

## Challenging times for evaluation of international development assistance

*Masafumi Nagao*  
*Hiroshima University*

### Introduction

Evaluation of development assistance first gained international prominence when an Expert Group on Aid Evaluation was created within OECD's Development Assistance Committee (DAC) in 1982. Behind its creation was an 'explosion of interest' among aid donor countries for institutionalizing its practice (Dabelstein and Rebien 2002). This group helped shape a common understanding among donor countries concerning evaluation methods, standardization of terminology and coordination of joint evaluation efforts. It also established the so-called 5 DAC Evaluation Criteria (i.e., relevance, effectiveness, efficiency, impact and sustainability), which are now used by most aid agencies and international organizations for evaluating their assistance projects. The work done by this group was quite instrumental in propagating evaluation activities in developing countries. It also helped spread evaluation practice in some of the donor countries, such as, for example, Japan.<sup>1</sup> This working group may, therefore, be said to have contributed, at least to some extent, to 'internationalization' of the practice of evaluation.

The scope of evaluation actually done on development assistance projects, however, remained quite modest for many years. It consisted mostly of project monitoring and output accounting, and not much attention was paid to systematic verification of outcomes achieved or impacts on the target beneficiary (Bamberger 2000). Even in the World Bank, where development evaluation had been practiced for many years, a review of the project appraisal documents showed that the proportion of projects with planned impact evaluations in 2000 counted only 10 % (Rawlings 2005). It was only in the late 1990s with the spread of the practice of Results-Based Management (RBM) in the donor country governments that evaluation came to assume greater relevance in development assistance. Under pressure from deteriorating public finances and unfavorable economic conditions, these governments subjected the programs and projects of their aid agencies through the same fiscal scrutiny as those of all the other ministries and departments. The result has been a thorough reform of organization and procedures for evaluation practice in many of these agencies (DAC 2000). RBM has

---

<sup>1</sup> This is evidenced by the fact that the greater majority of the members of the Japan Evaluation Society, numbering around 500 today, specialize in evaluation of international development assistance.

also been adopted by the multilateral organizations such as the World Bank and the UNDP. Furthermore, through the aid programs for public sector reform and improved governance, as well as for evaluation capacity building, of these organizations, the practice of public sector evaluation has begun to spread to developing countries (Weisner 1997; Guerrero 1999)..

With increased relevance and acceptance of its practice, evaluation of development assistance must still face challenges arising from the complexities inherent to international development assistance as well as from the rapidly changing context in international development. The challenges are quite formidable. There is, to start with, a widely-held apprehension that years of ongoing aid have not produced visible and sustainable impact on development in many developing countries (Chapman and Nagao 2006). There is also a somber reminder that evaluation may not affect actual policy or behavior of the donor agency (Lindahl and Catterson 2005).

This paper proposes to review some key issues arising from these challenges for the practice of evaluation, touching on practical as well as theoretical concerns. . It argues that there exist certain contextual factors, some old and others new, which complicate this practice, but that the increasing trend toward local ownership of the development process and accompanying emphasis on capacity development of the beneficiary from development, that is, developing countries, when coupled with empowerment approaches in evaluation, may give rise to new and promising roles for evaluation in development and development cooperation. This will be illustrated by a mathematics and science education project in South Africa in which the author participated as a leader of a Japanese technical assistance team and in which a locally-owned formative monitoring system was instituted for establishing an innovative teacher training practice.

### **Some complicating factors in evaluating international development assistance**

Before entering into the discussion of issues and concerns preoccupying those studying or engaged in evaluation of international development assistance, it may be useful to reflect briefly, and in somewhat generic terms, on the distinguishing features of evaluation in this area. What is considered here is not a diffusion perspective on the global development of evaluation practice, which is well-presented in a 21-country review by Frubo, Rist and Sandahl (2002). Rather it concerns a negotiation perspective in evaluating joint development efforts of various kinds between developed country donors and developing country recipients. What distinguishes it from the usual stakeholder-based approach to evaluation a la Patton (1997) and Weiss (1998) is that there is a basic asymmetry of interest regarding evaluation between the two principal parties and that these parties are no less than sovereign states with their respective agenda (Garaway 2003).

#### *Asymmetry of interest*

The donor-recipient asymmetry of interest manifests itself in a number of ways, generally tending to complicate the evaluation exercise. To start with, there is

divergent disposition in respect to the purpose of evaluation (Nagao 2001). The party that insists on evaluation is usually the donor side who must account for the development assistance funds and who need recognition from the beneficiary for their cooperation in order to satisfy the demand from the tax-payers. The aid agencies are hard-pressed to show that they are engaged in outcome-oriented planning and results-based evaluations of project effectiveness. The recipient side, on the other hand, is generally not interested in evaluation of projects that are ending or have ended, unless the exercise could lead to future project possibilities, and even then their interest would be a passive one. The participation of multilateral development agencies and non-governmental organizations as donors may modify but would not alter drastically the asymmetric character of the donor-donee relationship.

