

PROJECT MONITORING AND EVALUATION- ISSUES AND CHALLENGES

Abstract

As our country struggles to enhance the development results at budgeting and capital projects, new challenges arise for the practice of monitoring and evaluation (M&E). This paper outlines issues and challenges in M&E based on UNDP's experience. Associated with a shift from a culture of control to delegation, the focus has moved from monitoring inputs and outputs to assessing outcomes within a results-based framework. M&E ensures UNDP's accountability for resource use and achieving objectives. The results are evaluated for their contribution to the partner country's development effectiveness.

M&E also facilitates learning and knowledge generation through the analysis and objective feedback of lessons from development experience. Challenges include assessing the contribution of individual projects to broader development results. Methodological rigour is required for setting baselines and measuring the impacts of soft assistance, such as policy advice and capacity development. It is important to promote national ownership and capacity development in M&E so as to ensure that it contributes to national development goals. The utility of M&E should be enhanced to support programming and implementation. Involvement of broad categories of in-country stakeholders in M&E is a priority. The UN system must enhance partnerships for a systematic assessment of impacts on poverty eradication, economic growth and sustainable development in the partner countries.

Key Words are: Monitoring, Evaluation, Challenges, UNDP's, and Delegation

Introduction

In recent years the importance of monitoring and evaluation (M&E) has been increasingly realised in development cooperation. This trend has coincided with the recognition that much of aid has been ineffective. In many cases, development projects have done what they were set out to do, but with little or no impact on the bigger picture. This situation could be seen as a parallel to the old joke about the operation having been successful but the patient dying: the project was a success but no improvement in the country's development situation could be detected. Partly, this case may be a result of the fact that projects as isolated interventions have not been significant enough. Another part of the reason may be that they have not been sufficiently tied to

a country's strategies and policies. In both cases, sustainability of results has been lacking. These unfortunate realisations have put the onus on the development partners to demonstrate that the cooperation does indeed make a difference. Consequently, there has been a marked shift in M&E from monitoring inputs and outputs of individual interventions, to assessing the results and impacts of projects and programmes, and their contributions to larger outcomes (UNDP 2005).

This paper makes an attempt to elaborate on the challenges associated with these changes. The arguments are based on experiences from the United Nations Development Programme (UNDP) and the factors that have led to the development of a new evaluation policy for the organisation. The first part of the paper briefly outlines the purpose and context of M&E as they pertain to UNDP. Following that, I provide a discussion of project-level M&E and how it contributes to results-based management. I then move on to the linkages of project-level M&E to higher level assessment of development results. In the final concluding part of the paper I outline some directions for the future in the context of assessing the development results at country level.

Monitoring and evaluation context in UNDP

Although monitoring and evaluation are frequently presented as a dual concept, they are distinct functions and serve different purposes. Monitoring is defined as a continuous function that provides managers and stakeholders with regular feedback on programme performance taking into account the external environment. It provides an early indication of progress or lack thereof in the achievement of intended results (UNDP 2006).

Evaluation, on the other hand, provides a judgment based on assessments of relevance, appropriateness, effectiveness, efficiency, impact and sustainability of development efforts. It involves a rigorous, systematic and objective process in the design, analysis and interpretation of information to answer specific questions. It highlights both intended and unintended results, and provides strategic lessons on what works and why to guide decision-makers and inform stakeholders (ibid.). Monitoring can provide critical inputs to evaluation by way of systematic collection of data and information.

There are a number of major developments that have recently influenced M&E in UNDP and the larger UN family. One would be the move towards UN system harmonisation, which will require greater cooperation and alignment of various UN agency programmes at the country level. This will also pose challenges for evaluation, which will need to provide a more holistic analysis of UN support to the partner country.

Secondly, there is a drive to strengthen and professionalise the evaluation function in the UN system. In 2005, the United Nations Evaluation Group (UNEG) adopted a comprehensive set of norms and standards for evaluation in the UN, which seek to facilitate system-wide cooperation in evaluation (UNEG 2005a & b). The norms are intended to ensure that all UN evaluation units follow agreed basic principles (Box 1).

