Evaluation Principles and Methodologies Re-Visited

—— Issues for Japan ——

Kimihiro KAKU*

Introduction

In Japan, it is a rather recent phenomenon that the evaluation practice has been introduced in the various aspects of economic and social activities. Initially such practices have been done in the area of evaluating official development assistance (ODA) programs to the developing countries. Recently, however, municipal offices have also started evaluation of public administration. In addition, because of the introduction of the Project Finance Initiative (PFI) projects, the evaluation methodologies started to receive more attention than before. In this paper, the author intends to discuss some of the practical issues related to the evaluation principles and methodologies that have been applied in Japan and an attempt would be made to review the lessons that would be learned from the past experiences of the International Finance Institutions (IFIs).

On Principles of Evaluation

In the first place, it would be useful to see a distinction between audit and evaluation. Audit is normally defined as to whether the money (e.g. tax) has been expended properly according to the regulations and procedures set by the concerned authorities, whereas evaluation is the exercise to examine to what extent the money has been effectively used to achieve the intended objectives (Note : 1). It appears evident that the evaluation requires some form of judgment. In this regard additional thoughts are required and certain methodologies are needed to make the judgment objective and coherent. But in both cases, there is one common guiding principle, that is, the independence of any such exercises.

In the case of the World Bank, it is noteworthy to mention that the department

*Professor, Department of Regional Studies, Toyo University, Japan

responsible for evaluation (i.e. the Operations Evaluation Department : OED) is directly under the supervision of the Board of Directors, not under the President. This institutional arrangement implies the importance of independence of the OED. In carrying out the evaluation tasks, it is of great importance that such departments would be given the status where independent and impartial judgments would be respected. This, however, does not indicate that these departments can pursue unilateral judgements. The WB case shows that any draft evaluation reports would be sent to the operational departments and their comments would be reflected in the final reports. If any disagreements remain unresolved, the comments of the operational departments would be attached as an appendix to the final report. In the case of the European Bank for Reconstruction and Development (EBRD), the Audit Committee comprising of the Board of Directors periodically supervises the PED. It should also be noted that the evaluation department in the international finance institutions (IFIs) like WB and EBRD have an "unlimited access" to any documents including the procurement records and Board discussion records. According to the author's view, this aspect has not yet shown satisfactory progress in Japan (Note : 2).

It should also be mentioned that in addition to the institutional matters as discussed above, the quality of the evaluation staff is important. When Japanese aid-related agencies such as Japan International Cooperation Agency (JICA) and Japan International Cooperation Bank (JBIC) are compared with other international organizations, one of the weaknesses of the Japanese organizations is that because of the regular staff rotation, there is no person who would pursue the work of evaluation continuously. While it is considered that independent aspect of the evaluation work has improved in Japan over the years, the staff quality is a matter of great concern. In addition to this problem, there is almost no professional consultants in the market who specialize in evaluation. This lack of human resources in the area of evaluation poses a serious issue when we consider the future prospects of the evaluation activities in Japan (Note : 3).

Another important aspect that should be taken into account in Japan is the cultural and traditional value aspects. In the past, at least, it has not been a common practice to provide an open criticism to the works done by someone else. Open criticism has been taken as a source of social shame. Therefore, any critical criticism has been passed on to the people concerned on an informal basis. The author has conducted a seminar at one of the leading public banks (government funded) in Japan, and it was a surprise to learn that, as far as evaluation results'