#### *Difference in time horizon*

One clear sign of asymmetry is the difference in time horizon. Donors clearly have a finite time horizon, and have a tendency to target tangible short-term results from aid efforts implemented over a limited time span. For developing countries development is not a time-bound phenomenon. They are obviously more interested in the long-term investment effects and developmental impact. This difference in time horizon may pose less problem for evaluation of single-purpose projects whose parameters have already been well-established (e.g., school construction, dental service). For projects with more comprehensive approaches, which are increasing especially in the area of social development and for which contextual factors vary greatly, the evaluation task becomes more complicated. Development assistance by donors still tend to push 'best practices' with fixed time scale and backed by conditionalities, leading often to neither active learning nor lasting institutional changes (Ellerman 2004).

#### *Evaluation of development or impact of aid*

The asymmetry of interest also shows in respect to what to evaluate. Both the donor and the recipient may look at the output, outcome and impact generated by the project. The donor's basic interest lies in assessing the impact of their assistance in generating the output and outcome. Unless they can demonstrate the utility of the assistance, they will not be able to stay long as a donor. The recipient's primary interest by contrast is in evaluating the development itself. How much contribution the donor assistance has made in bringing about that development is of secondary importance. Furthermore, the recipient side is paying increasing attention to capacity building in the process of development. This has raised their interest in evaluation of project process (Minamoto and Nagao 2006). In some technical cooperation projects, the donor side retains the tight control of the project process so as to make sure that the intervention will generate output or outcome. Typically such projects remain as 'add-on' projects outside the regular development program of the recipient governments. How to organize the development process with aid intervention has surfaced as a point of contention between the donor and the recipient.

#### *Evaluator selection*

According to evaluation textbooks, evaluators should serve as an independent entity to mediate between the funder of evaluation and the beneficiary and to mitigate the negative impact of the asymmetry of interests between the two. This, however, is difficult in the case of evaluation of development assistance. Evaluation of aid projects is done usually by consultants selected and sent by the donor agency from home. Evaluator selection is made from among a relatively small group of professionals who get commissions from the agencies not only for evaluation but also for feasibility and planning studies and who, as a result, 'think' like the agencies. The recipient side is not consulted by the donor agency about the choice of evaluators, or their terms of reference in most cases. Some donor agencies have hired local evaluators or evaluators from other developed countries, but this is still a rare occurrence. Besides, they may not be as independent from the thinking of donor agencies because of their dependence on contracts from the agencies. There is, therefore, no 'right' answer as to who should do the evaluation (Stewart 2005).

### *Effectiveness-efficiency or Sustainability*

In terms of the 5 DAC Evaluation criteria, the donor agencies' concern for results-based evaluation may be best represented by a relative emphasis placed on the combination of effectiveness and efficiency (Svensson 1997). Evaluation should indicate that the project has achieved the intended purpose (i.e., effectiveness) and has done so cost-effectively (i.e., efficiency). A recurring though relatively neglected theme in development aid debate, however, is the sustainability of the impact of the project after its termination (Brinkerhoff and Goldsmith 1992; CIDA 2002). Although the sustainability criterion of evaluation checks different conditions under which the impact may be sustained, how it actually turns out would really depend on the disposition of the developing country concerned. Donor agencies are increasingly inclined to an intervention approach that includes the creation of a mechanism that ensures that the results are firmly cemented and widely disseminated based on the efforts of the recipient countries. The central focus of such efforts tends to be placed on capacity building, supporting the steps which the recipient country should take to develop its own problem-solving capability, capacity to conduct evaluation being an important component of the latter.

### **Challenges from the Changing development context**

The foregoing discussion focused on the distinguishing features of the development assistance evaluation which implies somewhat greater complexity, and perhaps a greater challenge, than the evaluation usually practiced in the domestic context of any country. In recent years, development evaluation has experienced additional, and quite demanding, challenges stemming from the changing international development context. These challenges include the increasing incidence of assistance in emergency situations involving political instability and armed conflicts, the adoption of global development targets such as Millennium Development Goals and the spreading practice of multi-donor assistance.

### *Assistance in complex emergencies*

Since the end of the Cold War, the occurrence of open intra-state conflicts has multiplied. The resulting situations often involving political instability and armed struggles has called for immediate humanitarian intervention for restoring peace, followed by extended assistance for reconstruction and development. Evaluation to be done in such 'complex emergencies' faces special conditions and requirements such as assessing the security situation and the protection needs of the affected population, coping with the lack of vital information, verifying coherence of assistance policies of the parties involved and workability of coordination mechanisms (DAC 1999). Given the volatility and fluidity of the situation on the ground, those assigned the task of evaluating the assistance would be hard pressed just to grasp what is happening and why, rather than theorizing and asking what caused what. Assistance policies may change quite often and in typically unpredictable ways, sometimes making evaluation only an after-thought. Furthermore, although apparent motive is humanitarian for such assistance, there is usually an underlying strategic objective of the donor governments concerned, which casts a strong political overtone for the evaluation exercise. The evaluator may be forced to engage in policy analysis instead of policy evaluation.