Box 1: Evaluation norms for the UN system

- Impartiality:** Absence of bias in due process, methodological rigour, consideration and presentation of achievements and challenges.
- Independence:** Pertains both to the location of the evaluation function independent from management, and to the independence of the evaluators.
- Evaluability:** Improving the ability to evaluate by building an evaluation plan during the planning stages.
- Quality:** Employ quality-oriented design, planning and implementation processes; evaluation reports must present in a complete and balanced way the evidence, findings, conclusions and recommendations.
- Competencies for evaluation:** The required evaluation skills must be present in the organisations.
- Transparency and consultation:** Involving major stakeholders at all stages of evaluation improves credibility and quality, facilitates consensus building and ownership of findings, conclusions and recommendations.
- Ethics:** Evaluators must have personal and professional integrity, respect the rights of individuals and institutions to provide information in confidence.
- Follow-up:** There should be systematic implementation of evaluation recommendations once these have been agreed to by the management and/or governing bodies.
- Contribution to knowledge building:** Evaluation findings and lessons should be available to target audiences in a user-friendly format.

(Adapted from UNEG 2005a)

Within this context, UNDP has revised its evaluation policy and sees evaluation as essential to ensure the organisation's accountability for the achievement of its objectives through assessing its programmes and operations. The new evaluation policy, still in a draft form, calls for the evaluation of UNDP's results for their contribution to the development effectiveness of partner countries. Evaluation is intended to support management for results and must be responsive to national priorities and concerns. Evaluation should strengthen partnerships with governments and key stakeholder groups in the country.

Another important role of evaluation is facilitating learning and improving knowledge, thus enhancing performance by providing objective feedback on lessons learnt from development experience. It should also support national capacities to manage, implement, monitor and evaluate of development programmes and projects.

Monitoring and evaluation at the project level

There are numerous modalities through which UNDP works with partner countries. Projects have in the past been a central modality and they continue to play an important role in delivering development results. However, there are also other modes of cooperation that may take different shapes, including advocacy, advisory services, knowledge networks, technical assistance, coordination and partnerships. Monitoring and evaluating the results of these various types of activities and modalities requires varying approaches.

Project development, management, implementation and M&E in UNDP have been mostly decentralised to the country level. The Country Offices (COs) are at the forefront working closely with the partner government and other stakeholders. Project-level M&E is intended to contribute towards results-based management by tracking that projects are on track in delivering the services and outputs in a timely and efficient manner, and that they are likely to achieve their set objectives. The CO results-based management systems are largely based on self-assessment of progress and results achieved, building upon M&E and reporting from individual operations and aggregated along UNDP's practice areas and service lines on an annual basis to provide an overall picture of the outcomes of the country programme.

Project-level evaluation has also been decentralised and is the responsibility of the COs. While the evaluation policy encourages evaluation of projects as an important tool for assessing results and impacts, and identifying lessons to improve performance, project evaluations are normally not mandatory. The justification for this is that there is a great range of sizes and types of projects and it would not be meaningful or cost-effective to arrange separate evaluations for all of them. Many projects can only be expected to contribute to broader goals that only can be achieved through a wider set of interventions.

In some cases, a partnership protocol calls for mandatory project evaluation. A case in point would be the Global Environment Facility (GEF), which requires that all full and medium-sized projects it funds must be evaluated through an independent mechanism. In the case of GEF-funded projects implemented by UNDP, the evaluation responsibility still rests with the CO, which must ensure that the projects are evaluated according to established norms and standards. A key requirement is independence of the evaluation, which is interpreted so as that the evaluation must be conducted by a team of consultants who have not been associated with the development or implementation of the project in question. Good practice would suggest that the consultant team should contain both international and local expertise to ascertain a proper balance of perspectives. The team should also combine evaluation expertise in addition to substance matter specialisation.

It is easy to look at projects as discreet interventions and evaluate them as such. In the past, project evaluation has largely focused on whether a project carries out the activities it was set out to; does this on time and within budget; and delivers the outputs that were predefined in the project document. Less often have evaluations actually looked at the results and impacts of the project. Even less often do they go beyond the immediate results of the project itself. However, it would appear necessary to do so in order to evaluate the relevance of the project and its long-term impacts. For example the GEF evaluation guidance goes beyond accounting for the outputs or immediate results of the project. It is required that evaluations also cover other, higher order dimensions, such as whether the project contributed to the achievement of global environmental benefits (the very justification for GEF funding!); whether the outcomes of the project are sustainable after the external funding and project interventions are completed; and whether the project approaches and results can be replicated elsewhere, so as to ensure broader impacts.