classification is concerned, there were only, "generally successful" and "partly successful" cases. There were no classifications called, "unsuccessful" or "failure" (Note : 4). The author mentioned to the senior staff of this bank that in the WB and other institutions like ADB and EBRD, the use of the term "unsuccessful" or "failure" is used and it is a standard practice to do so, and it is not only unrealistic but also it would damage the integrity of evaluation works if they avoid these classifications. The answer to my enquiry was such that they were very much concerned that if "unsuccessful" or "failure" cases are known to the general public, the reputation of this bank would be affected. Also, this bank would have to face serious criticisms at the parliament. This is nothing but psychological and cultural assets inherited from the past that are still inherent in Japanese society, including public administration offices and banking institutions. In such working conditions, it would appear evident that the meaning of the term "independence" would most likely to be compromised when compared with the standards set up in Europe and USA (Note : 5). The author is of the view that there is a strong need to nurture the social culture and social systems in Japan that any criticism should be taken as a source of information for better social life. Also, criticisms should be based on "reason" and expressed "rationally" and not be based on "sentiment and emotion". Improvements in this area, while observed though slowly, would require additional training at school not only at primary and secondary education but also at university education levels.

According to the author's observation, amongst the various evaluation reports reviewed, noteworthy improvements have been noted in the Japanese Board of Audit annual reports. Traditionally this agency's reports have been focused on the "inappropriate" use of money expended by the government (including local governments) and government affiliate agencies. The inappropriate use of tax payers' money has meant that use of the money did not follow the guidelines set by the government, or used for different purposes. For the last five years or so, auditor's reports have gone beyond the traditional norm of auditing. For instance, the auditor's report in 1999 reviewed the administration of local airports managed by local municipality offices. This report has analyzed the traffic volume, particularly the number of passengers using local airports, and concluded that traffic projections made at the time of planning were too optimistic. Thus the auditor's recommendation included the need to improve the traffic projection methods. Viewed overall, the author is of the view that, supported by various data and records, the auditor'

s reports have become convincing and objective. These new trends indicate that evaluation methods have been introduced in the auditing area. The author feels that this is an encouraging movement and it is expected that such trends would be expanded to the area of evaluation of public administration which recently started at local municipality offices.

On Evaluation Methodologies

When it comes to the evaluation methodologies, several steps have been developed by the various institutions. In the WB and other international organizations, the fund contributors (mainly developed countries) expressed concern sometime ago regarding how effectively their contributed funds had been used. This was one of the driving forces making evaluation tasks as top of the agenda for these institutions (because if the confidence of the fund providers is lost, these organizations could not expect sufficient funds for their activities). Particularly in the area of social infrastructure, such as, roads, railways, telecommunications, irrigation, dam, and ports, the evaluation methodologies had been developed by using the concept of net "present value" and "internal rate of return" (IRR) based on the cash flow analysis. This is partly due to the fact that for these projects, the cost records such as construction and maintenance had been kept and made available to the evaluators, and revenue records such as toll gate incomes and traffic records of road projects had also been made available. When these data are properly recorded, the costs benefits analyses have become rather straightforward exercises.

One of the important concepts developed in appraisal/evaluation of the development projects is the "Time Value of Money" (Note : 6). In a simple term, it is important to recognize that ¥1000 we have today would not be the same ¥1000 in the one year after from now. Similarly, the value of ¥1000 in the next year would not be the same ¥1000 for today. In order to calculate the future value, we use the interest rate. If the interest rate is 10% per year, following the formula of $(1+r)^n$, (r=interest rate, n=the number of years), the interest rate factor would be 1.1, thus the value of the present ¥1000 in the next year would be ¥1100. The process of finding the present worth of a future value is called "discounting". Since the discount rate is an inverse of the interest rate, it can be obtained by the formula of $1/(1+r)^n$, thus the value of ¥1000 in the next year would become ¥909 in the present value, following the formula using the discount rate factor of 0.909. As it

would appear it is rather laborious to calculate $(1+r)^n$ and $1/(1+r)^n$ for each year and each percentage, "discount factor tables" often used are usually found in the appendixes of financial analysis text book.

Gittinger once mentioned that "as the folk wisdom of people through the ages has recognized, present values are better than the same values in the future, and earlier return are better than later" (Note : 7). We can apply the idea of these proverbs to overcome the weaknesses of the undiscounted measures of project worth, and include a time dimension in evaluation through the discounting. In this regard, it is important to recognize that, any investment that would be made today would mean : (1) other alternative consumptions of today would have to be abandoned ; and (2) any consumption that would be brought about in the future (as a result of such an investment today) would be enjoyed by the generations that are not yet in existence today. Taking these aspects into consideration, in order to make the future value and the present worth consistent, it is necessary to introduce the concept of "discounting".

