

Independent Advisory Committee for Development Impact (IADCI)

Evaluation independence at DFID

An independent assessment prepared for IADCI

Robert Picciotto

August 29, 2008

Independent Advisory Committee for Development Impact (IADCI)

Evaluation independence at DFID

Foreword

This independent assessment, prepared by Robert Picciotto¹, is the final version of a paper designed to facilitate consideration of reforms aimed at strengthening evaluation independence at DFID. It was commissioned by the Independent Advisory Committee for Development Impact (IADCI) after its first meeting. An earlier draft provided the basis for the Committee's discussion of evaluation independence at DFID during its second and third meetings.

At its third meeting Committee members recognized that the creation of IACDI had signaled DFID's intent to increase the independence and impact of its evaluation function. DFID had already responded to one recommendation from the Committee – a regular report on actions taken following evaluations and the appointment of Directors responsible for follow up in each case. However, the Committee concluded that a number of further steps were needed to match good practice in evaluation. Once taken, such steps would help the Committee attest to the independence, impact and effectiveness of evaluation in DFID, as envisaged in its terms of reference.

The Committee's package of proposals is shown below. It consists of actions to be taken in two phases. The first phase of actions (a to f) would make the Evaluation Department (EVD) more influential and add to its budget resources. Next, IACDI would expect to actions g to k to be implemented over the next 2 – 3 years. Some of these actions might best be taken when the current head of EVD is replaced.

A. Proposals to be implemented more immediately

- (a) An agreed Departmental policy on evaluation that meets internationally recognized criteria, and is endorsed by IACDI. (The Committee acknowledged and welcomed progress being made on this).
- (b) Evaluation in DFID undertaken outside EVD to be greatly strengthened. This would require a cultural change in DFID and the creation of the right incentives. There is a case for EVD taking on appropriate responsibility for oversight, quality assurance and guidance on all evaluations carried out in DFID, as at present this function seems fragmented and largely left to the judgment of line managers. This would require extra resources for EVD, an issue to be resolved as part of the new evaluation policy.
- (c) A formal role for IACDI, or its chair, in agreeing the job description, protocols and arrangements for performance review of the Head of EVD would help buttress his independence.
- (d) Clear written protocols for unimpeded access by EVD, and their consultants, to information in DFID; for rules of engagement with DFID staff in discussing draft reports; for rules for avoiding staff conflicts of interest; and a written policy on disclosure of reports.
- (e) As agreed by the Head of EVD, an annual report drawing lessons and common themes from reports completed during the year and to discuss it with IACDI in draft form. It would be published

¹ Robert Picciotto, Professor, King's College, London was Director General, Evaluation at the World Bank (1992-2002)

and appended to the Chair's annual letter to the Secretary of State, which is to be copied to the IDC. The Committee welcomed this.

(f) DFID Management Board to hold at least one annual discussion of evaluation work, perhaps to coincide with the annual report envisaged in (e).

B. Proposals to be implemented over a 2 – 3 year period

(g) To match good practice the Head of EVD would be accountable to the Permanent Secretary of DFID, and be able to make reports without clearance from the management line

(h) The post would also have a more senior grade and status to give greater visibility and clout.

(i) The contradiction between the need for more, high quality evaluation and a declining administrative budget for evaluations would be resolved by exploring ways to protect the budget for evaluation. IACDI would have a central role in deciding EVD's budget in future.

(j) There would be changes in modalities for (and control of the Head of EVD over) staffing, probably with more staff recruited from outside DFID. Over time EVD would need to change the balance towards an increased role for EVD staff and a lesser role for external consultants.

(k) Future heads of evaluation would probably be appointed on an understanding precluding employment elsewhere in DFID, and the contract should be for a fixed term, possibly renewable on the advice of IACDI.

Independent Advisory Committee for Development Impact (IADCI) Evaluation independence at DFID

Introduction

This background paper addresses independence of evaluation at DFID, an issue of central concern to the Independent Advisory Committee for Development Impact (IADCI). It was prepared at the request of the IADCI to help frame its discussions on this topic².

During the first committee meeting members opined that independence should not be equated with isolation and that independent evaluation needs to have “clout”. They noted that the credibility of evaluation hinges on public perceptions as well as on reality. The Committee concluded that reviewing the reporting line alone would not be sufficient: complementary actions would be needed. Beyond organizational independence, the Evaluation Department (EVD) should have mechanisms in place to protect the integrity of evaluation work by consultants and the views of developing countries and beneficiaries should be systematically sought.

Accordingly, this report covers governance structures, evaluators’ behavior, conflicts of interest and protection from external interference. It is in three parts. First, evaluation independence is defined and its relevance for evaluation quality and impact is specified. Second, the current status of independent evaluation at DFID is examined. Third, conclusions and recommendations are outlined.

What is evaluation independence?

The meaning of independence has long been a subject of debate within the auditing and evaluation communities. The Handbook of International Auditing by the International Federation of Accountants (2003) distinguishes between *independence of mind* (“the state of mind that permits the provision of an opinion without being affected by influences that compromise professional judgment, allowing an individual to act with integrity, and exercise objectivity and professional skepticism”) and *independence in appearance* (“the avoidance of facts and circumstances that are so significant that a reasonable and informed third party, having knowledge of all relevant information, including safeguards applied, would reasonably conclude (that) integrity, objectivity or professional skepticism had been compromised”). Both meanings are intertwined within the concept of auditors’ independence.

