Credibility and Independence of Evaluation in UNDP: A Response to David Lempert

Juha I. Uitto

Deputy Director, Evaluation Office, United Nations Development Programme

Why government and non-governmental policies and projects fail despite 'evaluations,' asks David H. Lempert in the *Journal of MultiDisciplinary Evaluation* (2010) and then puts forth a stinging critique of evaluation practice in public organizations. While Mr. Lempert has many good insights and makes a number of interesting and valid points, there seems to be some fundamental misunderstandings about what independence means in evaluation and the role of evaluation in oversight.

As he uses the United Nations Development Programme (UNDP) evaluation function as a case example, I feel compelled to respond to some of the points, which indicate a lack of understanding of the evaluation architecture in the organization. My purpose is not to defend UNDP, which indeed does face a number of challenges related to monitoring and evaluation. Rather, I will focus on clarifying some misconceptions about the organization of the evaluation function in UNDP, while making some broader points about evaluation in public sector organizations.

Lempert's article contains a number of valuable observations regarding the prerequisites of effective evaluation. These start from the design of programs and results frameworks that either are or are not evaluable. When goals, outcomes and outputs have not been clearly defined or when data and information on progress are not collected, evaluation is difficult, to say the least. On the other side there are the internal requirements for an effective evaluation function. Lempert's indicators overlap to a large extent with the norms of the United Nations Evaluation Group (UNEG), a professional network of evaluation units in the UN system, which include independence, intentionality, transparency, ethics, impartiality, quality, timeliness and utility.[1] The UNDP evaluation policy builds upon these principles as well.

Mr. Lempert's premise is, rightly, that the performance of organizations should ultimately be judged by how well their programs, products and services serve the people who are intended to benefit from them. I do however disagree with his basic assumption that public organizations (and

---

[1] http://www.unevaluation.org

indeed NGOs and even private foundations that are not based on a profit motive) behave somehow less accountably than private companies. Competitiveness of private firms is ultimately tested by their profits and market shares, but this can hardly be compared with evaluation, *per se*, as Lempert seems to suggest. As he himself points out (p. 91), private firms judge their own success only by the above criteria, while their entire business model may be mostly detrimental to human wellbeing. Such social concerns—or "external harms" (p. 91) they may cause to public interest—do not often feature in their 'evaluations.'

I fully agree that UN agencies should be evaluated against their performance in promoting development that underlies their policies and operations in countries that have ratified relevant international conventions and agreements. Again, the ultimate stakeholders are the citizens of the developing countries, which UNDP and other UN agencies are set out to support. Therefore, such 'downward accountability' to the intended beneficiaries must form the basis for evaluation. The proof of the pudding must be whether the lives of the people are improved and whether the work of the UN agencies partnering with their governments and others has contributed to such improvements.

As an intergovernmental organization set up under the Economic and Social Commission of the UN, UNDP works primarily through governments and its Executive Board consists of government representatives from both donor and program countries. The direct accountability of UNDP for achieving results is in the first instance to the Executive Board, as its governing body, but UNDP is also accountable to the program country governments and citizens where the organization works. And as a public organization, it also has a broader accountability to the public. Having stated this ideal situation, I recognize the specific challenges that face international organizations that are operating somewhat removed from direct control by any single democratic system.

For this purpose, evaluation serves a particularly important role. In UNDP, the Executive Board has significantly strengthened its oversight role and supported the strengthening of the evaluation function over the past few years. The evaluation policy approved by the Executive Board in 2006, which Lempert scrutinizes in his paper, was an important step into this direction. In 2009, there was an independent review of the policy the findings of which were presented to the Executive Board in June 2010 (Gariba et al., 2010). The review found that while the policy by and large was relevant, its implementation left room for improvement.

The UNDP evaluation policy (UNDP, 2006), defines the organization's evaluation function as follows:[2]

> Evaluation in UNDP will provide an objective assessment of contributions to development results, through assessing its programmes and operations, including advocacy, advisory services, knowledge networks, technical assistance, coordination and partnerships. Evaluation will address what works and why, as well as what does not work and unintended outcomes. This will support accountability, inform decision-making and allow UNDP to better manage for development results.

Lempert criticizes the policy for not providing a definition of the mission of

---

[2] All UNDP evaluation documents are publicly available at http://www.undp.org/evaluation.

the Evaluation Office in terms of sustainable human development as defined in treaties, such as the UN Conference on Environment and Development or the 1992 Rio Declaration (p. 98). This is strictly correct, especially as the two instances mentioned above do not form the basis of UNDP's mandate. These conventions evolve over time and are part of the normative architecture of the UN. Therefore, the principles embodied in organizational structures, like that of UNDP, are more fundamental. However, we do take specific conventions into account in relevant evaluations. For example, the recent evaluation of UNDP's role and contributions in environment and energy did refer to the above mentioned environmental agreements (EO, 2008).