For example, if we have a project of 5 year period, and an initial capital would be invested in the first year, and the maintenance cost and gross revenue to be generated from the send year, we can make the cash-flow stream of the net benefit for the five year period. And if we can apply a certain discount ratio which would make the incremental cash flow equal zero. This discount rate is called "the internal rate of return : IRR". The IRR, in other words, means the marginal efficiency of capital, or the effectiveness of additional one unit of capital. Thus it can be observed that the higher the percentage of IRR, the higher the number of capital turnover, thus the greater productivity. One example of cash flow and calculation of IRR is shown below :

Project A

	Investment Cost	Mainte-nance Cost	Gross Revenue	Net Revenue	35% Discount Rate Applied (NPV)
1st Year	2000	0	0	−2000	−1480
2nd Year		500	1500	1000	548.7
3rd Year		500	1500	1000	406.44
4th Year		500	1500	1000	301.07
5th Year		500	1500	1000	223.01
			IRR	35%	Total : −0.78

***Note** : It should be noted that 35% of IRR is estimated in this case. This means that, if 35% discount rate factors are applied to the net revenues, its summation of Net Present Value (NPV) should become zero. But in this case, it is indicated as -0.78. This is mainly due to the discount factor decimals used. Thus there would be no problem in concluding that by applying 35% discount rate, the present net worth would become zero.*

It is also important to note the distinction and difference between Financial IRR (FIRR) and Economic IRR (EIRR). Take roads project for example, which aims at improvement from gravel roads to concrete pavement roads, initial investment cost and subsequent maintenance costs could be used for both FIRR and EIRR (however, there is a difference regarding the cost of labor which would be discussed later). For FIRR, the main revenue would be the income from the road users by way of collecting toll-gate fees. But in EIRR, the main benefits would consist of shortening (savings) of transportation time for both commuters and transportation of goods such as agricultural products. Vegetables would be less damaged under concrete roads compared with gravel roads. Maintenance costs of vehicles would also become less because the tires would last longer (because of less attrition with the pavement) and the travel distance per one litter of gasoline would become longer. These economic savings derived from concrete roads, compared with gravel roads (this is called a comparison of "with" project and "without project") could be expressed in monetary terms, and thus these economic savings could be used in the costs and benefit stream of the project cash flow analysis. One advantage of calculating FIRR is that it is possible to examine whether the project could be sustained without subsidy and also for setting of appropriate level of toll gate prices. On the EIRR aspect, the question is what should be the appropriate and acceptable benchmark percentage figure for any social infrastructure projects. In the international aid agency like the WB and ADB, EIRR is given more importance than FIRR and the minimum of 8% is expected for EIRR. However, for such projects as "poverty alleviation projects" and "life line" projects like water supply, below 8% is accepted.

Amongst the various aspects that need to be taken into consideration in calculating the EIRR, the use of international price (or known as Border Price) deserves special attention. In FIRR calculations, the prevailing market prices are normally used. But this rule does not apply to the EIRR. The main reason is that the prevailing market prices are sometimes distorted because, for instance, by the subsidy provided by the government. Thus, in calculating EIRR such items as subsidies and taxes (what is known as "transfer payments") have to be removed. In estimating EIRR, it is also important to recognize that the prices of the commodities, equipment, machineries, and services that would be procured and used under the project should be based on the internationally competitive price. Theoretically, the prices of commodities and services to be used in EIRR calculation should be the one