There is little doubt that independence is as central to the credibility of evaluation as it is to the credibility of auditing. The Evaluation Thesaurus³ indicates that by increasing independence in evaluation “*we can decrease certain types of bias (including) ...extreme conflicts of interest where*

² Although the author greatly benefited from the Evaluation Department’s comments on a previous draft, the judgments contained in this document are his and his alone.

³ Michael Scriven. *Evaluation Thesaurus*. Fourth Edition. Sage Publications. Newbury Park, London and New Delhi. 1991.

the evaluator is 'in bed with' the program being evaluated...typical of much program monitoring by agencies and foundations where the monitor is usually the godfather of the program, sometimes its inventor, and nearly always its advocate at the agency..."

Within the development evaluation profession, an authoritative definition of evaluation independence is found in the *Glossary of Key Terms in Evaluation and Results Based Management* issued by the Development Assistance Committee of the OECD. It specifies that an evaluation is independent when it is “*carried out by entities and persons free of the control of those responsible for the design and implementation of the development intervention*”. It also indicates that independent evaluation presumes “*freedom from political influence and organizational pressure*”, “*full access to information*” and “*full autonomy in carrying out investigations and reporting findings*”.

These and other sources suggest that, along with quality assurance and ethical guidelines, independence is a critical feature of evaluation credibility and reliability. This is why persons that have reason to fear the outcome of an evaluation will frequently throw doubt on its independence. To be sure, independence on its own does not guarantee quality (relevant skills, sound methods, adequate resources and transparency are also required) but there is no necessary trade-off between independence, quality or credibility.

Indeed, evaluation quality without independence does not assure credibility. Furthermore, in open and accountable working environments, evaluation independence induces credibility, protects the learning process and induces program managers and stakeholders to focus on results. Thus, evaluation independence, quality and credibility are complementary characteristics that together contribute to evaluation excellence.

The limits of independence

Optimum independence is not maximum independence. Accurate and fair evaluations combine intellectual detachment with empathy and understanding. The ability to engage with diverse stakeholders and secure their trust while maintaining the integrity of the evaluation process is the acid test of evaluation professionalism. This is why diminishing returns set in when evaluation independence assumes extreme forms of disengagement and distance. Independence combined with disengagement increases information asymmetry, ruptures contacts with decision makers and restricts access to relevant sources of information. It leads to isolation, a lack of leverage over operational decision making and a chilling effect on learning. Thus, the basic challenge of evaluation governance design consists in sustaining full independence without incurring isolation.

External evaluation is often equated with evaluation independence. Yet, external evaluators often lack an appreciation of the operating context. Furthermore, their judgment may be impaired by their dependence on funding from the very managers in charge of activities being evaluated. Specifically, fee dependence often threatens the integrity of the evaluation process. By contrast, internal evaluations funded and controlled by a supreme governance authority can be protected from management influence while enjoying proximity with the programs being evaluated. Such evaluations are more likely to overcome information asymmetries while protecting the objectivity of the process from management influence.

In sum, evaluators should not be so detached as to shirk interaction with program managers, staff or beneficiaries. But they need resilient protection from threats to their impartiality as well as mandatory access to the operational information they need to carry out their work. They should be immune to capture by any of the parties that share in the responsibilities of operational management. This implies special organizational safeguards but, in addition, it means that distinctive personal characteristics, attitudes and behaviors⁴ should be cultivated among evaluators. Good evaluation just as good science calls for a frame of mind characterized by curiosity, skepticism and a hunger for evidence. It also calls for a fair and balanced approach that does not shirk from pointing out problems and performance shortfalls but also recognizes success and achievement.

Criteria

How then can one assess the adequacy of independent evaluation? The answer lies in guidelines that reflect the above considerations and draw legitimacy from a participatory design embedded within the profession. Fortunately, such guidelines for independence in development evaluation are available. Drawing on the good practice standards of official audit and evaluation agencies that span government and the corporate sector, four dimensions of evaluation independence have been recognized by the Evaluation Cooperation Group (ECG), a network composed of the heads of evaluation of the multilateral development banks (MDBs) and such observers as the DAC Evaluation Network head and the Director of Independent Evaluation at the International Monetary Fund⁵.

The criteria (Annex 1) include: (i) organizational independence; (ii) behavioral independence; (iii) protection from external influence and (iv) avoidance of conflicts of interest. The four criteria are interrelated. Protection from outside interference is the object of organizational independence. Conflicts of interest are frequent absent organizational independence. Behavioral independence is a function of organizational independence as well as avoidance of conflicts of interest and protection from external interference.

Organizational independence ensures that the evaluation unit and its staff are not under the control or influence of decision-makers who have responsibility for the activities being evaluated and that they have full access to the information they need to fulfill their mandate. *Behavioral independence* measures the extent to which the evaluation unit is able and willing to set its work program, produce high quality and uncompromising reports and to disclose its findings to the Board without management imposed restrictions. *Protection from outside interference* keeps the evaluation function free to set its priorities, design its processes and products, reach its judgments and administer its human and budget resources without intrusion by management. *Conflict of interest safeguards* guarantee that current, immediate future or prior professional and personal relationships and considerations are not allowed to influence evaluators' judgments or create the appearance of a lack of objectivity.