#### *Adoption of global development targets*

Against the background of the development of the 'ownership-partnership' discourse concerning the nature of the development process initiated by the New Development Aid Strategy of DAC countries in the mid-1990s (DAC 1996), the international development agenda has come to embody globally-set development targets as typified by the Millennium Development Goals. These targets represent a development consensus, with a series of time-lines for achieving certain targets, such as poverty reduction, achievement of universal primary education and gender equity in educational opportunity. This new trend is accompanied by increasing reliance on program- and policy-based assistance, rather than project-based assistance. All this would require major changes in policies at the national level in developing countries and on the donor side, which are difficult to bring about (Picciotto 2003). With policy-related conditions in flux, evaluation of the assistance involved would have to be made against much uncertainty (White 2005). Achieving a measure of transparency and assuring accountability to the national audience for the essentiality of the assistance poses quite a challenge to the donor agencies involved.

#### *Spreading practice of multi-donor assistance*

The global targeting in development assistance is increasingly associated with a multi-donor assistance structure of sector-wide approaches, direct budget support to developing country governments and the Poverty Reduction Strategy Papers. In spite of the ownership discourse that underlies this structure, a sizable number of developing countries, especially in Sub-Saharan Africa, are becoming more and more dependent on external financial aid, with such aid running at between 40 and 50% of the government's entire recurrent budget. Thus, while evaluation is asked to assess effectiveness and efficiency of the multi-donor assistance involved in achieving the international development targets and to verify that this process takes place

increasingly with developing country governments exercising control, a more fundamental question of sustainability of the development assistance effort and, for that matter, the development process itself need to be seriously questioned on a global scale (King 2004).

### **Capacity development and the role of evaluation**

It is out of reflection of these changes in the international development context by the developing countries and the donor community alike that serious attention has come to be paid to capacity development and more self-reliant approaches this may encourage within developing countries. Capacity development concern at the level of individual development projects is nothing new; as it has been promoted in a narrow technical sense to accumulate local knowledge and skills for specific tasks which technical cooperation should account for as an added dimension. What is new is that this development is now increasingly viewed as a process for enhancing more general problem-solving capacity in an institutional and project management sense. What is even more significant is that the key to the successful implementation of this process is believed to be the country ownership of the capacity development itself (DAC 2006).

The ownership focus in the development assistance context usually leads to participatory trends in the project evaluation practice. The participatory nature may vary depending on the purpose of evaluation and the degree of local contribution to the evaluation work. A comprehensive representation of different alternatives is provided by Owen's notion (or form) of interactive evaluation (Owen 2006). The principal focus of interactive evaluation is project delivery, especially its improvement, but its application may involve different approaches, such as responsive evaluation, action research, developmental evaluation and empowerment evaluation. In relation to the ownership discourse and local capacity building in development evaluation, what is most relevant is empowerment evaluation which is designed to facilitate the control of the process by the local project participants, including the direct beneficiaries, through their engagement in the development and evaluation of the projects and programs involved.

In the negotiating framework of empowerment evaluation, the funder and the recipient must overcome the asymmetric nature of their relationship and arrive at a shared definition and understanding about performing an evaluation that emphasizes dialogue on an equal footing (Fetterman 2000). On that basis, a project can be designed to incorporate evaluation function into its implementation process in such a way as to raise both evaluation capabilities and the consciousness of the concerned parties with respect to that process. Here the role of the evaluator is critical. It differs from that of conventional evaluations, in which the evaluator performs judgment based on evaluation techniques at his disposal. Instead the techniques required here are those which enhance collaboration with, or cede leadership to, the people involved with respect to the work of evaluation. A significant feature of this approach is that the evaluation process promotes the 'buy-in' of the stakeholders, assuring the utilization of the evaluation results (Patton, 1997). Adoption of this approach, however, calls for



facilitation capability, interpersonal communication skills, negotiating techniques, and so on. In addition, because it is important to make evaluation an integral part of the project, the evaluator must often act as a trainer or a technical advisor.

Application of empowerment evaluation in development assistance setting, however, faces additional difficulties. Above all the donor agency, while agreeing to adopt empowerment thinking in the project to be implemented, would not be persuaded to assign an independent evaluator to facilitate the process for a project which may take place over several years and in a distant location. It is simply not in their mind set. The task of organizing the process then would fall on the project manager or advisor, usually a consultant dispatched by the donor agency from home. These professionals may have a thorough understanding of empowerment approaches and may even agree to implementing them as part of their project. However, they would not be able to escape from the fact that their principal task is to plan and implement the project according to the expectation of the donor agency which contracts them, often resulting in relative neglect of the evaluation function. Should empowerment evaluation in development assistance setting, then, depend on the human factor – on the particular disposition of the project manager? This is not necessarily so. A project may be structured in such a way as to make application of an empowerment approach both feasible and useful. The key lies in elaborating a project design that works towards establishment of a systemic effort by all the stakeholders involved, taking account of the particular context in which the project is carried out. In the next section, this is illustrated in terms of a secondary mathematics and science teacher retraining project in South Africa in which the author participated for 7 years from 1999 to 2006 as a member of the Japanese technical cooperation team.