Much effort has been invested in the development of tracking tools and indicators that would form an objective basis for assessing the performance and results of projects within the GEF context. All GEF focal areas have made considerable progress in developing project- and programme-level indicators that allow for the tracking of progress and for aggregation of results beyond individual projects. For example, in the International Waters programme of the GEF, indicators are tracked at three levels: processes, stress reduction, and environmental status (Box 2).

Box 2: Types of indicators for GEF international waters projects

- Process indicators: Track the implementation of the agreed processes that are intended to lead to the desired outcomes (examples would include the establishment of policies, institutional arrangements or legal frameworks necessary for the achievement of the project's objectives).
- Stress reduction indicators: Focus on the actual implementation of measures that will reduce the environmental stress on the particular water body in question (examples would include the instalment and operation of a sewage treatment system for an industrial or urban centre).
- Environmental status indicators: Measure the actual changes in environmental quality in the water body.

(See Uitto 2004 for more details)

In this particular case, the actual improvement in the environment is the goal of the intervention which justified the existence of the project. However, the achievement of such positive changes requires a long time and often the interaction of several project and non-project interventions. Therefore, these impacts may only materialise fully long after the project is gone; hence the need for intermediate measures of the intervention's success. Naturally, this is not unique to environmental interventions but would pertain to any complex development efforts.

What thus is needed is a theoretical framework that sets out the logical connections between specific actions and the higher level outcomes. Theory-based evaluations attempt to do this through using programme theory or programme logic that articulates the intervention logic and the assumptions that are necessary for the achievement of the higher objectives (Davidson, 2000). Theory-based evaluation is particularly useful in the context of complex projects and programmes where the causal linkages and intervening factors are not direct or simple. In such cases, theory-based evaluation can help to detect both the desired and unanticipated outcomes of projects or programmes and also to identify what works and what doesn't (Rogers et al., 2000; Weiss, 2000).

In order to assess the overall results of its project and non-project interventions at the country-level, UNDP undertakes outcome evaluations that seek to determine the contributions towards desired development results. Outcome evaluations normally cover the cluster of projects and interventions in each of UNDP's programmatic areas in the respective programmatic framework. They may also cover a particular modality or crosscutting theme. They are intended to reinforce the accountability of operational units for development results. Outcome evaluations that at the country level are organised by the CO should also feed into independent programme evaluations undertaken by the Evaluation Office (EO).

The real challenge for evaluation is how to assess the contributions of a large number of disparate projects, activities and programmes to the larger issue of development at the country level

Assessing development results

Development must be based on the aspirations of the people concerned. Externally funded interventions are intended to support the overall development goals of the country where they take place. At the same time, these must be placed into the context of internationally agreed principles, such as the Millennium Development Goals (MDGs). UNDP has formulated its contribution to the development results in terms of poverty reduction, economic growth, and sustainable development. At the same time, UNDP advocates a people-centred approach to development that enhances the capabilities, choices and rights for all men and women. Ultimately, the success of UNDP's programmes must be evaluated against these principles: i.e., whether UNDP has been successful in supporting the recipient country's development goals and whether it has been able to incorporate its own guiding principles into its programming.

An evaluation tool to address the issues outlined above is the Assessment of Development Results (ADR) managed by the EO. ADR's are country-level evaluations that are intended to address the issue of UNDP's contributions to the country's development results. Their scope normally includes UNDP's responsiveness and alignment to country challenges and priorities, strategic positioning, use of comparative advantage, and engagement with partners. These evaluations focus on the outcomes and critically examine achievements and constraints in the UNDP thematic areas of focus, draw lessons learnt, and provide recommendations for the future. ADRs also provide a strategic analysis for enhancing performance and positioning UNDP support within national development priorities and UNDP corporate policy directions. ADRs thus look at the totality of UNDP's programme in the country over several years (Box 3).