⁴ The Yellow Book of the General Accounting Office of the United States that sets out criteria for auditing and evaluation identifies "an independent attitude and appearance" as desirable characteristics.

⁵ The Evaluation Cooperation Group criteria are based on standards developed in the auditing profession and government

All four dimensions are important and none implies necessary trade-offs with the operational relevance or influence of evaluation. Evaluation quality cannot be divorced from its independence: poorly designed evaluations can be misleading, disruptive and costly. Conversely, quality is undermined when the evaluation process is captured and vested interests: (i) constrain information so that evaluation products cannot have any critical content; (ii) control the content of the evaluation program so that it does not contribute any new knowledge; (iii) delay the evaluation process (or the disclosure of evaluation results) until after the decisions that might have been informed by the evaluation are taken; (iv) induce evaluators to focus on irrelevant or marginal aspects of the program or policy being evaluated.

It follows that behavioral independence is a privileged dimension of evaluation excellence. It can be ascertained not only by assessing governance structures, processes and practices but also by examining whether the independent evaluation unit produces evaluations that are endowed with one or more of the following characteristics: (i) *criticality*: the ability and willingness to judge performance in an objective and transparent fashion; (ii) *additionality*: a distinctive contribution to operational knowledge creation or dissemination; (iii) *timeliness*: the delivery of operations evaluation findings and lessons early enough to inform decision making; and (iv) *materiality*: a deliberate focus on topics and issues that have substantial relevance to development effectiveness.

Where does evaluation independence stand at DFID?

Based on the above criteria, this section provides an assessment of EVD's independence; highlights major threats to its independence and identifies specific obstacles that need to be addressed. Decisive action is needed since according to an opinion survey carried out by the National Audit Office (NAO)⁶, more than four fifths of DFID staff consider the independence of the evaluation function to be extremely or very important while less than a quarter believe that EVD is entirely or very independent and almost forty percent view organizational pressures to water down critical reports as significant threats to high quality, independent evaluation work.

The Evaluation Department

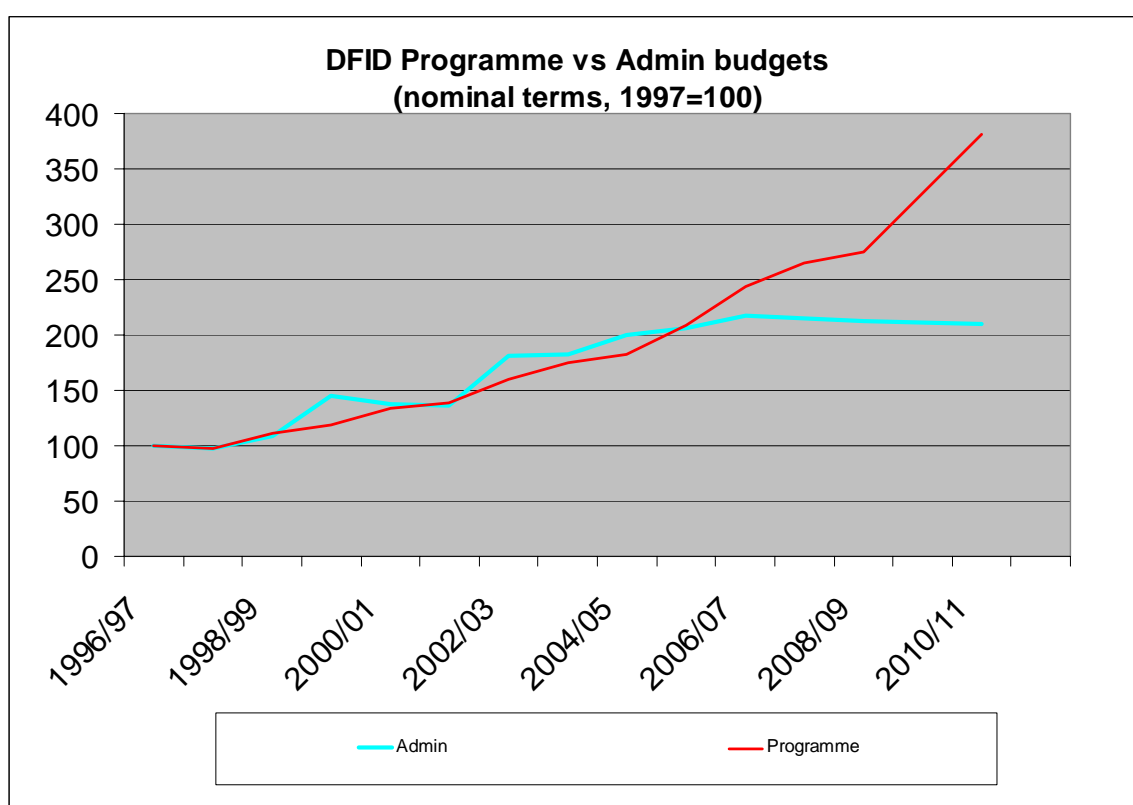
EVD is a centralized unit of 20 staff based in East Kilbride. It is tasked with the production of independent evaluations of country assistance plans; policies (e.g. gender, HIV) or themes/processes (e.g. budget support). The department is composed of three teams: (i) Policy and Development Effectiveness; (ii) Country Programme Evaluation; (iii) Partnerships and Support.

