

Challenging times for evaluation of international development assistance

The spread of results-based management in developed countries, combined with increasing taxpayers' concern for non-performing projects, has increased the relevance of evaluation in development assistance. There are challenges, however, stemming from the 'more than usual' complexities of evaluation circumstances (e.g. asymmetry of interest in evaluation between an aid donor and a recipient) as well as the changing development context, such as increasingly strategic orientation of donor policies, growing weight of global development targets such as the Millennium Development Goals, and spread of the practice of multi-donor assistance. The paper reviews the theoretical and practical issues involved in coping with these challenges. It argues that the increasing emphasis placed on capacity development in development policy and programs, when coupled with empowerment approaches in evaluation, may give rise to a new and promising role for evaluation in development and development cooperation.

Masafumi Nagao

.....
Masafumi Nagao is a Research Professor at the Center for the Study of International Cooperation in Education (CICE) at Hiroshima University, Japan. Email: <nagaom@hiroshima-u.ac.jp>.

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 & 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 five 'DAC Evaluation Criteria': relevance, effectiveness, efficiency, impact and sustainability, which are now used by most aid agencies and international organizations for evaluating the impact of their assistance projects. The work done by this group was instrumental in propagating evaluation activities in developing countries. It also helped spread evaluation practice in some of the donor countries, such as Japan¹. This 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 practised for many years, a review of the project appraisal documents showed that the proportion of projects with planned impact evaluations in 2000 amounted only to 10% (Rawlings 2005).

It was only in the late 1990s with the spread of the practice of results-based management (RBM) in donor country governments that evaluation came to assume greater relevance in development assistance. Under pressure from deteriorating public finances and unfavorable economic conditions, donor governments subjected the programs and projects of their aid agencies to 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 (OECD/DAC 2000). RBM has also been adopted by multilateral organizations such as the World Bank and the United Nations Development Program (UNDP). Furthermore, through their aid programs for public sector reform and improved governance, the practice of public sector evaluation has begun to spread to developing countries (Weisner 1997; Guerrero 1999).

Even 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 concern that years of ongoing aid have not produced a visible and sustainable impact on development in many developing countries (Chapman & Nagao 2006). There is also a somber reminder that evaluation may not affect actual policies or behavior of the donor agency (Lindahl & Catterson 2005).

This paper reviews some key issues arising from these challenges, touching on practical as well as theoretical concerns. It argues that there exists certain contextual factors, some old and others new, that complicate this practice. However, the increasing trend towards local ownership of the development process and local capacity development, when coupled with empowerment approaches, may give rise to new and promising roles for evaluation in development and development cooperation.

This is illustrated by a mathematics and science education project in South Africa in which I participated as a leader of a Japanese technical assistance team. This project included a locally owned formative monitoring system that was instituted to guide 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 this from the usual stakeholder-based approach to evaluation suggested by Patton (1997) or the negotiation perspective offered by Owen (2006), 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 agendas (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, who must account for the development assistance funds to their government and ultimately, to their taxpayers. Donor agencies are pressed towards outcome-oriented planning and results-based evaluations of project effectiveness.

The recipient, 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 is likely to be a passive one. The participation of multilateral development agencies and non-governmental organizations as donors may modify this, but would not alter drastically the asymmetric character of the donor-recipient relationships regarding evaluation.

Difference in time horizon

Another clear sign of asymmetry is the difference in time horizon. Donors clearly have a finite time horizon, and a tendency to target tangible short-term results from aid efforts. Typically these efforts involve imposing of projects geared to quick transplanting of best practices from the donor countries. Evaluation needs only to trace their implementation using indicators associated with these practices. However, for developing countries development is not a time-bound phenomenon. They are obviously more interested in real developmental impact made possible by these practices. Depending on the learning requirements

or institutional changes involved, the realization of such impact may have to be sought on a different time scale (Ellerman 2004). Hence evaluation may need to be looked at in a totally different light.

Evaluation of development or impact of aid

The asymmetry of interest is also apparent in respect to what to evaluate. Both donor and recipient may look at the output, outcome and impact generated by the project. The donor's basic interest lies in ensuring the influence of their assistance projects in generating outputs and outcomes. They even try to set up these projects outside the regular development program of the recipient governments so as to facilitate the task of their monitoring and evaluation. The recipient's primary interest, by contrast, is not in assessing the impact of the 'add-on' element, but that of the developmental effort as a whole. How much contribution the donor assistance has made in bringing about that development is of secondary importance. As I indicated above, the recipients are also paying increasing attention to capacity building within the process of development, which has raised their interest in evaluation of project process (Minamoto & Nagao 2006). How to organize the controlling and monitoring of the development process is surfacing as a point of contention between the donor and the recipient.