### **The Mpumalanga Secondary Science Initiative (MSSI) Project**

The Mpumalanga Secondary Science Initiative (Mpumalanga Dept. of Education 1999) was a project carried out by the Department of Education in the Mpumalanga Province, South Africa from 1999 to the beginning of 2006. It aimed at improving the quality of mathematics and science (hereafter to be referred to as 'M & S') teaching in classrooms in the province's 540 secondary schools (Please see the project summary in Figure 1). The Department proposed to do this by retraining its M & S teachers. The retraining was needed not only to compensate for whatever gaps and deficiencies that existed in their instructional capacity owing to the training shortfall dating back to the apartheid times, but also to facilitate the introduction of a major curriculum reform which was based on an ambitious outcome-based approach. The retraining was done with the collaborative support of the Japan International Cooperation Agency (JICA), the Japanese government agency for technical cooperation, and the University of Pretoria (UP), a local university, by applying a cascade model of training. The project was designed to establish a school-based in-service training system for all secondary M and S teachers.

### Figure 1. Main Elements of the MSSI Project

- Goal: Improved M & S understanding of secondary students
- Aim:
  1. Improvement M & S teaching via teacher retraining
  2. Development of a Province-wide system of School-based in-service training (INSET)
- Duration: Phase 1 November 1999 ~ March 2003  
Phase 2 April 2003 ~ March 2006
- Partners: Mpumalanga Dept of Education / JICA / University of Pretoria (UP)
- Target population: M & S teachers in all (540) secondary schools
  
- Characteristic approaches:
  1. Retraining for teacher capacity improvement and curriculum reform
  2. Cascade model of training targeted at school-based INSET
  3. Promotion of Peer Teacher Learning as a project instrument
  4. Individual teacher incentives through UP accreditation scheme
  5. Extensive use of monitoring and evaluation as a tool of project development

MSSI had several distinct characteristics. It tried to combine different teacher training concerns into a single in-service training program. It targeted all the secondary schools and all the M & S teachers in the Province, rather than employing a pilot approach. It aimed at establishing a system of in-service training by which teachers would come together once a month at their own schools to engage in peer collaboration exercise for improving lesson plans and teaching methods in M & S. The practical working of this system was visualized in terms of monthly meetings organized at each school by the M & S teachers to engage in peer learning activities.

The project sought to utilize technical inputs from both foreign (i.e., JICA) and domestic (i.e., University of Pretoria) sources. This last was by necessity. Although Japan possessed a lot of accumulated knowledge and experience in M & S education, JICA's experience in providing M & S education assistance was still limited – only since the mid 1990's, so it could not claim to be an accomplished service provider. Besides, for JICA, this was the first major cooperation engagement in South Africa, and its knowledge base about education in the target country was no more than a few hastily conducted studies. For these reasons the department had some skepticism about the utility of Japanese technical assistance. The University of Pretoria could not only serve as a local guide for the Japanese side but also, and more importantly, as a project-based researcher for interpreting as well as extracting lessons from the Japanese experience. JICA's presence, on the other hand, could help lessen the historical tension that could have existed between the university and the department.

All these facets of the project were worked out in a tentative manner in the initial process of planning and elaborated in the process of project implementation. What evolved from the initial effort was a new model of technical cooperation termed 'experience-sharing' model (Nagao 2004). This model contrasted with the usual



technology transfer model of technical cooperation in a number of respects as shown in Figure 2. The object of the cooperation activity in this model is not transfer of a technology but of an ‘experience’. The principal mode of intervention would typically be exposure of a group(s) of individuals from the developing country to the relevant experience of the cooperating country rather than on-site instruction by dispatched experts from the latter. The target outcome would not be an autonomous use of the transferred technology as in the case of the technology transfer model, but more likely be the formation of an autonomous system and practice utilizing the experience. And most importantly, the success of the experience-sharing model would not depend on the quality of the technology supplying transfer agent but on the quality of learning of both sides, especially the developing country side.

**Figure 2 Technical cooperation model: technology transfer vs. experience sharing**

| <b>Contrasting points</b> | <b>Transfer of technology model</b>                      | <b>Experience sharing model</b>                                    |
|---------------------------|----------------------------------------------------------|--------------------------------------------------------------------|
| Typical field             | Manufacturing / Transport                                | Education / Health / Social welfare                                |
| Cooperation aim           | Transfer of technology                                   | Transfer of experience                                             |
| Principal means           | Dispatch of expert team by donor for instruction on site | Exposure of aided country team to relevant donor experience        |
| Donor agents              | Technical experts                                        | Individuals, organizations and area communities holding experience |
| Recipient agent           | Technical counterparts                                   | Groups, organizations and local communities                        |
| Target outcome            | Autonomous use of technology                             | Formation of autonomous system and practice                        |
| Key to success            | Quality of transfer agents                               | Quality of learning by both sides                                  |