Rather than carry out evaluations directly, EVD commissions independent evaluations by external consultants and manages the evaluation process. In addition, EVD sponsors – or participates in – evaluations carried out in concert with other agencies (e.g. the on-going joint evaluation of the Paris Declaration).

⁶ The NAO survey was issued to 204 staff in DFID selected randomly from a population of 1405 staff with grades judged to be relevant to the assessment. The response rate was 32%. Similarly, 300 DFID stakeholders (NGOs, academics, donor agencies) were surveyed with a response rate of 14%. There were no obvious biases in the characteristics of response group vs. the non response group.

EVD also communicates and disseminates evaluation findings and recommendations to facilitate their utilization in decision making. It provides a policy lead, evaluation training and guidance to staff at HQ and in country offices, especially the three Regional Divisions charged with mid-term reviews, progress monitoring, completion reporting and lessons learnt. Finally, EVD represents the United Kingdom in international evaluation networks and forums.

EVD’s annual budget (2007-8) amounted to £3.7 m (0.08% of the DFID budget), of which £2.7m for administration and £1m for programme costs. Given that the overall administrative budget envelope of DFID is planned to be reduced even as the volume of DFID funding and the scope of its operations subject to evaluation expands rapidly there are risks that EVD’s budget will not keep up with priority evaluation needs.



No formal evaluation mandate

According to the criteria used for this independent assessment, the first criterion of organizational independence stipulates a mandate statement that makes clear that the unit’s scope of responsibility extends to all operations of the organization and that its reporting line, staff and budget functions are organizationally independent from the operational, policy and strategy departments. Strictly speaking, DFID does not comply with this standard since the Evaluation Policy note produced in July 2002 was never officially approved even though its content is unobjectionable.

The Policy note lists two key objectives: (i) strengthening accountability by documenting the allocation, utilization and results of DFID’s development assistance; and (ii) promoting learning

through feedback of lessons learnt from experience. It is deemed to guide EvD's activities. Is the lack of a formally approved evaluation policy significant? Not according to EvD's self assessment which asserts that everyone in the organization recognizes that EvD can challenge DFID's operational management.

Thus, a guidance note to staff⁷ refers to EVD's outputs as "*independent, impartial, external evaluations conducted on completion of a project or program*". According to EvD, the Department is able to examine critically the full range of DFID's operational activities and it is 'relatively independent'. On the other hand, the ODI benchmarking study⁸ concludes that DFID does need a widely understood evaluation policy that clarifies the roles of the different divisions in DFID in evaluation processes: "*Such a document would certainly be useful to the EVD not only as a back up tool when evaluation is slipping off other Divisions' agendas and priorities but also as a formal document laying out the key features and function of EVD, such as its independence*".

The absence of a formal evaluation policy may have contributed to organizational uncertainty. It may also have fed perceptions of an ambiguous relationship between independent and self evaluation. Formal approval of an evaluation policy that complies with ECG criteria is imperative – and as was stressed by IADCI members at its first meeting. Fortunately, EVD is currently engaged in the development of a new evaluation policy and strategy in consultation with IACDI.

The independent- self evaluation interface needs strengthening

Evaluation is most effective in organizations that promote learning and accountability. Just as financial auditing requires the prior production of accounts, independent evaluation is heavily reliant on the capacity and willingness of line departments to carry out credible self evaluation of their activities.

The need to give priority to enhanced self evaluation is highlighted by the NAO opinion survey. It discloses that both DFID staff and stakeholders consider improvement in the "evaluability" of programs, projects and strategies at the outset to be extremely important. It is the highest ranking factor among changes that would enhance the quality of evaluation work at DFID according to survey respondents. The second most important factor is the definition of clear evaluation standards and the third is the need to invest in M&E staff training.

The need to fill the *self evaluation gap* has been recognized by DFID management. The new Investment Committee is expected to emphasize results based management. Clear responsibility for self evaluation has been assigned to line managers. A quality assurance group has been set up in a central location and EVD is now more regularly consulted. But much remains to be done to improve the quality and rigor of self evaluations.

Recent reviews confirm that the log frames prepared by operational staff for new operations are not regularly vetted by EVD. Nor is the quality at entry and supervision quality of DFID operations independently measured and tracked in real time. No substantial effort is devoted to projecting

⁷ Guidance on Review and Evaluation for DFID staff, July 2005

⁸Overseas Development Council, 2006. Benchmarking Study on Evaluation Policies and Practices, DFID Case Study

likely development impact. Completion reports are not mandatory for operations of under £1 m. Even for other operations, completion reporting coverage is incomplete. In general, quality assurance of self evaluation reports is patchy and independent validation of completion ratings is not carried out.

The relatively low priority traditionally accorded to self evaluation in DFID is linked to the absence of a focal point within senior management for oversight of monitoring and self evaluation⁹. EVD is not resourced to do this work in a systematic way since the independent review of Finance and Corporate Division monitoring and Regional Divisions' self evaluation standards, methods and processes has not been considered a core EVD function. Nor is the 'traffic light' system that signals progress towards strategic objectives transparently connected to the self evaluation system. Ensuring coherence in self evaluation practices across managerial units and facilitating coordination in the interaction between the operational units and EVD would help remove a significant weakness in DFID's current evaluation governance.