Furthermore, the evaluation policy is not intended as a standalone document and is complementary to other UNDP policy statements. At the same time, the evaluation policy does outline the principles that underlie evaluation in the organization, namely human development and human rights, UN system coordination, national ownership, and managing for results (UNDP, 2006).

I fully agree with Lempert that protecting the professionalism and objectivity of the evaluators is essential (p. 71). He goes on to identify factors that are important in this respect, including looking at the evaluator's work in its entirety and having safeguards against undue interference at every step from hiring to firing. This is indeed valuable insight.

It is worth noting that the evaluation architecture in UNDP consists of two distinct and complementary parts, a distinction that appears to elude Mr. Lempert. On the one hand, there is the independent evaluation function as practiced by the central Evaluation Office reporting directly to the Executive Board. On the other, in the highly decentralized organization operating on the ground in 166 countries, the evaluation function has also been decentralized. Most of the evaluations are commissioned by the country offices and other program units in the organization. Like UNDP, many development organizations (including other UN agencies, like UNICEF, as well as bilateral donors such as USAID and DFID) that have country-level presence rely on decentralized evaluations (some call them self-evaluations). Their role is primarily not for accountability, but helping program managers and decision-makers on the ground to manage better for development results. Such decentralized evaluations also form building blocks for independent evaluations that have a greater emphasis on accountability.

While these evaluations are carried out by outside consultants, they are not considered to be independent because they are indeed commissioned by the very units whose performance they are evaluating. It is my understanding that Mr. Lempert himself has in the past undertaken such consultancy assignments for certain UNDP country offices. Frustrations caused by them may in part explain the livid tone in his article. Such frustrations are understandable, as the decentralized evaluation system still leaves much to be desired, as has been identified in both the independent review mentioned above as well as analyses carried out by the Evaluation Office (Gariba et al., 2010; EO, 2010). Both the coverage as well as the quality of the decentralized evaluations has been found to be deficient. Both UNDP management and the Executive Board are well aware of these problems and steps are being taken

to enhance the compliance with evaluation policy and to strengthen the decentralized evaluation system. While the decentralized evaluations will not become independent evaluations, they can still play an important role in program oversight and, in particular, improvement of programming and implementation by feeding in lessons about what has worked and what has not.

As for the independent evaluation function, this rests with the Evaluation Office whose independence has been established through both structural and administrative processes. However, it is worth noting that the above mentioned independent review (Gariba et al, 2010) highlighted the same observation as Lempert (p. 98) that the evaluation policy is not explicit in protecting the independence of the Evaluation Office. This omission will be rectified in a revision of the evaluation policy that will be presented to the Executive Board in January 2011.

The director of the Evaluation Office is appointed by UNDP's administrator in consultation with the Executive Board and her/his service is limited to a maximum of two four-year terms after which she/he cannot return to any other function in UNDP. The director is in charge of recruitment of staff into the office. She/he also has full authority to propose a program of work for independent evaluation to the Executive Board for approval, reflecting both the need for accountability for results and ensuring that the organization is strategically positioned vis-à-vis the defined development priorities, its own comparative advantages, and other partners. The director also decides on the scope, conduct and disclosure of evaluations. The evaluations submitted to the Executive Board are considered

reports by the Evaluation Office, not of individual consultants. For these reasons, the Evaluation Office is responsible for the quality of the evaluations.

While the Evaluation Office director and staff are UNDP employees and located inside of the organization, it is a complete misunderstanding that the lack of independence of the evaluation function is proven by the fact that the reports produced by the Evaluation Office go out under the signature of the director and that each of the evaluations lists the name of the task manager from the office. It is a misunderstanding, because the office itself is an independent unit solely responsible for the quality and credibility of the evaluations. None of the reports are submitted to UNDP management for vetting or clearance before they are made available to the Executive Board and to the general public.

Of course, one could argue, as Lempert seems to, that evaluations should be carried out by complete outsiders who have nothing to do with the organization, program, policy or whatever the evaluand may be. However, evaluators too far removed from the field may not have the needed knowledge and insights into the area, which would make their learning curve very steep. It is furthermore difficult to understand who would commission such evaluations and how their results would be fed back into organizational decision-making. If the focus is on utilization of evaluation, this would be counterproductive. Naturally, outside researchers, journalists and others are free to study, say, UNDP programs and feed their findings back to their own audiences, but these would not be considered evaluations.