Source: Nagao (2004)

The experience-sharing model is characterized by two distinguishing features - namely, symmetry of relationship between the technology ‘supplier’ and ‘recipient’, centrality of the learning function, and the importance of managing the cultural factor. The symmetry of relationship in any experience sharing is essential, since whether or not actual sharing takes place would depend on the disposition of the ‘recipient’, rather than the ‘supplier’. The latter may offer what it considers the ‘best’ experience, but it will simply not be ‘shared’ if the former does not find it ‘attractive’. The experience-sharing model, therefore, implies the receiving side’s ownership of the project, which reflects the rethinking on the nature of the donor-recipient partnership towards greater national ownership of aided projects by developing countries (King, 1998).

The other critical and instrumental factor of the experience-sharing model is the learning associated with the cooperation project. Learning here means a process by which an individual or an organization grasps elements of a new experience, reflects on them, and uses and adapts them so as to improve the performance. In the MSSI project, the University of Pretoria team spearheaded the learning effort, but equal emphasis was placed on peer teacher learning in regular in-service training sessions

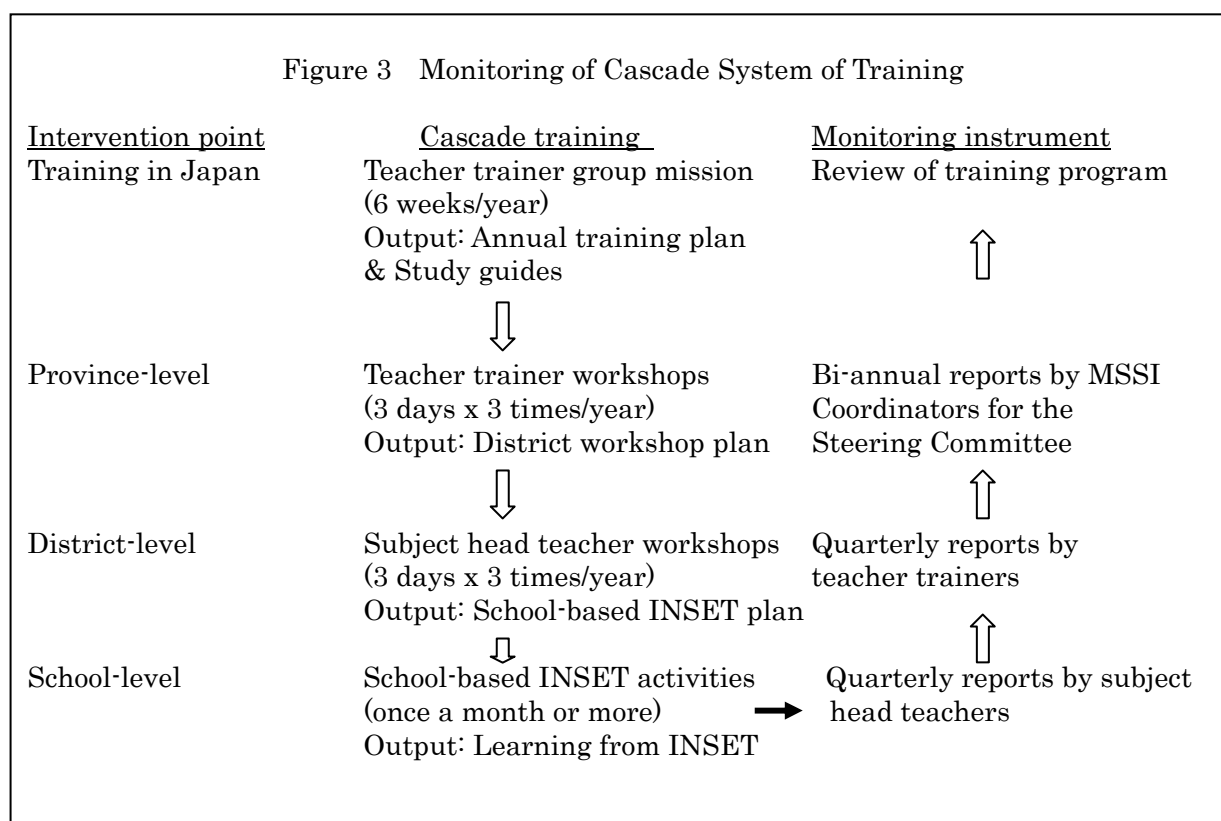
by the M & S teachers. Particularly instrumental in the latter was the adoption by the teachers of the practice of 'group reflection' as a daily routine whenever MSSSI activities took place, taking after the Japanese practice of reflection for continuous improvement. It should be pointed out also that the learning required does not relate only to technical contents of the experience to be studied. The smooth functioning of the experience-sharing model is predicated on mutual understanding and close communication between the partners involved, which should require active learning and management of the cultural factor in the partnership relationship.