Evaluator selection

According to traditional evaluation theory, evaluators should serve as a neutral and independent entity to stand between the funder of evaluation and the beneficiary, and to mitigate the negative impact of the asymmetry of interest between the two. This, however, is difficult in the case of evaluation of development assistance. Evaluation of aid projects is usually undertaken by consultants selected from the country of the donor. 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 about the choice of evaluators, or their terms of reference. Some donor agencies have hired local evaluators or evaluators from other developed countries, but this is still a rare occurrence. Whatever the location of consultants, 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 five 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). Of the five criteria, the most neglected criteria appears to be sustainability of impact after project termination (Brinkerhoff & 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 depends 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 that the recipient country should take to develop its own problem-solving capability, and the capacity to conduct evaluation as an important component of the latter (Mackay 2002).

Challenges from the changing development context

The foregoing discussion focused on the distinguishing features of development assistance evaluation which implies somewhat greater complexity, and perhaps a greater challenge, than the evaluation usually practised 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 the 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 intrastate conflicts has multiplied. The resulting situations, often involving political instability and armed struggles, have 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 the coherence of assistance policies of the parties involved and workability of coordination mechanisms (OECD/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

afterthought. 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 (OECD/DAC 1997), 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 timelines 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 has to be made against much uncertainty (White 2005). Achieving a measure of transparency and assuring accountability to the national audience for the critical need 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, 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 by developing countries and 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 whole-of-government sense. What is even more significant is that the key to the successful implementation is believed to be country ownership of the capacity development itself (OECD/DAC 2006).

The ownership focus in development assistance 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, and 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 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 which lead to an independent judgment of program worth.

Instead the techniques required here are those that 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 to project decision-making as it is rolled out (Patton 1997). Adoption of this approach, however, calls for leadership to assure 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 lead evaluator must often act as a trainer or a technical adviser.

Application of empowerment evaluation in development assistance settings faces additional difficulties. Above all the donor agency, even if they agree to adopt empowerment thinking in the project, are usually not inclined to assign an independent evaluator to facilitate the process for a project that may take place over several years and in a distant location. It is simply not in their mindset. What is required for those responsible for planning a long-term intervention is to formulate a project design that incorporates empowerment thinking into the structure of the project, by engaging the project participants in monitoring and continuous review of the project implementation process. This requires that donors appreciate the advantages that empowerment evaluation can make to effective implementation, the attainment of project outcomes as well as ensuring that the effects of the project will be sustained over the longer term.

In the next section, this is illustrated in terms of a secondary mathematics and science teacher retraining project in South Africa in which I participated for seven 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 Department 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 (see the project summary in Figure 1). The Department of Education 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 apartheid times, but also to facilitate the

introduction of a major curriculum reform that 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&S teachers.

MSSI had several distinct characteristics. It tried to combine different teacher training concerns into a single in-service training program. It targeted all 540 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. 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-1990s, 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 of Education 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.

FIGURE 1: MAIN ELEMENTS OF THE MSSI PROJECT

Goal	Improved M&S understanding of secondary students
Aims	1 Improvement in M&S teaching via teacher retraining 2 Development of a province-wide system of school-based in-service training (INSET)
Duration	November 1999 – March 2006
Partners	Mpumalanga Department of Education/Japan International Cooperation Agency (JICA)/University of Pretoria (UP)
Target population	Mathematics and science teachers in all 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

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 effort was a new model of technical cooperation termed 'experience-sharing' model (Nagao 2004), which contrasted with the usual technology transfer model of technical cooperation. The object of the cooperation activity in this model is not transfer of a technology but 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.

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, including 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 MSSSI 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.

Use of an 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 2, the project's intervention was organized as a cascade chain of training activities—starting with (1) a study-cum-training mission to Japan for a teacher trainer group, which prepared an annual training plan and study guides on the basis of material development

FIGURE 2: MONITORING OF CASCADE SYSTEM OF TRAINING

Intervention point	Cascade training	Monitoring instrument
<i>Training in Japan</i>	Teacher trainer group mission (6 weeks/year) Output: annual training plan and study guides	Review of training program
<i>Province level</i>	Teacher trainer workshops (3 days x 3 times/year) Output: district workshop plan	Bi-annual reports by MSSSI Coordinators for the Steering Committee
<i>District level</i>	Subject head teacher workshops (3 days x 3 times/year) Output: school-based INSET plan	Quarterly reports by teacher trainers
<i>School level</i>	School-based INSET activities (once a month or more) Output: learning from INSET	Quarterly reports by subject head teachers

work, (2) 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, (3) district-level workshops for M&S subject head teachers from the participating schools, and (4) 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 that 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-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 collates 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.

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. The 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, a question remained about whether the positive practices initiated could be sustained. The project initially envisaged combining the systemic effort for promoting cascade training activities and accompanying monitoring with the provision of 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 states:

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 closely with 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 organized 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 public 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 donor and recipient, and different orientation to development as reflected in objectives and time horizons.

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 the Millennium Development Goals and the need to align the evaluation framework to macro-level goals and policies, and the emergent 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 international development assistance, this paper has pointed out that developing countries and the donor community alike are paying increasingly serious attention to capacity development and other self-reliant approaches. Moreover, there is an emerging consensus that successful implementation lies in 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 and participatory thinking, which gives greater control about the planning, implementation and evaluation of development undertakings to the developing countries.

