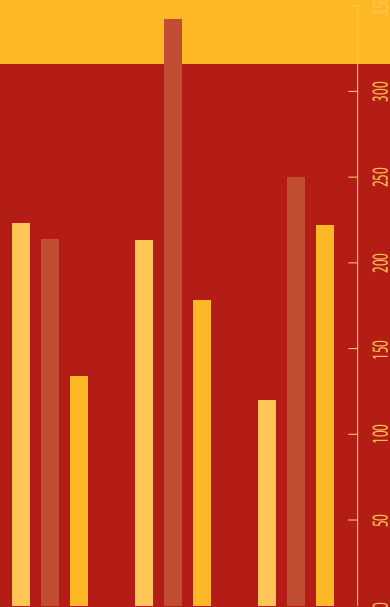
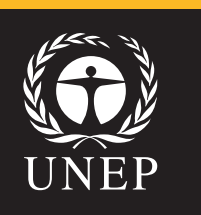




Internal and External Needs for Evaluative Studies in a Multilateral Agency:

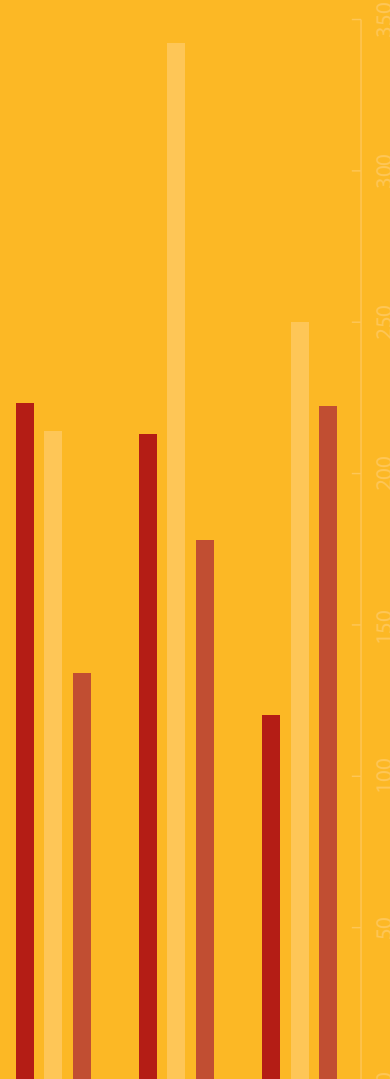
Matching Supply with Demand in UNEP

Special Study Papers No. 1



United Nations Environment Programme

September 2006





UNEP
United Nations Environment Programme

**Internal and External Needs for
Evaluative Studies in a Multilateral Agency:
Matching Supply with Demand in UNEP**

M. J. Spilsbury, G. Rauniyar, C. Perch and S. Norgbey

Evaluation and Oversight Unit
United Nations Environment Programme
Nairobi

September 2006

List of acronyms

CPR	Committee of Permanent Representatives
EOU	Evaluation and Oversight Unit
GC	Governing Council
GEF	Global Environment Facility
MDG	Millennium Development Goal
OECD	Organisation for Economic Co-operation and Development
DAC	Development Assistance Committee
UN	United Nations
UNEG	United Nations Environment Group
UNEP	United Nations Environment Programme

Table of Contents

List of acronyms.....	ii
Executive Summary.....	1
Key Findings and Implications for Evaluation in UNEP	1
Introduction.....	3
Evaluation in UNEP	3
Independence of the Evaluation Function.....	4
Types of Evaluation Activities carried out by EOU	4
Staffing and Resource Allocation.....	4
Objectives of the Study.....	6
Methods	6
Findings and Discussion	7
Organisation of the Evaluation Function.....	7
Perceptions about the Purpose of Evaluation.....	8
Levels of Accountability Expected of UNEP.....	9
Factors Affecting the Credibility of Evaluations	9
Factors Affecting the Utility of Evaluations	11
The Importance of Different Evaluative Parameters	12
The Importance of Indicators for Demonstrating the Results of UNEP Activities	13
Allocation of Resources Across Different Types of Evaluation.....	15
Perceptions on Utility of UNEP Evaluative Products for Resource Allocation Decisions	17
Perceptions about Resourcing Evaluation and Oversight Activities.....	18
Readership of Evaluation Products.....	21
Perceptions on the Quality of UNEP Evaluation Reports.....	21
Sharing of UNEP Evaluation Products	22
Conclusion.....	22