### **The use of empowerment approach in the MSSSI project**

In the MSSSI project, the partners agreed from the beginning that systematic monitoring of the project activities would be essential for the success of the project. As the project's objective was to establish a system of school-based INSET, this monitoring function was to be built around capturing what went on at the province's schools in terms of INSET activities. As shown in Figure 3, the project's intervention was organized as a cascade chain of training activities – starting with (i) a study-cum-training mission to Japan for a teacher trainer group, who prepared an annual training plan and study guides on the basis of material development work, (ii) a province-level workshop for all the teacher trainers to share what the Japan mission group prepared, the result of which was District-level workshop plans, (iii) District-level workshops for M & S subject head teachers from the participating schools, and (iv) school-based INSET activities (e.g., once a month) organized by the subject head teachers. The interventions at the Provincial and District level, as well as training in Japan, were generally well accounted for.

The key monitoring activity consisted of capturing the INSET activities which went on at each school. Each time the M & S subject head teacher organized an INSET activity at his/her school, he/she was to write a half-a-page account of the activity using a standard format (i.e., date, place, topic, presenter's name, activity content, resources/materials used/assessment of the session quality) generated by the coordinators' team. At the end of each quarter the subject head teacher collate all the session accounts and prepares a quarterly report. This report is presented by the principal of respective schools in a joint periodic review exercise organized on the first day of each District-level workshop. These reports from the schools are collated and compiled into a quarterly District-level report by the teacher trainers concerned and sent to the MSSSI secretariat, which will prepare a status report twice a year for submission to the Steering Committee.

Figure 3 Monitoring of Cascade System of Training



The whole scheme worked in the image of Patton’s collaborative evaluation practice with its emphasis on a group approach (Patton 1982). Although the formal, systemic construction of the monitoring system was slow, the adoption of group practice, such as peer teacher learning and group reflection, apparently took roots much sooner. MSSSI project internalized monitoring and continuous reflection in its project process and, thereby, helped foster empowerment of both the officials and organization of the department more or less in the manner described by Fetterman (2000). Thus, the internal system of formal monitoring was accepted by all the key stakeholders as the shared project instrument. The experience-sharing model of the MSSSI project was, thus, set in the formative monitoring system of the project activities.

There remained a question of validation associated with the functioning of the formative character of the monitoring system in the longer run. The monitoring system was established not only to capture the extent and kind of project activities that took place at various levels of educational administration but more importantly to encourage sharing of improved practices among different schools through peer teacher learning activities and exchanges of session notes covering improved practices. With the termination of the partnership project, if the positive practices initiated could be sustained remains a question. The project initially envisaged combining the systemic effort for promoting cascade training activities and accompanying monitoring with the provision of a custom-made study opportunities for individual teachers backed by an accreditation scheme supported by the University of Pretoria. The latter was to provide the incentives the teachers would need to sustain interest in

continued learning. The university started a scheme but it did not take root and a new scheme is being debated.

How would the empowerment approach followed in the MSSSI project fare in the longer run? Would the formative monitoring system built around school-based INSET system evolve to make the experience-sharing model a viable option? As the MSSSI project was terminated officially in March 2006, the three partner institutions of the project conducted a joint evaluation (MSSSI 2006). Its concluding section read, in part, as follows:

*“The MSSSI, which was initiated by the MDoE in November 1999 with the aim of strengthening the secondary M&S education in the Province, comes to an end on March 31, 2006. Through two phases of active intervention, the department has been able to make significant advances in achieving its specific project purpose – to establish a Province-wide system of INSET for secondary M&S teachers. In so doing the department has benefited from the collaborative partnership of JICA and UP. The work to build the INSET system, however, is not yet complete. In fact, the nature of this work is such that the system-building effort must continue in the spirit of the MSSSI-nurtured principle of ‘learning on the way’.”*

The Mpumalanga Department of Education instituted a ‘sustainability strategy’ for further promotion of the system-building effort on the basis of continued collaboration with the University of Pretoria to follow up on the evaluation exercise. Some of the comments shown below which were made by the department officials working close to schools during the interviews for the evaluation exercise may provide indications as to the optimistic forecast the department has for the fate of this strategy (MSSSI 2006):

*“MSSSI has provided the support to M&S education in an organised way – from planning to implementation and evaluation of interventions. Its aim to establish a Province-wide system of INSET may have already been achieved in the sense that the systemic learning process is established”*

*“The INSET system has taken root. The discussion among teachers for improvement of classroom teaching will continue.”*

*“The specific thing that one would say MSSSI succeeded in doing was to ignite that passion among the [M&S] educators to conduct INSET and also to work on cluster in order to have common programmes in M&S.”*

*“For the first time in our history I saw educators of M&S in our schools going extra miles whereby they will come together to conduct INSET among themselves so that they could improve their performance.”*

### **Some concluding remarks**

The spread of results-based management in developed countries, combined with an increasing tax-payers’ concern for non-performing development assistance projects, has increased the relevance of evaluation in international development assistance. Its

practice, however, faces challenges stemming from the ‘more than usual’ complexities of evaluation circumstances, such as asymmetry of interest in evaluation between an aid donor and a recipient and different orientation to development as reflected in objectives and time horizons. More recently the changing international development contexts are posing new and complicating challenges. They include growing assistance in complex emergencies accompanied by increasingly strategic orientation of donor policies, the adoption of global development targets such as Millennium Development Goals and the need to align evaluation framework to macro-level goals and policies, and the spreading practice of multi-donor assistance with reduced transparency for achievement of national assistance objectives. Development evaluation is indeed facing challenging times.