To illustrate the application of these approaches to a concrete development project, this paper has described a mathematics and science teacher training project in South Africa in which I 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 learning by those on 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.

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), '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.

Note

- 1 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.

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.
- Brinkerhoff, D & Goldsmith, AA 1992, 'Promoting the sustainability of development institutions: a framework for strategy', *World Development*, vol. 20, no. 3, pp. 369–383.
- Canadian International Development Agency (CIDA) 2002, 'Assessing sustainability', *What we're learning no. 2*, Canadian International Development Agency, Ottawa.
- Chapman, D & 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.
- Dabelstein, N & Rebien, CC 2002, 'Evaluation of development assistance: its start, progress, and current challenges', in J-E Furubo, RC Rist & R Sandahl (eds), *International atlas of evaluation*, Transaction Publishers, New Brunswick, New Jersey, 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 DL Stufflebeam et al. (eds), *Evaluation models: viewpoints on educational and human services evaluation*, 2nd edn, Kluwer Academic Publishers, Boston, pp. 395–408.
- Furubo, J-E, Rist, RC & Sandahl, R (eds) 2002, *International atlas of evaluation*, Transaction Publishers, New Brunswick, New Jersey.
- Garaway, Gila 2003, 'Evaluating educational programs and projects in the Third World', in T Kellaghan & DL Stufflebeam (eds), *International handbook of educational evaluation*, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 701–720.
- Guerrero, RP 1999, 'Evaluation capacity development in developing countries: applying the lessons from experience', in R Boyle & D Lemaire (eds), *Building effective evaluation capacity*, Transaction Publishers, New Brunswick, New Jersey, pp. 177–194.
- 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 & Catterson, J 2005, 'What causes aid dependency? Evaluating thirty years of Swedish bilateral aid to Tanzania', in GK Pitman, ON Feinstein & GK Ingram (eds), *Evaluating development effectiveness*, World Bank series on evaluation and development, vol. 7, Transaction Publishers, New Brunswick, New Jersey, pp. 133–151.
- Mackay, K 2002, 'The World Bank's evaluation capacity building experience', in DW Compton, M Baizerman, & S Hueftle Stockdill (eds), *The art, craft, and science of evaluation capacity building, New directions for evaluation*, no. 93, Spring, pp. 81–100.
- Minamoto, Y & 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, no. 1, pp. 53–70.
- OECD/DAC 1997, *Development Cooperation 1996 Report*, OECD, Paris.
- OECD/DAC 1999, *Guidance for evaluating humanitarian assistance in complex emergencies*, Working Party on Aid Evaluation, OECD, Paris.
- OECD/DAC 2000, *Results based management in the development co-operation agencies: a review of experience—executive summary*, Working Party on Aid Evaluation, OECD, Paris.
- OECD/DAC 2006, *The challenge of capacity development: working towards good practice*, DAC Network on Governance, OECD, Paris.
- Owen, J 2006, *Program evaluation: forms and approaches*, Allen & Unwin, Crows Nest, NSW.
- Patton, MQ 1982, *Practical evaluation*, Sage Publications, Newbury Park, California.
- Patton, MQ 1997, *Utilization-focused evaluation: the new century text*, 3rd edn, Sage Publications, Thousand Oaks, California.
- Picciotto, R 2003, 'International trend and development evaluation: the need for ideas', *American Journal of Evaluation*, vol. 24, no.3, pp. 227–234.
- Rawlings, L 2005, 'Operational reflections on evaluating development programs', in GK Pitman, ON Feinstein & GK Ingram (eds), *Evaluating development effectiveness*, World Bank series on evaluation and development, vol. 7, Transaction Publishers, New Brunswick, New Jersey, pp. 193–204.
- Stewart, F 2005, 'Evaluating evaluation in a world of multiple goals, interests, and models', in GK Pitman, ON Feinstein, & GK Ingram (eds), *Evaluating development effectiveness*, World Bank series on evaluation and development, vol. 7, Transaction Publishers, New Brunswick, New Jersey, pp. 3–29.
- Svensson, K 1997, 'The analysis and evaluation of foreign aid', in E Chelimsky & WR Shadish (eds), *Evaluation for the 21st century: a handbook*, Sage Publications, Thousand Oaks, California.
- White, H 2005, 'Challenges in evaluating development effectiveness', in GK Pitman, ON Feinstein & GK Ingram (eds), *Evaluating development effectiveness*, World Bank series on evaluation and development, vol. 7, Transaction Publishers, New Brunswick, New Jersey, pp. 33–54.
- Wiesner, E 1997, 'Evaluation, markets, and institutions in the reform agenda of developing countries', in E Chelimsky & W Shadishi (eds), *Evaluation for the 21st century: a handbook*, Sage Publications, Thousand Oaks, California, pp. 189–200.