under "perfect" competitive market situation. This idea actually comes from the neo-classical economic school of thought. But in practice, it is extremely difficult to define what conditions would qualify the "perfect" competitive market. Therefore, in practice, the international market price is used instead. For example, if a project is intended to increase the production of rice in Indonesia for domestic consumption purpose, it has been the standard practice in EIRR estimation that the export price of rice produced in Thailand has been used in the past, because Thailand has been the largest exporter of rice in the world market. More specifically, CIF (cost, freight, insurance) price of the rice produced in Thailand has been used. The actual rice price to be used in EIRR calculation has been the CIF arrival cost at Jakarta port, plus the transportation cost from Jakarta to the project site. (this is known as "farm gate price"). If the project is intended for export, FOB (free on board) price would be used. In this context, it should be mentioned that for internationally traded goods, the international prices are relatively easy to know. But, one problem arises in pricing of labor that would be used in the project. The term "labor" could be divided into two categories. One is known as "skilled" labor in which case it is relatively easy to determine the international price because the international mobility of such labor force would be high. But, for "unskilled" labor, the pricing is not easy because such mobility would not exist. Therefore, for pricing of unskilled labor, what has been commonly done in the past is to use the "conversion factor" developed by the WB. Pricing of labor cost would become an important issue for those aid-related projects in the developing countries including the former socialist countries.

There are other technicalities that need to be taken into account in EIRR calculations such as the use of GDP deflator. At the time of post-evaluation, for instance, the project implementation period is already completed. Here, it is necessary to make price adjustments of commodities for the project implementation period, taking the completion year as starting point. All current costs and benefits would be converted into constant prices by applying the WB's manufacturer's unit value (MUV) index for the traded goods and GDP deflator for all local components. The constant price arrangement is needed in order that all expenditures and benefits should be expressed in real purchasing power, using a particular year as the base for comparison in a consistent manner. While there are some other factors that need to be considered in EIRR calculation, the author would not go into more details because of the shortage of space. But, amongst the items discussed, the most

important aspect in calculating EIRR is the use of international price. By using the international price, we can only ascertain regarding the economic viability of the project. Unfortunately, these basic rules, according to the view of the author, have not been properly applied in Japan in the evaluation works..

It should be added that the importance of the concept of "present value" and the corresponding "discount rate" has increased in the current international market. According to the "International Accounting Standards" (IAS), the price of the value registered in the fixed assets in the balance sheet of a private company, should not be based on the book value registered at the time of purchase. It should be based on the "present value" discounted at certain discount ratio. This discount rate would be equal to incremental income generation capacity (or expected return) of that fixed asset. This arrangement is needed, since the book value would not ensure the expected revenue that would be generated by that fixed assets. This is one of the aspects that the private sector company has to follow if it is intended to be listed in the international stock exchange markets. Only limited number of private firms have complied to the IAS.

For those projects where cash flow analysis are not applicable

A few words need to be mentioned regarding the projects in which cash flow analysis is not possible, thus the calculation of IRR cannot be made. One such example would be the engineering project in the Philippines done by ADB (Note : 8). This project aimed at improving the quality of education at the engineering departments of major universities in the Philippines. This project was post-evaluated in 1987. In 1992 (five years after post-evaluation) it was decided to re-evaluate this project with a view to reexamine whether the finings observed at the time of post-evaluation were still valid. At the time of post-evaluation, it was concluded that the project was "partly successful" mainly due to the two reasons : (1) about 40% of the equipment and materials procured under the project remained defective ; (2) except for few cases, most engineering departments were financially not sustainable. These conclusions were somewhat endorsed at the time of re-evaluation, except that use of some equipment had shown improvement. The re-evaluation team consisted of the author and a consultant who was a retiree/ ex-engineering professor. The basic issue raised by the re-evaluation team was how to evaluate the quality of education and this was not an easy task. After observing

