release after approval from the board of directors.

(Distribution)

17 The evaluation study is distributed to all members for free and distributed to non-members for a charge.

(Distribution of the Printed Manuscript)

18 30 copies of the respective paper are reprinted and distributed to the authors. The authors must cover any costs incurred by author's requests for printing more than 30 copies.

(Release on the Internet)

19 The papers published in the evaluation study are released on the internet with approval from the authors.

(Copyright)

20 Copyright of papers which appear in the evaluation study is attributed to the respective authors. Editorial right is attributed to the evaluation society.

(Office)

21 The office is in charge of administrative works for editing, distribution, and accounting.

Information for Contributors
(For English Papers)

1 "The Japanese Journal of Evaluation Studies" is the publication for reviews, articles, study notes, and reports relating to evaluation.

2 "The Japanese Journal of Evaluation Studies" is primary published to provide opportunities for members of the Japan Evaluation Society (hereinafter referred to as "members") to exchange findings, and to contribute to further development of the study of evaluation both domestically and internationally. As a principle, this journal publishes the contributions submitted by the members. With the exception of requested papers, the first author must be a member.

3 Adoption judgments of the manuscript are made at the discretion of the editorial board. Comments from two referee readers who are appointed for every manuscript are referred to in the screening process (the editorial board requests referee readers without notifying the author of manuscript).

4 Payment for the manuscript is not provided.

5 Papers published in "The Japanese Journal of Evaluation Studies" are released on the Internet at homepage of this academic society.

6 Regarding submission, manuscripts must be identified as one of the following categories: 1) article, 2) review, 3) study note, 4) report, and 5) others. However, the final decision of the category is made by the editorial board.
"Article" is considered as a significant academic contribution to the theoretical development of evaluation or understanding of evaluation practice.
"Review" is a paper which provides an overview of evaluation theory or practice.
"Study note" is a discussion equivalent to the intermediate outputs of a theoretical or empirical study in the process of producing an "article".
"Report" is the study report related to a practical evaluation project or evaluation.
"Others" are manuscripts for special editions requested by the editing committee.

7 Manuscript Submission
   1 Manuscripts may be written in either Japanese or English.
   2 Correction by the author is only for the first correction.
   3 English manuscripts should be submitted only after the English has been checked by a native speaker.
   4 Submit four hard copies (A4 size) of the manuscript. Contact information including mailing address, telephone number, fax number, and e-mail address, and the category of the manuscript should be clearly stated.
     For approved manuscripts, after necessary rewriting, the author needs to submit two hard copies of the final paper as well as a text file saved on a DOS/V formatted floppy disk. Original figures, charts, and maps should be provided.
   5 Total printed pages should not exceed 14 pages. Any cost incurred by printing more than 14 pages must be covered by the author.
   6 The layout for English papers should be 30 mm of margin at left and right side, 10pt for font size, 43 lines on A4 paper (about 500 words per page). An abstract of 150 words should be attached to the

front. 14 pages are equivalent to 7,000 words but the body should not exceed 6,000 words to allow for the title, header, figure, chart, footnotes, and references. Please note that the number of pages may be more than expected depending on the number of figures included.

8   Mailing address
Office of Japan Evaluation Society at International Development Center of Japan
Hitachi Soft Tower B 22nd Floor, 4-12-6, Higashi-Shinagawa, Shinagawa, Tokyo,
140-0002, Japan
Phone: +81-3-6718-5931, Facsimile: +81-3-6718-1651
E-mail: jes.info@idcj.or.jp

Writing Manual of the Japanese Journal of Evaluation Studies
(For English Papers)

Revised on 18th September 2002

1   Text, Charts, Figures, Graphs, Diagrams, Notes, and References

(1) The paper should be written in the follow order:

First page: Title; the author's name; Affiliation; E-mail address; Abstract (150 words); Keywords (5 words)

Second page: The main text; acknowledgement; notes; references
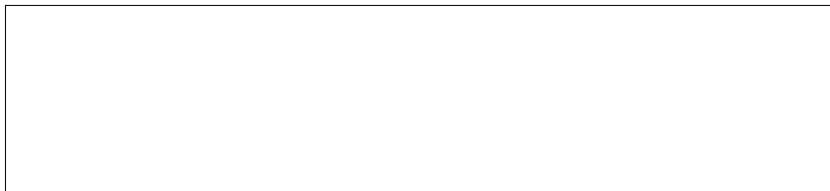
(2) Section of the text should be as follow:
1.
1.1
1.1.1
1.1.2

(3) Source of the charts, figures, graphs, and diagrams should be clarified. Submitted charts and others will be photoengraved, therefore it is important that the original chart is clear. Pictures shall be treated as figures.

Figure 1 Number of Students in the State of

Note:
Source:

Table 1 Number of Accidents in the State of

Note:
Source:

(4) Citation of literature in the text should be, (Abe 1995, p.36) or (Abe 1995).
(5) Note in the text should be, (------.[1])

(6) Note and references should be written all together in the end.

Note

1 --------.

2 --------.

(7) Reference should list the literature in alphabet order, and arranged in chronological order. Follow the examples:

Book: author (year of publication). *Title of the book*. Published location: publishing house.