In this new setting for evaluation of international development assistance, this paper has pointed out that both the developing countries and the donor community alike are paying increasingly serious attention to capacity development and more self-reliant approaches this may encourage within these countries. Moreover, there is an emerging consensus that the key to the successful implementation of this process lies in the country ownership of the capacity development itself. In terms of evaluation practice this new concern has signified the increasing relevance of the interactive approaches, especially the application of empowerment thinking, which gives greater control about the planning, implementation and evaluation of development undertakings to the developing countries.

To illustrate the application of empowerment approaches to a concrete development project, this paper has described a mathematics and science teacher training project in South Africa in which the author participated as a leader of the Japanese technical cooperation team. This project was based on an ‘experience-sharing’ model, rather than the usual technology transfer model, of technical cooperation, and was designed to promote the learning by the South African side about the relevant elements in the mathematics and science education development experience of Japan. In order to ensure that this learning process was conducted on the basis of local control and in an empowering fashion, a local university joined the project as a full partner with a research and evaluation function. Moreover, in order to establish a school-based in-service training system using a cascade model of training and a peer teacher learning approach, a formative monitoring mechanism was established to stimulate the systemic learning effect. The experience showed some promise for the application of the empowerment approach based on a more symmetric relationship between the aid donor and the recipient, although whether or not the developing country partner can sustain the positive gains made with the project in the longer run is yet to be verified and changing reality of the development context can always threaten such development. This is but a micro reflection of the big picture of the development evaluation, where there is still a widely-recognized gap in the knowledge base on what works and what does not in different contexts. In the words of Francis Stewart (2005), a well recognized authority on development, “evaluation will always be work-in-progress, and contested”. What we evaluators can do is to continue and search for a better product of work in progress.

## References

- Bamberger, M. (2000) "The Evaluation of International Development Programs: A View from the Front" *American Journal of Evaluation*, Vol. 21, No. 1, pp. 95-102.
- Bamberger, M., Rugh J., Church, M. and Fort, L. (2004) "Shoestring Evaluation: Designing Impact Evaluations under Budget, Time and Data Constraints" *American Journal of Evaluation*, Vol. 25, No. 1, pp. 5-37.
- Binnedijk, A. (2001). *Results Based Management in the Development Co-operation Agencies: A Review of Experience*. Paris: OECD.
- Brinkerhoff, D. and Goldsmith, A. A. (1992) "Promoting the Sustainability of Development Institutions: A Framework for Strategy, *World Development*, Vo. 20, No. 3, pp. 369-383.
- Canadian International Development Agency (CIDA) (2002) "Assessing Sustainability" What We're Learning NO. 2, Ottawa: Canadian International Development Agency.
- Chapman, D. and Nagao, M. (2006) "Editorial: The Evaluation of International Cooperation in Education: Assuring Quality and Effectiveness during Rapid Change" *Journal of International Cooperation in Education*, Vol. 9, No.1, pp. 1-5.
- DAC (1999) *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies*, Evaluation and Aid Effectiveness 1, Paris: OECD.
- DAC (2000) *Results Based Management in the Development Co-operation Agencies: A Review of Experience, Executive Summary*, Working Party on Aid Evaluation, Paris: OECD.
- DAC (2006) *The Challenge of Capacity Development: Working Towards Good Practice*, DAC Network on Governance, Paris: OECD.
- Dabelstein, N. and Rebien, C. C. (2002) "Evaluation of Development Assistance: Its Start, Progress, and Current Challenges" in Furubo, J-E., Rist R. C. and Sandahl, R. (eds.), *International Atlas of Evaluation*, New Brunswick, NJ: Transaction Publishers, pp. 393-405.
- Ellerman, D. (2004) "Autonomy in Education and Development" *Journal of International Cooperation in Education*, Vol. 7, No. 1, pp. 3-14..
- Fetterman, D. (2000). Steps of Empowerment Evaluation: From California to Cape Town. In Stufflebeam, Daniel L. et. al. (Eds.) *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, 2<sup>nd</sup> Edition, Boston: Kluwer Academic Publishers (pp.395-408).
- Furubo, J-E., Rist R. C. and Sandahl, R. (eds.), (2002) *International Atlas of Evaluation*, New Brunswick, NJ: Transaction Publishers.
- Garaway, Gila (2003) "Evaluating Educational Programs and Projects in the Third World" in Kellaghan, Thomas and Stufflebeam, Daniel L (Ed.) *International Handbook of Educational Evaluation*, Dordrecht: Kluwer Academic Publishers, pp. 701-720.
- Guerrero, R. P. (1999) "Evaluation Capacity Development in Developing Countries: Applying the Lessons from Experience" in Boyle, R. and Lemaire, D. (ed.) *Building Effective Evaluation Capacity*, New Brunswick, NJ: Transaction Publishers, pp. 177-194..
- Ingram, G. K. (2005) "Overview" in Pitman, G. K., Feinstein, O. N. and Ingram, G. K. (eds.)