the actual conditions at each engineering department and discussions that followed, the consultant proposed that one of the ways to evaluate the quality of education was to see the passing rate of national engineering exams conducted by the government. This has shown an interesting result. While the overall exam passing rate was downward, those departments participated in the project had shown, though moderately, better results. One has to be careful in interpreting these results. Equipment and materials procured under the project were "input", and not "output". Outputs (quality) can be partly judged by the passing rate of national exams. But in the context of the Philippines, particularly during 1900s, drop-out ratio of students was also high particularly in those universities that are located outside of urban cities. Scholarship programs also proved to be deficient. Under such circumstances, there were several cases that the students before graduation would go to Middle East as overseas workers, because the tractor drivers would not require high engineering skills, but could earn the salary several times higher that that of the Philippines. If we take into these conditions, the evaluation exercise would become more difficult. Indeed, more issues were raised in the re-evaluation than in the post-evaluation. The author came to the conclusion that one of the ways to evaluate the project like this is to review the cost of education against the expected life time income, so that some form of cash-flow analysis would become possible.

The above-mentioned story is only one example where evaluation methodologies need to be further developed. From these experiences, the author has come to the view that in future evaluation exercise, particularly like re-evaluation and a more broad based impact evaluation study, the inclusion of a "sociologist" and/or "anthropologist" should be seriously considered. One of the reasons is that, in a project like education, the incentives of the students would be an important area for review, this is something that has not been done. Also, socio-economic environment of the students need to be investigated. If considered in this manner, it would appear that the evaluation practice is a costly exercise if it is to be done properly.

Concluding remarks

From the above arguments, it would have become clear that those engaged in the evaluation would require wider perspectives. It is felt that it has come to a stage where a single economist or engineer would not be sufficient to understand and analyze the complex situation. The author has suggested the use of a sociologist for

gaining better view of the projects (whatever the sector may be). Secondly, it should be emphasized that generally speaking the general public's attention is likely to be focused on whether the project was success or not. But from the overall view point of the society that we live in, the real value of the evaluation practices and its results rest on the fact that to what extent the lessons learned from the past experiences have been effectively used for improving the future projects. The WB has spent a large amount of money for making easy access to the lessons learned data base through the use of computer programs, so that retrieval of information is available to every staff instantaneously. The EBRD has introduced a system that in the appraisal report, there is a section called, "lessons learned from past experience". These arrangements which may be called an effective "feedback system" is the area that the Japanese institutions have to improve. In order to improve the areas that the author has discussed so far, it is considered imperative that the Japanese institutions increase its interactions with IFIs. And eventually it is desirable that the evaluation practices would be accepted by the society as beneficial fruits emanating from past experiences, and that the socio-cultural barriers would be gradually removed.

(Note : 1)

According to the Oxford Advanced Learner's Dictionary (fourth edition, 1989), audit is defined as official (usually yearly) examination of accounts to see that they are in order. Evaluation is defined to find out or form an idea of the amount or value of something.

(Note : 2)

In Japan, the Ministries concerned have a final authority to determine the large scale ODA projects, but the evaluators of aid-related projects do not have access toe the records and files of the concerned Ministries.

(Note : 3)

It should be noted that the independence of evaluation professionals as well as that of auditing firms has received an international attention since the collapse of ENRON in the States (where the auditor's role has been questioned as to why the window dressings of the financial statements had not been critically examined and detected by the auditing firm engaged). Similar case was spotted in Japan where

the Mitsubishi Motors have been hiding the records of those defective cars which should have been reported to the concerned ministry for a recall. This kind of situation, if continues, it is likely that the government would have to issue a new bill to rectify the situation (but such action by the government would also likely to undermine the initiative that should have been taken by the private sector company). Therefore, it would appear that the independence and impartiality of the evaluation/audit professionals are getting more critical for developing the greater confidence desired in the market economy.