More important than formal independence, I would argue, is the credibility of the evaluation. Such

credibility is based several factors of which independence is but one. Others include the quantity and quality of evaluative evidence, the methodological rigor of the evaluation and, importantly, the professional skills of the evaluators. Like Lempert indicates (p. 74), the objects of evaluation in international development are complex and involve intricate linkages between multiple factors and sectors. Therefore they are usually beyond single individuals and require teams that encompass various types of expertise. Normally, UNDP Evaluation Office evaluations are conducted by teams consisting of four to up to a dozen staff and consultants with necessary expertise, not only in the relevant subject matter(s) but also in evaluation. Also, for the sake of objectivity, in selecting our teams we vet the consultants carefully to ensure that hey have no conflict of interest, through having been involved in the development or implementation of the initiatives that will be evaluated or expectations of employment with the unit that is being evaluated.

From a methodological point of view, the Evaluation Office has constantly worked on improving its guidance on methods adapting both quantitative and qualitative evaluation methods to the types of evaluations that it conducts.[3] The office also provides guidance to other units in the organization who commission evaluations (UNDP, 2009).

With regard to the UNEG norms, I've already discussed independence and credibility above. In addition, I would like to highlight two other norms, namely intentionality and transparency. Intentionality refers to being clear why an evaluation is carried out and what it is expected to contribute to. Evaluation should not be a goal in itself and should not pose undue burden on those responsible for the evaluand. Instead, it is important to define clearly whether the evaluation would contribute to a particular cause, be it accountability for results achieved to benefit the intended benefits, effective use of resources, program improvement or learning from the past for future operations.

Transparency is also important, both in planning and conduct of evaluations, as well as sharing their results and what will be done about them. For the purpose of transparency, the UNDP Executive Board has decided that all evaluations produced by UNDP—be they independent or commissioned by program units—and the management responses thereto are public documents available to all and every organization or individual. The public website, Evaluation Resources Center,[4] contains all evaluation reports and management responses, as well as the monitoring system of the implementation of the agreed follow-up actions to each of the evaluations. This transparency has allowed also Mr. Lempert to scrutinize the UNDP evaluation-related documents that he refers to in his analysis.

It is true that the management does not have to agree with all the conclusions and recommendations put forth by an evaluation. After all, evaluators are not super-humans with absolute knowledge and understanding of the evaluand. Rather, they make evidence based findings and suggestions on how to address problems or improve performance (or to close down a program, change strategy, etc.). The lame management responses referred to by Lempert surely exist, but I believe that in

[3] http://www.undp.org/evaluation/
methodologies.htm

[4] http://erc.undp.org

UNDP's case they are getting rarer and occur mostly in case of the decentralized evaluations. It is required that the management responses address each of the evaluation recommendations and give reasons in cases when the management does not agree with the recommendation. I could give several examples from evaluations conducted by the Evaluation Office where the Executive Board has rejected the management response and requested a stronger one to be presented in the next session.

Finally, what is the role of evaluation? Lempert seems to lump all oversight functions together into one pot under the lid of evaluation. Many of the indicators and questions proposed seem to be more the function of audit or even investigations than evaluation (a blatant example of such would be *Question I.B.3. System 3. Cost Accounting and Cost Control Systems for Procurement and for Personnel*—surely this is not an evaluation function!).

Certainly, the role of evaluation is neither to make decisions about how a policy or program should be changed or terminated. Lempert complains (p. 60) about cases "where evaluators are not empowered to stop or replace projects or redesign failing systems, or to hold anyone accountable for performance…" Under no circumstances can or should the evaluator act as the prosecutor, judge, jury and executioner all in one. Furthermore, getting into "redesigning failing systems" would be a direct conflict of interest and would undermine the very independence of the evaluators. Who would be the watchdog to ward against this and who would then evaluate what the evaluators have created?

The role of evaluation should be to provide objective and credible information on the relevance, effectiveness, efficiency, sustainability and performance of the evaluand, be it a strategy, policy, program, project, organization, unit, etc. Based on evidence, solid methods and professional judgment, evaluators will give options or propose possible solutions to identified issues or problems, but it is up to the management to decide how to act on such recommendations. It is of utmost importance that such discussions are made in the public domain and that the decisions are transparent.

References

Gariba, S., Balogun, P. & Thanh An, P.T. with Hildenwall, V. (2010). *Independent review of UNDP evaluation policy.* New York: UNDP.

EO. (2008). *Evaluation of the role and contribution of UNDP in environment and energy.* New York: UNDP Evaluation Office.

EO. (2010). *Annual report on evaluation in UNDP 2009.* United Nations DP/2010/19.

Lempert, D. H. (2010). Why government and non-governmental policies and projects fail despite 'evaluations': An indicator to measure whether evaluation systems incorporate the rules of good governance. *Journal of MultiDisciplinary Evaluation, 6*(13), 58-108.

UNDP. (2006). *The evaluation policy of UNDP.* United Nations DP/2005/28.

UNDP. (2009). *Handbook on planning, monitoring and evaluation for development results.* New York: United Nations Development Programme.

UNEG. (2006). *Norms for evaluation in the United Nations system.* United Nations Evaluation Group.