Executive Summary

1. Many organisations supported through donor financing face the dual problem of satisfying donor demands for accountability and securing continued funding. A first step in addressing these problems is to understand the types of evaluative evidence that satisfy accountability demands and those that increase the likelihood of future funding.
2. This study explores how evaluations are used within the United Nations Environment Programme (UNEP) and, to a limited extent, how they influence donor funding decisions. It also provides indications for future direction of the evaluation function of the organization. The study is based on a survey of UNEP Governing Council representatives, UNEP's donor agencies and UNEP project and programme managers. The survey examines preferences for different types of evaluation approaches and methods as well as products, and their perceived credibility, reliability and utility in relation to the resources required to produce them. The findings reveal the relative importance of different evaluation approaches and highlight important tradeoffs, for example, the need to satisfy donor demands for accountability whilst meeting management needs for operational improvement. The study reveals perspectives on the influence of evaluation findings on resource allocation decisions.
3. In light of evolving changes in the evaluation functions within the United Nations System and elsewhere UNEP's Evaluation and Oversight Unit (EOU) developed a new Evaluation Policy to be considered by the organisation's senior management and the Committee of Permanent Representatives (CPR) during 2006. This study was initiated after in-house comments on the draft policy were received in order to explore broader perceptions from key stakeholders relating to the major aspects of the evaluation function articulated in the draft Evaluation Policy.

Key Findings and Implications for Evaluation in UNEP

4. The following represent the key findings of the study:
 - i) The importance of the evaluation function in UNEP is recognized by the CPR and UNEP staff with the survey findings revealing strong support for dedicated evaluation budgets at both the organizational and project levels.
 - ii) The majority of respondents would like to see *considerable independence for the evaluation function* in UNEP. This finding is consistent not only with the recently adopted '*Norms and Standards for Evaluation in the United Nations System*' but also with the draft evaluation policy of the UN's Chief Executive Board and the World Bank's criteria for independence of the evaluation function. Respondents also clearly linked the independence of the evaluation function to higher levels of credibility being afforded to the evaluations it undertakes.
 - iii) UNEP Evaluation activities need to further emphasise improved programme/project delivery and impacts of UNEP activities. Over two-thirds of all survey respondents agreed that '*UNEP activities like any other forms of publicly funded development assistance should have discernable benefits that should be documented.*'
 - iv) The UNEP evaluation function further needs to strive for excellence by improving mechanisms for quality control of evaluation products. The survey revealed that public disclosure of evaluation findings, application of international norms for evaluation standards, peer review of evaluation

products and the independence of the evaluation function are the key factors that can enhance the credibility of UNEP evaluations. Similarly, the most important factors affecting the utility of evaluations are *'timeliness', 'rigour' and 'relevance [of the evaluation] to current organisational priorities'*. EOU needs to ensure that quality control measures applied to UNEP evaluations address these key factors.

- v) There is also a demonstrated need for improved efforts in the dissemination of evaluation findings and products. Specific improvements are required in upgrading the profile of the UNEP EOU web page (to be located on the UNEP's main webpage like other agencies) and better access to evaluation findings, recommendations and lessons.
- vi) There is strong support for the evaluation parameters used by UNEP. Greater focus on such evaluation parameters as "project implementation approach", "financial planning and management", "attainment of outputs", "achievement of objectives", "impacts" and "sustainability" implies that evaluation rating parameters should be weighted.
- vii) Furthermore, given the importance placed by the respondents on indicators of impact, UNEP programmes/projects need to be evaluated more specifically in terms of *'reduced risk and vulnerability', 'influence on international environmental policy processes', 'changes in human capacities and/or levels of empowerment', 'uptake and use of project/assessment outputs' and 'economic valuation of changes in environmental factors'*.
- viii) While recognizing the importance of current evaluation activities, additional demand for studies that demonstrate uptake of proven technologies, policies, new knowledge and/or management practices and evaluations of impact require the expansion of activities currently undertaken by the evaluation function. This is further reinforced by the revealed preferences expressed by the Governing bodies as well as UNEP programme and project managers for conducting evaluations that determine impacts or benefits of UNEP activities. Such studies are perceived as important for informing funding decision-making.
- ix) The study found that as currently configured and deployed, the existing resources for evaluation in UNEP are insufficient to meet the increasing donor needs for accountability in terms of programme as well as impact results. The analysis shows that while the organization established a clear mechanism for funding project evaluations, the current levels of funding for other critical UNEP evaluation activities such as thematic studies and impact evaluations are far too low given the expanded variety of evaluation outputs demanded by UNEP staff and the Governing bodies. This, in combination with the requirements for very high standards in the quality and rigour of evaluations, creates a considerable challenge for EOU to address as UNEP moves forward.