(Note : 4)

In the International Finance Institutions (IFIs) like WB, ADB and EBRD, the evaluation results are usually classified into three categories ; "generally successful", "partly successful", and "unsuccessful". Some IFIs use additional categories like "highly successful", but basic criteria could be classified into three. It would be of some interest to note the ratings results. According to the Operations Evaluation Department of ADB, of 976 projects and programs, evaluated by the end of 2003, and completed from 1974 to 2002, distribution of ratings indicates : 62% rated as highly successful, generally successful, or successful ; 30% rated as partly successful ; and 8% rated as unsuccessful. For EBRD, during 1996-2003, about 45-50% rated as highly successful, successful ; about 25-30% rated as partly successful ; and about 15-20% rated as unsuccessful. The main reason for relatively high unsuccessful cases is that in the case of EBRD, about 60% of its portfolio has to be in the private sector operation (meaning that there would be no sovereign guarantee) according to the EBRD Charter, thus operation risks are higher compared to other IFIs. It is also noteworthy to mention that in the case of EBRD, in addition to success ratings, "transition impact" is also evaluated. For details, see "Annual Review of Evaluation Activities in 2003", Project Evaluation Department, European Bank for Reconstruction and Development (EBRD), 2003

(Note : 5)

When I reviewed the evaluation reports of JICA and JBIC, it was noted that at the initial stage, most of the evaluation works had been done by the staff who belonged to the associations and institutions that are affiliates of the Ministry of Foreign Affairs (the one that supervises the JICA). In such a situation, it is difficult to expect something critical to the works done by JICA. Recently,

however, the professors who are members of Japan International Development Society and Japan Evaluation Society have been deployed for the evaluation assignments. This can be considered as an improvement, but the shortages of professional evaluation consultants still persists.

(Note : 6)

It should be noted that the initial contributions in the field of evaluation methodologies were done by *Little*, I.M.D., and *Mirrlees*, J.A. (Manual of Industrial Project Analysis in Developing Countries, 1969, OECD). The theoretical framework was further developed by *Gittinger*, J. Price, (Economic Analysis of Agricultural Projects, 1982, The World Bank, published by the Johns Hopkins University Press).

(Note : 7)

For details, see pp.304-329 of the below listed reference of *Gittinger*.

(Note : 8)

For details, see Impact Evaluation Study : Re-Evaluation of Engineering Education Project in the Philippines, Asian Development Bank, mentioned in the reference list.

Reference Materials

Manual of Industrial Project Analysis in Developing Countries, I.M.D. Little, and J.A. Mirrlees, OECD, 1969.

Economic Analysis of Agricultural Projects, J. Price Gittinger, The Johns Hopkins University Press, 1982.

Cost-benefit analysis, E.J. Mishan, George Allen & Unwin, 1975.

Project Evaluation : Techniques and Practices for Developing Countries, Heng-Kang Sang, Wilson Press, New York, 1988.

Guidelines for Project Evaluation, United Nations, 1972.

Economic Analysis of Projects, Lyn Squire and Herman G. van der Tak, The World Bank, published by the Johns Hopkins University Press, 1975.

Assessing Development Effectiveness, World Bank Operations Evaluation Department, 1998.

Guidelines for the Economic Analysis of Projects, Asian Development Bank, 1997.

Investment Appraisal and Financial Decisions, 5th Edition, Steve Lumby, Chapman & Hall, 1995.

Guidelines for the Preparation of Project Performance Audit Reports, Asian Development Bank, September 2000.

Annual Review of Evaluation Activities in 2003, Project Evaluation Department, European Bank for Reconstruction and Development (EBRD), 2003.

A Review of the New Role of Project Evaluation, Kimihiro Kaku, Government Auditing Review Volume 4, Board of Audit, Japan, 1997.

Impact Evaluation Study : Re-Evaluation of Engineering Education Project in the Philippines, Asian Development Bank, 1992.

Author's Note

In addition to the above-mentioned references, it should be added that several individual evaluation reports have been reviewed, particularly those from the *World Bank*, *Asian Development Bank* (ADB), *European Bank for Reconstruction and Development* (EBRD), *Japan International Cooperation Agency* (JICA), Japan Bank for International Cooperation (JBIC). But since the numbers of reports are numerous, the author would like to note that each individual evaluation report is not mentioned in this paper.