Nor is evaluation transparency adequate: all EVD reports are public but self assessment reports are not routinely disclosed. And while quarterly and annual management and PSA reports include project scores, quality ratings for projects under implementation (as well as development effectiveness ratings for completed operations) are not regularly published.

To be sure, DFID's grant resources are increasingly devoted to country wide programs, partnership activities and capacity building operations. While these activities are hard to evaluate let alone rate for development effectiveness, appropriate methods are available within the evaluation tool kit. In recognition of this, EVD now gives pride of place to country program evaluations. Similarly, DFID's new Quality Assurance Group has been entrusted with the task of ensuring that country programs comply with agreed sector and thematic policies.

Without the regular and timely production of adequate self evaluation documents, EVD cannot be expected to exercise an effective operational quality oversight role. Conversely, the *quality of self evaluation* hinges significantly on the extent to which intellectual guidance and authoritative oversight is provided by EVD.

Effective evaluation oversight requires *regular reporting* to the public on the overall status of independent and self evaluation processes within DFID. In addition, the *dissemination* of lessons learnt within the organization should be mandated (with implications for publications, training and data base management). Equally, EVD's role in *follow up* should be specified since the value of independent evaluation hinges on the influence it exerts on operational management and staff behavior. Finally, EVD's role in *evaluation capacity development* in borrowing countries should be enhanced since country led evaluations have begun to be viewed as central to country ownership of development programs and projects.

⁹ EVD reports are widely circulated within DFID but there is no formal procedure for the presentation of individual reports to any oversight body and there is no overtly designated champion for evaluation on the Management Board

In sum, the current EVD review should formulate a formal evaluation policy that covers the *evaluation responsibilities of all departments*, assigns to EVD the role of *attesting to the adequacy of the overall evaluation system* within DFID; ensures that the implementation of agreed evaluation recommendations is *monitored*, takes effective steps to *disseminate* lessons of experience to the development community and promotes *evaluation capacity development* in member countries.

Organizational independence is not assured

The second ECG organizational independence criterion explicitly prescribes accountability of the function to the head of the organization, its deputy head – or its governing board. Prior to the set up of IADCI three degrees of separation were interposed between EVD and the Secretary of State. Specifically, EVD reported to a Divisional Director (Finance and Corporate Performance Division) who, in turn, reported to a Director General (Corporate Performance) – who reported directly to the Permanent Secretary, who reports to the Secretary of State. This unusual location made EVD an ‘outlier’ in comparison to other development agencies.

Recently, it was decided to have the EVD head report to the Director General, Corporate Performance who reports directly to the Permanent Secretary. This central position facilitates engagement across the organization and it complies with the third indicator of organizational independence that prescribes an arm’s length relationship between operational management and evaluation. However, a further step would be needed to comply with good international practice, i.e. the EVD head should report directly to the Permanent Secretary.

The fourth independence criterion stipulates regular reporting to the larger organization’s Audit Committee - or a similar oversight body. To be sure, EVD reports regularly to the Development Committee (DC). But it is composed of managers and advisers¹⁰ and it cannot be considered akin to an independent audit or oversight committee. Whether IADCI’s remit (and resources) should develop to enable it to fulfill an effective detailed oversight role of EVD is an issue worthy of consideration and the approval of EVD programs by IADCI instead of the DC might be viewed as a first step in this direction.

While EVD is not directly accountable to a political authority, it may not be sufficiently immune to ‘peer group’ or other pressure to be able to report findings without fear of repercussion. In particular, since the personnel system is the same as for all DFID employees it may deserve fine tuning and the introduction of special arrangements that ensure adequate autonomy and protection to the independent evaluation function and its staff.

General personnel rules that decree that remuneration, training, tenure and advancement are based on merit only go so far. An independent process would be warranted to ensure that full protection is offered to evaluators who deliver unvarnished and unwelcome assessments of performance. In addition, the criteria governing the selection, evaluation, remuneration and dismissal of the EVD head need to be designed with care in a way that helps to guarantee independence for the function given that the current conditions of service for the EVD head are identical to those of other senior civil servants.

¹⁰ DFID Directors, General Directors, Chief Advisers, Head of Development Policy, Head of Information Department.

No explicit instruction to operational staff appears to have been issued to ensure that all needed information is made available to EVD so that it can carry out its work efficiently. Yet, good evaluation practice specifies that the evaluation unit has “*unrestricted access to the staff and records of the organization*” and also that operational managers should “*facilitate visits to project sites and meetings with client, partner and government representatives*”.

Behavioral independence should be promoted

To demonstrate behavioral independence is to display a willingness and capacity to issue strong, high quality and uncompromising reports – and to do so more often than “occasionally”. Producing such reports is the acid test of behavioral independence. In part, such behavior reflects organizational incentives. Unless an evaluation unit feels secure enough to report its findings without fear of reprisal, it cannot be deemed to be independent. Equally, behavioral independence is shaped by the willingness and capacity of evaluation managers and staff to stand up publicly to their operational colleagues and defend their evaluation methods and findings.

This can be very demanding both professionally and personally since operational managers have greater knowledge of the detailed features of programs and projects subject to evaluation than the evaluators. Public disclosure, a key criterion of behavioral independence, adds to the stress since transparency offers other stakeholders some of whom have intimate knowledge of the program under evaluation (as well as interests to defend) a chance to comment on the quality of the evaluation. From this perspective, transparency in the reporting of evaluation findings raises the stakes for evaluation quality.