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

Article from magazine: author (year of publication). Title. *Title of the magazine*, volume (number), page-page.

(e.g.) Rossi, P. H. (1999). Measuring social judgments. *American Journal of Evaluation*, 15(2), 35-37.

Article in Book: author (year of publication). Title. In editor (Eds.), *Title of the book*. Published location: publishing house, page-page.

(e.g.) DeMaio, T. J., and 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 Fransisco, Calif: Jossey-Bass, 177-196.

Book by two authors: surname, first name, and surname, first name. (year of publication). *Title of the book*. Published location: publishing house.

(e.g.) Peters, T., and Waterman, R. (1982). *In Search of Excellence: Lessons from America's Best Run Companies*. New York: Harper & Row.

Book by more than three authors: surname, first name, surname, first name, and surname, first name. (year of publication). *Title of the book*. Published location: publishing house.

(e.g.) Morley, E., Bryant, S. P., and Hatry, H. P. (2000). *Comparative Performance Measurement*. Washignton: Urban Institute.

(note 1) If some references are from the same author with the same publication year, differentiate by adding a,b,c as (1999a), (1999b).

(note 2) If the reference is more than a single line, each line from the second should be indented by three spaces.

(e.g.) DeMaio, T. J., and 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 Fransisco, Calif: Jossey-Bass, 177-196.

Referee-Reading Guideline

1. Content of the Referee-Reading Guideline

This Referee-Reading Guideline is to provide explanation of the main publication judgment, procedure of the referee-reading, to the members who submit the manuscript and for the members who are requested to conduct referee-reading in order to carry out the procedure efficiently and effectively.

2. Purpose of Referee-Reading and the Responsibility of the Author

Referee-reading is necessary for the editorial board to make decisions of whether submitted manuscripts are appropriate to publish in the Japanese Journal of Evaluation Studies or not.

If there is doubt or obscurity identified in manuscripts during the referee-reading corrections may be required. Therefore, referee-reading also contributes to the improvement of the submitted manuscripts. However, although the manuscripts are requested corrections, the author is still solely responsible in regards to the contents and it is not attributed to the referee-readers.

Referee-readers are two persons who are requested by the editorial board depending on the specialty or the field of the submitted manuscript. People who are not members of this academic society also may be requested.

3. Items of Consideration in Referee-Reading

Five points are considered in referee-reading, however, the importance of each may be different depending on the type of manuscript.

(1) Importance and utility of the theme
(2) Originality of the study
(3) Structure of the logic
(4) Validity of verification and methodology
(5) Contribution to evaluation theory and practice

- For the article, all of above five are considered.
- For the study note, especially (1), (2), (3), and (4) are considered.
- For the report, especially (1), (3), and (5) are considered.
- For the review, especially (3) and (5) are considered.

4. Attentions in submission of manuscript

Besides above five viewpoints, basic completeness as a paper is also considered, for example;
- appearance of the paper is organized
- written according to the writing manual
- described simply and distinctive
- verification data is appropriately used
- notes and references are corresponding with the text
- terminology is appropriately used
- no wording and grammatical mistakes
- no errors and omission
- no punctuation mistakes
- expression in English abstract is appropriate
- word count is according to the manual

The above mentioned forms and contents are also considered. There have been cases in which graduate students and practitioners posted without organizing the manuscripts as a paper. On those occasions, referee-reading was not conducted. Necessary consultation is strongly recommended prior to submission.

5. Judgment Cases in Referee-Reading

(1) In the case of the manuscript which is considered acceptable for the publication but is not yet complete:
The referee reader should evaluate carefully whether the paper can contribute to the development of evaluation theory or evaluation studies.

- Verification is lacking but the theory and formulation are useful for academic development.
- Analysis lacking but useful for formation and promotion of new theory.
- The literature review is not of a high standard but, the overall study is meaningful.
- Comparative study is not up to standard but is meaningful as an example of application.
- Analysis is lacking but it is meaningful as an evaluation of socially and historically important cases.
- Analysis is lacking but it is meaningful as an evaluation of particular social activities.
- Organization and expression are not up to standard as a paper but the contents are worthy to evaluate.
- Logic is not strong enough but useful in practice.
- The paper has significance as a report.

(2) In case of the manuscript which is considered as difficult for publication:
- Awareness of the issue or setting of the problem is indecisive.
- Understanding or analytical framework of notion of basic terminology is indecisive or inappropriate.
- There is a lack in credibility of data for the grounds of an argument.
- There is no clear point of an argument or appropriateness of proof.
- Organization of the paper and presentation (terminology, citation, chart, etc) are inappropriate (or not consistent).

6. Judgment

The final decision will be made on publication at the standing editors committee following one of four patterns (listed below). However, these judgments are not based on the number of errors but on the strength of the overall report. In the case of (3) and (4), there is a possibility to be published as a different type of paper. If it is published as a different type of paper, major rewrite concerning the number of words may be required.

(1) The paper will be published.
(2) The paper will be published with minor rewrite.
(3) The paper will be published with major rewrite, however as a different type of paper (review, article, study note, or report).
(4) The paper will not be published; however there is the possibility that it will be published as a different type of paper (review, article, study note, or report).

Japanese Journal of Evaluation Studies