Box 3: Scope of ADRs

The purpose of the evaluation is to assess UNDP's contributions to development results and strategic positioning in the country, draw lessons learned and outline options for improvements. Specifically the ADR will:

- Provide an overall assessment of the results achieved through UNDP support and in partnership with other key development actors during the last five to seven years.
- Provide an analysis of how UNDP has positioned itself strategically to add value in response to national needs and changes in the national development context.
- Based on the analysis of achievements and positioning above, present key findings, draw key lessons, and provide a set of clear and forward-looking options for the management to make the necessary adjustments in the current strategy and next Country Programme applied by UNDP and partners towards intended results.

This modality of evaluation was commenced in 2000 and thus for fifteen ADRs have been completed and a number are underway. There are, therefore, already a number of conclusions and lessons emerging from experience. A basic constraint in evaluations of this nature would relate to attributing the country-level changes to particular assistance provided by the external agency. The purpose is not only to evaluate UNDP's activities. The external agencies are there only to support the development aspirations of the country in question. Therefore, the evaluation questions should relate to whether the agency has had a plausible contribution to the development results of the country.

With that in mind, it is important to gauge the relevance and success of the contributions of that particular external agency. This is important primarily for the accountability purposes viz. the partner country, and also for the accountability viz. the governing bodies of the agency. There are, however, a number of methodological challenges if we want to move beyond mere impressionistic or politically correct assessment.

A major one is how does one assess the impacts of soft assistance, which is often the main mode of operation of UNDP. This refers to modalities such as policy dialogue and advice, and capacity development. In many countries, these are the most common modalities employed by UNDP in its role to support the government in achieving the MDGs. Undoubtedly, UNDP plays a central role in many countries in supporting the government in devising and implementing policies that are conducive to development. While these may be results at least partly attributable to UNDP they are primarily the results of the efforts of the government and other partners in that particular country. Consequently, methodological rigour is needed in order to isolate the impacts that would be attributable to specific actions. Oftentimes, direct attribution will not be possible. In that case, it may be more relevant to identify what are the contributions of an activity or an agency to a specific outcome. Detecting the plausible contributions requires establishing a logical link between the two.

Another methodological challenge faced by higher-order evaluations, such as the ADRs, pertains to the availability of data and evaluative evidence. In order to measure the impacts of a programme, it would be important to have baseline information against which change could be measured. Unfortunately, such baselines frequently do not exist. Similarly, carrying out contrafactual analysis – what would the outcome have been without the interventions – is often methodologically impossible.

Programmatic and country-level evaluations should be able to build upon earlier evaluations of projects and, in the case of UNDP, the sectoral or thematic outcome evaluations. Outcome evaluations are built into UNDP's programming, so they would normally be available as building blocks for ADRs. As project evaluations are as a rule not mandatory – except in specific cases, such as the GEF-funded projects – their availability tends to be patchier and those that are available are not always high-quality independent evaluations.

As a conclusion, it is important to ensure early on that baseline data is being collected in the programme areas, and that evaluations at outcome and project level are encouraged. This may require some quality assurance support from the EO. However, building the data and evidence base for the evaluation to come is crucial.

Promoting use of evaluation

Although it may sound like a truism, it is worth noting that more emphasis and importance will be given to evaluation by various stakeholder groups if the evaluations are useful for their purposes. Promoting use of evaluations is a key feature of UNDP's evaluation policy. Evaluations are not carried out for their own sake, but as tools to promote development results. They are intended to serve both accountability and learning purposes. These purposes are not mutually exclusive, as both must be based on an analysis of the results achieved and an understanding of why they were or were not achieved. There can only be learning through a systematic study of what worked, what didn't, and why.

As their purpose is to enhance development results, evaluations should be useful primarily to the partner countries, as well as for UNDP to improve its operations. Additionally, evaluation enhances UNDP's accountability for its contribution to such development results. The uses of evaluation should be made clear from the outset when they are commenced so as to enhance ownership of the evaluation process. Similarly, those responsible for the functions or programmes which have been evaluated should systematically implement evaluation recommendations when these have been agreed to by management.

The way forward

Benchmarking evaluations firmly against national priorities will enhance national ownership and utility of the evaluations. It will then become clearer that evaluations are not there to satisfy procedural needs of individual projects, but are essential management tools for ensuring the efficiency and effectiveness of programmes and projects to support the government in achieving its development goals.