Introduction

5. We have entered an era where development programmes and activities are subject to intense scrutiny by funding agencies. Competition for funding is increasing and the expectations of funding agencies have shifted from the production of outputs towards an insistence that their contributions generate measurable positive impacts in the 'real world'. Demands for accountability and transparency are commonplace and this is motivating an increasing number of institutions to invest in a rigorous and comprehensive portfolio of evaluations, - a trend that is set to continue (Wenar 2006).
6. However, the relative demands for, and the utility of, different types of 'accountability-oriented' and 'learning-oriented' evaluative products have not been well documented and thus remains unclear. Intuitively, one would expect that higher standards of accountability and demonstrated benefits from prior investment (e.g. from *ex post* impact assessment) should increase the likelihood of maintaining or increasing funding levels. Nevertheless, informal discussions with funding agencies suggest that the linkages between evaluation findings and any subsequent resource allocation patterns are often indirect or weak. However, the organisations that are able to demonstrate their effectiveness and efficiency are better placed than those that are not.

Evaluation in UNEP

7. The evaluation function in UNEP was established in 1976 when the organization responded to Governing Council Decision 20 (III) in which the Council requested the Executive Director of the organization to "*present to the Governing council at its fourth session his views and recommendations regarding the methodology he has developed for the evaluation of projects.*"¹ In a subsequent GC Decision (6/13), 1978 the Council recognized "*the importance of the evaluation unit being involved at the outset in programme planning but urged that it be made more managerially independent of the programming and project implementation sections and that it be strengthened by reallocations of existing resources to ensure the realization of its objectives*". The function has undergone numerous changes over the years. As currently constituted, the Evaluation and Oversight Unit (EOU) in UNEP is functionally located within the Office of the Executive Director with management responsibilities delegated to the Deputy Executive Director. Executive management of UNEP fulfils its reporting requirements to the Governing Council by submitting an intersessional annual evaluation report to the Governing Council. The evaluation function also conducts briefings of the Committee of Permanent Representatives on the evaluation activities of the organization.
8. The primary reasons for the existence of the evaluation function in UNEP are to provide the required evidence-based information for policy decision making and to assist UNEP project/programme managers to measure and demonstrate performance, identify where improvements can be made to design or delivery methods, identify good practices and lessons for the future, and in general, provide a tool for adaptive management and positive learning². Another key purpose for evaluations is to determine how UNEP's activities have impacted environmental policy-making and management at the, national, regional and global levels. Evaluations also serve as a basis for substantive accountability to the organization's governing bodies and other stakeholders. Besides substantive accountability, evaluations, through disclosure, provide

1 Governing Council, Fourth Session, Nairobi, 30th March-14 April 1976

2 Regulations and Rules Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods of Evaluation ST/SGB/200/8.
Draft UNEP Evaluation Policy, August 2006

a means for transparency in the way the organization implements its programme activities and uses its resources in an efficient and effective manner.

Independence of the Evaluation function

9. The extent to which the evaluation function needs to be independent has been a focus of continuing debate not only in UNEP but within the United Nations system and the international evaluation community. The level of independence of the evaluation function can be measured against four criteria: **organizational** independence; **behavioural** independence; protection from **external influence**; and avoidance of **conflict of interest**.³ The evaluation function in UNEP enjoys independence in the following areas:
 - a. developing its work program in collaboration with various divisions;
 - b. recruitment of evaluation consultants within the guidelines of the organization;
 - c. conducting evaluations without interference from management ;
 - d. reporting of evaluation findings, and
 - e. follow-up on the implementation of evaluation recommendations.
10. However, measured against the full set of criteria for independence, the evaluation setup at UNEP fails because, as currently constituted, the evaluation function reports directly to management, and therefore lacks ‘organizational independence’. It is clear that in spite of the level of independence currently enjoyed, the evaluation function does not have sufficient independence of resources to commission studies which, in its judgment, might assist in providing strategic advice to senior management and governing bodies and contribute to policy formulation.