*Evaluating Development Effectiveness*, World Bank Series on Evaluation and Development, Vol. 7, New Brunswick, N.J.: Transaction Publishers, pp. xiii-xx.

King, K. (1998). New Challenges to International Development cooperation in Education. *Journal of International Cooperation in Education*, Vol. 1, No. 1, pp. 3-16.

King, K. (2004) "The External Agenda of Educational Reform: A Challenge to Educational Self-Reliance and Dependency in Sub-Saharan Africa" *Journal of International Cooperation in Education*, Vol. 7, No. 1, pp. 85-96.

Lindahl, C. and Catterson, J. "What Causes Aid Dependency? Evaluating Thirty Years of Swedish Bilateral Aid to Tanzania" in Pitman, G. K., Feinstein, O. N. and Ingram, G. K. (eds.) *Evaluating Development Effectiveness*, World Bank Series on Evaluation and Development, Vol. 7, New Brunswick, N.J.: Transaction Publishers, pp. 133-151.

Minamoto, Y. and Nagao, M. (2006) "Process Considerations in Evaluating Educational Cooperation Projects," *Journal of International Cooperation in Education*, Vol. 9, No. 1, pp. 89-106.

Mpumalanga Department of Education (1999). Mpumalanga Secondary Science Initiative (MSSI). Project Document, Middelburg, South Africa.

Mpumalanga Secondary Science Initiative (MSSI) (2006) "*Joint Evaluation of the Mpumalanga Secondary Science Initiative (MSSI)*" Nelspruit, South Africa.

Nagao, M. (2001) "Purposes and Uses of Evaluation in Aid Projects: An Illustration by Japanese aid to mathematics and science education in South Africa," *Journal of International Cooperation in Education*, Vol. 4, No. 1, pp. 89-100.

Nagao, M. (2004) "Could Japan be a Good Math and Science Teacher for Africa", *Journal of International Cooperation in Education*, Vol.7, No1, pp53-70.

Patton, M. Q. (1997) *Utilization-Focused Evaluation*, The New Century Text, 3<sup>rd</sup> edition, Sage Publications.

Picciotto, R. (2003) "International Trend and Development Evaluation: The Need for Ideas" *American Journal of Evaluation*, Vol.24, No.3, pp. 227-234.

Platteau, J. P. (2005) "Institutional and Distributional Aspects of Sustainability in Community-Driven Development" in Pitman, G. K., Feinstein, O. N. and Ingram, G. K. (eds.) *Evaluating Development Effectiveness*, World Bank Series on Evaluation and Development, Vol. 7, New Brunswick, N.J.: Transaction Publishers, pp. 275-297.

Rawlings, L. (2005) "Operational Reflections on Evaluating Development programs" in Pitman, G. K., Feinstein, O. N. and Ingram, G. K. (eds.) *Evaluating Development Effectiveness*, World Bank Series on Evaluation and Development, Vol. 7, New Brunswick, N.J.: Transaction Publishers, pp. 193-204.

Stewart, F. (2005) "Evaluating Evaluation in a World of Multiple Goals, Interests, and Models" in Pitman, G. K., Feinstein, O. N. and Ingram, G. K. (eds.) (2005) *Evaluating Development Effectiveness*, World Bank Series on Evaluation and Development, Vol. 7, New Brunswick, N.J.: Transaction Publishers, pp. 3-29.

Svensson, K. (1997) "The Analysis and Evaluation of Foreign Aid" in Chelimsky, E. and Shadish, W. R. (eds.) *Evaluation for the 21<sup>st</sup> Century, A Handbook*, Thousand Oaks, CA: Sage Publication.

Weiss, C. H. (1998) *Evaluation*, 2<sup>nd</sup> Edition, Upper Saddle River, N.J.: Prentice Hall, Inc.

White, H. (2005) "Challenges in Evaluating Development Effectiveness" in Pitman, G. K., Feinstein, O. N. and Ingram, G. K. (eds.) *Evaluating Development Effectiveness*, World Bank Series on Evaluation and Development, Vol. 7, New Brunswick, N.J.: Transaction Publishers, pp. 33-54.

Wiesner, E. (1997) "Evaluation, Markets, and Institutions in the Reform Agenda of Developing Countries" in Chelimsky, Eleanor and Shadishi, William (Ed.) *Evaluation for the 21<sup>st</sup> Century: A Handbook*, Thousand Oaks: Sage Publication, pp. 189-200.