These considerations regarding quality suggest that behavioral independence is inseparable from evaluation transparency, skills, methods and resources. Rather than a once for all organizational characteristic, behavioral independence is an ideal towards which evaluation managers should continually strive. The outcomes are contingent on the performance of individual evaluators but they also hinge on organizational structure and contracting processes that offer adequate protection and incentives for “telling it like it is” and promote intellectual rigor, honesty and firmness in evaluation.

Creating such a working environment is a feature of good evaluation governance and organizational learning. Reportedly, EvD’s evaluations have become more challenging and quality assurance standards for EvD’s evaluations have been upgraded. However, the reliance on quality oversight by steering committees which (along with external experts and NGO representatives) includes officials representing different DFID divisions may be perceived as lacking objectivity. A careful assessment of evaluation quality by IADCI as well as an independent stakeholder survey would confirm the extent to which EvD has achieved a distinguished record of behavioral independence.

In the past it has been felt that the use of external consultants is supportive of behavioral independence. This is because EvD management and staff encourage consultants to be rigorous and critical; to comply with guidelines and to respond to the guidance of evaluation steering groups. However, given the dependence of major UK development oriented consulting firms on DFID funding, external evaluation cannot always be assumed to be equivalent to independent evaluation. Ultimately, behavioral independence hinges on the willingness and capacity of EVD as contracting party to take full charge of quality assurance and ensure that evaluations are not unduly influenced by vested interests through the review process.

Most operational staff value evaluation but there are concerns that in practice units sometimes give a lower priority to evaluation than should be the case. Since evaluation delayed is evaluation denied additional efforts are needed to make clear to staff that the current managerial emphasis on results implies timely evaluations.

Quality standards and ethical guidelines strengthen behavioral independence and help guarantee that evaluation findings are grounded in evidence. But good practice guidelines should also provide country officials and other stakeholders with a full opportunity to comment on drafts and to record their disagreements in dissenting footnotes and statements. Full interaction among evaluators, country officials and other stakeholders enhances evaluation quality, and increases the probability of ownership and use of evaluation findings by stakeholders. The ability to report candidly is also a function of the evaluation practices through which reports are prepared, reviewed and disseminated. Disclosure and outreach are equally critical and to this end greater dissemination efforts by EVD may be necessary.

The NAO survey highlights inadequate resources devoted to evaluation as the most significant perceived threat to high quality independent evaluation work according to staff as well as stakeholders. Thus, the issue of budget allocations for evaluation needs careful consideration from a behavioral independence perspective. Though management has apparently not used its budget authority to interfere with evaluation priorities, the aggregate budget constraint imposed on EVD has been severe especially before reports by the National Audit Office and a DAC peer review four years ago induced the Management Board to take action and double the resources allocated to EVD.

This experience confirms the utility of regular independent checks on the adequacy of evaluation budget resources. Greater independence in evaluation at DFID ultimately means that the budget resources allocated to EvD should be consistent with its evolving mandate. Furthermore, suitable arrangements should be made for independent validation (perhaps by IADCI) that EVD has the resources it needs to do its work within an expanded mandate.

Evaluation “clout” implies an appropriate *grade* and *title* for the EVD head position in order to signal the degree to which the sensitivity, complexity and importance of the function are recognized by the institution. The need to reassess the status of evaluation at DFID is illustrated by comparing the level and the title of the EvD head with that of the head of independent evaluation at the World Bank where the Director General, Evaluation position¹¹ is at the Senior Vice President level (just below the President). Similarly, the head of evaluation at the IMF is graded immediately below the level of the deputy managing directors. Taking account of the proposed upgrading of the evaluation function, DFID may wish to revisit the grade and title of the EVD position to ensure that it reflects the importance it places on independent evaluation for corporate accountability and learning.

Equally, there should be a clear cut DFID policy statement on the full disclosure of evaluation documents.

¹¹ The Director General, Operations Evaluation in the World Bank Group oversees the work of three evaluation units (the Operations Evaluation Department in IBRD/IDA, the Operations Evaluation Group in IFC and the Operations Evaluation Unit in MIGA)

Last but not least, an important dimension of behavioral independence has to do with the extent to which the programming of evaluation activities is subject to management interference. Before 2008 the evaluation program was approved by the Development Committee which, according to the current informal evaluation policy document of 2002 is “*the principal client and audience for studies produced by EVD*”. While this has not led to substantive changes in EVD’s programs, a potential threat to behavioral independence remained until it was decreed that IADCI would approve the work program.

Protect EVD from outside influence

Evaluation can be hindered either through active as well as passive interference in the conduct of evaluations. As indicated above, there have been delays imposed on the evaluation process by some operational managers and overseas offices. Improving this state of affairs may require greater senior management emphasis on the priority of independent evaluation activities in order to remove any perception of operational interference in EVD’s work.

Program funding for country evaluations might provide another potential avenue for outside influence in the evaluation process that needs to be closed. Furthermore, clear evidence is needed that judgments made by evaluators are invariably protected. Evaluators’ judgments are only accepted where the evidence base is robust and there are no errors of fact. Inevitably, differences of views on these matters arise. Hence, transparent contestability processes that give EVD the final word should be put in place in order to make clear that outside influence is not allowed to affect the integrity of the evaluation process.