It is also important to involve broad stakeholder groups in the countries where evaluations are carried out. Different groups have different expectations and experiences on the programmes and projects that are subject to evaluation. It is important to capture their perspectives to understand the distribution of benefits and to identify any unintended outcomes, positive or negative.

It is also important to enhance the linkage between evaluation and programming. M&E is an integral part of programme and project implementation and adaptive management. Similarly, evaluation should effectively feed into new programming decisions. In the case of UNDP, attention is paid to timing the ADRs so that they can provide inputs to the development of the next country programme. The ADRs analyse the strategic positioning of UNDP vis-à-vis the partner country's development priorities, the results of the programmes, and the strengths and weaknesses of the approach. These analyses and resulting recommendations can thus be used to strengthen the focus and delivery of the next country programme.

As the UN system moves further towards harmonisation and alignment in its country-level programmes, joint evaluations between agencies will increase in importance. Four key purposes for such evaluations have been identified (Box 4). In fact, such joint country-level evaluations will be piloted in 2006 under the aegis of UNEG. Partnerships in evaluation need to be promoted so as to assess the contributions of various agency programmes to the partner countries' development results. Eventually, the goal should be that the partner countries themselves take the lead in initiating and undertaking such evaluations in cooperation with the UN.

Box 4: Purpose of UN country-level evaluations

- To provide developing countries, UN organisations and other development partners with objective and transparent information on the effectiveness of the UN system's developmental, humanitarian and normative activities at the country level.
- To assess the contribution of UN organisations to achieving results in developing countries and to identify the extent to which these results stem from developmental, humanitarian and normative activities.
- To examine the ways in which UN agencies and their partners programme effectively (or not) in situations of transition from humanitarian relief to development.
- To investigate themes and issues that are the focus of collaborative or joint programming by different UN agencies working at country level with development partners.

(Freeman 2005)

Conclusion

Great strides have been made in the past years in advancing results-based management and an evaluation culture. There has been a marked shift in emphasis from monitoring processes towards evaluating results and impacts of development programmes and projects. Evaluation has also moved from being primarily a donor-driven activity to become a management and accountability tool for the partner country. This trend needs to be further strengthened. Evaluation should be a priority for the partner country and there is a need to promote national ownership of evaluation. An important part of this would be to link evaluation more strongly to country-level perspectives on development results and outcomes.

References

- Davidson, E.J. (2000). Ascertaining causality in theory-based evaluation. In: Rogers, P.J., Hacsí, T.A., Petrosino, A. and Huebner, T.A. (eds.), *Program Theory in Evaluation: Challenges and Opportunities*. New Directions for Evaluation No. 87, pp. 17-26.
- Freeman, T. (2005). *Issues and Options Paper on UN Country Level Evaluations*. UNEG Working Group on Country-level Evaluations: New York.
- Rogers, P.J., Petrosino, A., Huebner, T.A. and Hacsí, T.A.(2000). Program theory evaluation: Practice, promise and problems. In: Rogers, P.J., Hacsí, T.A., Petrosino, A. and Huebner, T.A. (eds.), *Program Theory in Evaluation: Challenges and Opportunities*. New Directions for Evaluation No. 87, pp. 5-13.
- Uitto, J.I. (2004). Multi-country cooperation around shared waters: The role of monitoring and evaluation. *Global Environmental Change* 14: 5-14.
- UNDP (2002). *Handbook on Monitoring and Evaluation for Results*. Evaluation Office, United Nations Development Programme: New York.
- UNDP (2005). *Monitoring & Evaluation Workshop for Country Offices in Asia & Pacific Region*. Workshop Report. Evaluation Office, United Nations Development Programme: New York.
- UNDP (2006). *Evaluation Policy Statement (Draft)*. United Nations Development Programme: New York.
- UNEG (2005a). *Norms for Evaluation in the UN System*. United Nations Evaluation Group.
- UNEG (2005b). *Standards for Evaluation in the UN System*. United Nations Evaluation Group.
- Weiss, C.H. (2000). Theory-based evaluation: Theories of change for poverty reduction programs. In: Feinstein, O. and Picciotto, R. (eds.), *Evaluation and Poverty Reduction*. Operations Evaluation Department, World Bank: Washington, DC.