Finally, EVD has up to now been treated as any other department in terms of hiring/firing, term of office, rotation, performance review of unit head and unit staff, compensation, etc. In particular, the head of EVD is eligible for appointment to other parts of DFID upon completion of his assignment. The acceptable reasons for dismissal of the EVD head (e.g. demonstrated inefficiency or malfeasance) have not been made explicit.

Conflicts of interest should be avoided

Providing a wide range of stakeholders with an opportunity to comment on the findings of an evaluation helps to correct evaluator bias. Participatory evaluation methods whereby beneficiaries of interventions and civil society representatives contribute to the evaluation strengthen evaluation quality and transparency. But good practice standards also require evaluation managers and staff to be excluded from evaluating programs, activities or entities that might involve a conflict of interest or create the perception that current or past associations and/or activities (whether personal or professional) could impair the objectivity and integrity of the evaluation process.

EvD appears to have followed the precept that precludes staff or consultants to evaluate operations in which they have been involved. For internal evaluations, DFID procurement rules and clauses in consultancy contracts seek to address this issue while EVD is sensitive to its implications in the management of independent evaluations. The guidance on evaluation and review in EVD covers conflict of interest but it needs additional detail, more profile and updating as well as dissemination to ensure that it is fully understood and systematically used.

What is to be done?

This review of evaluation independence at DFID comes at a propitious time. The resources available to DFID are growing. The decentralization to overseas offices is generating new energies, capturing new synergies and making new partnerships feasible. This augurs well for more coherent approaches to country programming as well as greater harmonization of practices among development partners. But given the heightened priority accorded to results, independent evaluation is critical to the credibility of DFID evaluations and the achievement of development impact.

Other aspects of evaluation excellence (quality of evaluation evidence, appropriate evaluation methods, etc.) will have to be addressed as well. Equally, the role of developing countries in DFID's evaluations will have to be revisited since performing developing countries are increasingly called upon by DFID to be in the "driver seat" of aid programs and projects. Consequently, developing countries should also be required to play a more active role in DFID's evaluations. In parallel, evaluation capacity development should become more prominent in DFID evaluation activities.

To be sure, the creation of IADCI is emblematic of a serious determination to enhance corporate accountability and organizational learning through enhanced evaluation independence. But the major conclusion of this review is that the evaluation function at DFID does not currently comply with internationally recognized independence standards. It is probable that other bilateral aid agencies do not meet them either. However, the standards used for this assessment reflect good practice within the broader evaluation and auditing community. Hence, in pursuit of evaluation excellence, it would be appropriate for DFID to implement them in full while respecting the special needs and circumstances of the organization and the government context within which DFID operates.

This said, and considering the numerous personnel and budget constraints within which EVD operates it has performed remarkably well. It has produced useful and numerous evaluation studies. It has shifted evaluation at DFID to the higher plane of country evaluations. It has also launched influential joint evaluations with other donors. Yet, it cannot yet be characterized as a fully independent evaluation unit and as a centre of evaluation excellence that can be fully relied upon for quality assurance, corporate accountability or organizational learning.

Nor within the current evaluation governance structure could it have been expected to fulfil these roles given the limited resources at its command; the ambiguities of its relationships with operational departments and its relatively low status and modest profile within the organization. Revealingly, EVD has not generated annual reports about the development effectiveness of DFID. It enjoys no clear cut oversight authority over self evaluation. Yet, there is a high effective demand for such a role: increasingly, operating departments are seeking authoritative advice regarding their self evaluation methods and tasks.

The major weaknesses that need to be addressed are the lack of a formal evaluation mandate currently being addressed by EVD's on-going review; the absence of effective oversight over self-evaluation; the still distant reporting relationship to the Secretary of State; the lack of connection to a truly independent oversight body (although the IADCI may as it develops be able to play this role); the modest grade and status of the EVD head; the variable priority assigned to evaluation by

operational departments; the prevailing culture of negotiation and compromise that do not favour the timely generation of candid and hard hitting evaluation reports; the weak autonomy from management regarding personnel matters and budgeting; the lack of an evaluation disclosure policy; and the absence of detailed and explicit rules regarding conflicts of interest.

The IACDI recommendations listed in the foreword would tackle all of these obstacles and enhance the rigor, objectivity and impartiality of evaluation at DFID while at the same time narrowing the professional distance between evaluators and operational departments. Once these recommendations are acted upon, the DFID evaluation function would comply with the good practices of other development comparator organizations while respecting DFID's distinctive character and traditions. Ultimately, the reforms outlined in this report would help DFID deliver better development results even as the volume of UK aid increases in line with its international obligations.

Template for Assessing the Independence of Evaluation Organizations

Criterion	Aspects	Indicators
I. Organizational independence	The structure and role of evaluation unit	Whether the evaluation unit has a mandate statement that makes clear its scope of responsibility extends to all operations of the organization, and that its reporting line, staff, budget and functions are organizationally independent from the organization’s operational, policy, and strategy departments and related decision-making
	The unit is accountable to, and reports evaluation results to, the head or deputy head of the organization or its governing Board	Whether there is a direct reporting relationship between the unit, and a). the Management , and/or b). Board or c). relevant Board Committee, of the institution
	The unit is located organizationally outside the staff or line management function of the program, activity or entity being evaluated	The unit’s position in the organization relative to the program, activity or entity being evaluated
	The unit reports regularly to the larger organization’s audit committee or other oversight body	Reporting relationship and frequency of reporting to the oversight body
	The unit is sufficiently removed from political pressures to be able to report findings without fear of repercussions	Extent to which the evaluation unit and its staff are not accountable to political authorities, and are insulated from participation

		in political activities
	Unit staffers are protected by a personnel system in which compensation, training, tenure and advancement are based on merit	Extent to which a merit system covering compensation, training, tenure and advancement is in place and enforced
Criterion	Aspects	Indicators
	Unit has access to all needed information and information sources	Extent to which the evaluation unit has access to the organization's a). staff, records, and project sites; b). co-financiers and other partners, clients; and c). programs, activities, or entities it funds or sponsors
II. Behavioral Independence	Ability and willingness to issue strong, high quality, and uncompromising reports	Extent to which the evaluation unit: a). has issued high quality reports that invite public scrutiny (within appropriate safeguards to protect confidential or proprietary information and to mitigate institutional risk) of the lessons from the organization's programs and activities; b). proposes standards for performance that are in advance of those in current use by the organization; and c). critiques the outcomes of the organization's programs, activities and entities

Criterion	Aspects	Indicators
	Ability to report candidly	Extent to which the organization's mandate provides that the evaluation unit transmits its reports to the Management/Board after review and comment by relevant corporate units but without management-imposed restrictions on their scope and comments
	Transparency in the reporting of evaluation findings	Extent to which the organization's disclosure rules permit the evaluation unit to report significant findings to concerned stakeholders, both internal and external (within appropriate safeguards to protect confidential or proprietary information and to mitigate institutional risk). Who determines evaluation unit's disclosure policy and procedures: Board, relevant committee, or management.
	Self-selection of items for work program	Procedures for selection of work program items are chosen, through systematic or purposive means, by the evaluation organization; consultation on work program with Management and Board
	Protection of administrative budget, and other budget sources, for evaluation function	Line item of administrative budget for evaluation determined in accordance with a clear policy parameter, and preserved at an indicated level or proportion; access to additional sources of funding with only formal

		review of content of submissions
III. Protection from outside interference	Proper design and execution of an evaluation	Extent to which the evaluation unit is able to determine the design, scope, timing and conduct of evaluations without Management interference
	Evaluation study funding	Extent to which the evaluation unit is unimpeded by restrictions on funds or other resources that would adversely affect its ability to carry out its responsibilities
	Judgments made by the evaluators	Extent to which the evaluator's judgment as to the appropriate content of a report is not subject to overruling or influence by an external authority

Criterion	Aspects	Indicators
	Evaluation unit head hiring/firing, term of office, performance review and compensation	Mandate or equivalent document specifies procedures for the a). hiring, firing, b). term of office, c). performance review, and d). compensation of the evaluation unit head that ensure independence from operational management
	Staff hiring, promotion or firing	Extent to which the evaluation unit has control over : a). staff hiring, b). promotion, pay increases, and c). firing, within a merit system
	Continued staff employment	Extent to which the evaluator's continued employment is based only on reasons related to job performance, competency or the need for evaluator services
IV. Avoidance of conflicts of interest	Official, professional, personal or financial relationships that might cause an evaluator to limit the extent of an inquiry, limit disclosure, or weaken or slant findings	Extent to which there are policies and procedures in place to identify evaluator relationships that might interfere with the independence of the evaluation; these policies and procedures are communicated to staff through training and other means; and they are enforced

Criterion	Aspects	Indicators
	Preconceived ideas, prejudices or social/political biases that could affect evaluation findings	Extent to which policies and procedures are in place and enforced that require evaluators: a). to assess and report personal prejudices or biases that could imperil their ability to bring objectivity to the evaluation; b). and to which stakeholders are consulted as part of the evaluation process to ensure against evaluator bias
	Current or previous involvement with a program, activity or entity being evaluated at a decision-making level, or in a financial management or accounting role; or seeking employment with such a program, activity or entity while conducting the evaluation	Extent to which rules or staffing procedures that prevent staff from evaluating programs, activities or entities for which they have or had decision-making or financial management roles, or with which they are seeking employment, are present and enforced
	Financial interest in the program, activity or entity being evaluated	Extent to which rules or staffing procedures are in place and enforced to prevent staff from evaluating programs, activities or entities in which they have a financial interest
	Immediate or close family member is involved in or is in a position to exert direct and significant influence over the program, activity or entity being evaluated	Extent to which rules or staffing procedures are in place and enforced to prevent staff from evaluating programs, activities or entities in which family members have influence

Sources: U.S. General Accounting Office, *Government Auditing Standards*, Amendment 3 (2002); OECD/DAC Working Party on Aid Evaluation, *Glossary of Key Terms in Evaluation and Results Based Management* (2002); OECD/DAC, *Principles for Evaluation of Development Assistance* (1991); INTOSAI, *Code of Ethics and Auditing Standards* (2001); Institute of Internal Auditors, *Professional Practices Framework* (2000); European Federation of Accountants, *The Conceptual Approach to Protecting Auditor Independence* (2001); Danish Ministry of Foreign Affairs, *Evaluation Guidelines* (1999); Canadian International Development Agency, *CIDA Evaluation Guide* (2000).