Types of evaluation activities carried out by EOU

11. UNEP commonly conducts evaluations at the project and sub-programme levels and also undertakes management studies designed to inform decision-making on key issues. The evaluation function manages the process of conducting evaluations in the organization and has an active monitoring and follow-up system in place to check that agreed evaluation recommendations are indeed implemented. As part of its responsibilities, the evaluation function advises management on key strategic issues that require management action and feeds back findings and lessons from evaluations into programme planning and project development.⁴

Staffing and resource allocation

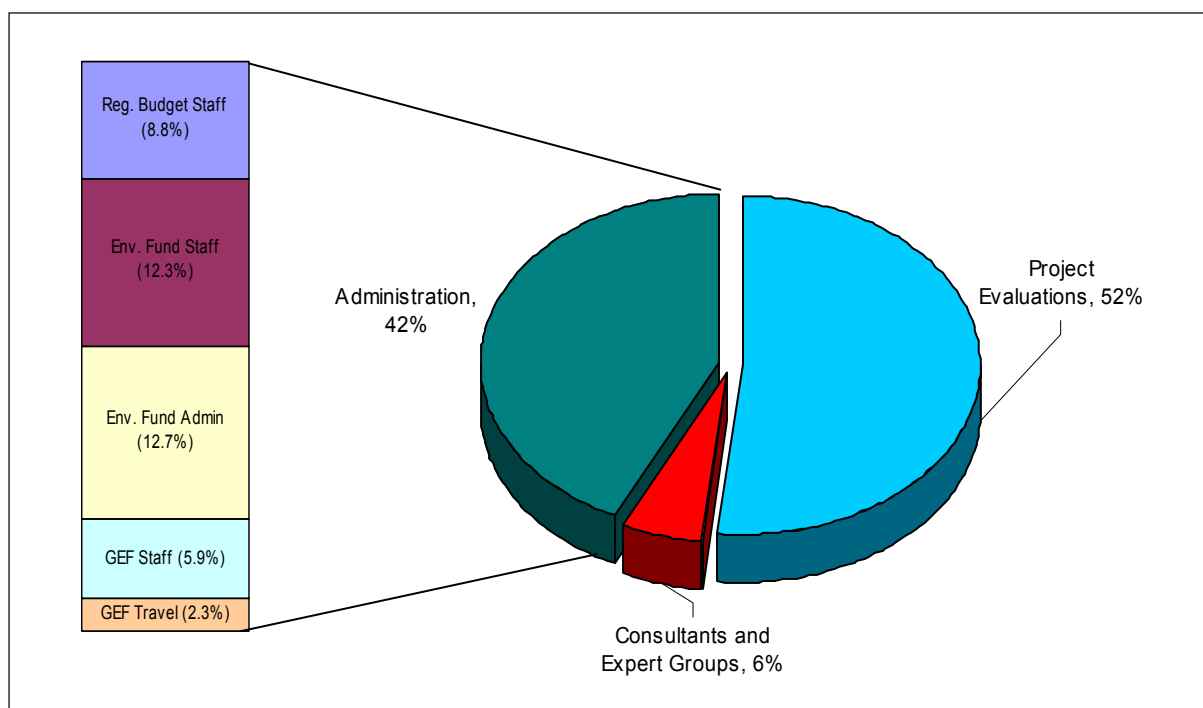
12. The staff resources of the Evaluation and Oversight Unit include a Chief of Evaluation, two Evaluation Officers, a Junior Professional Officer and three Administrative Support staff. All evaluation and oversight work related to the Environment Fund, Counterpart Contribution and Trust Funds including Global Environment Facility (GEF) activities are managed using the above staff resources. Evaluations of project activities are often supported by external experts while the set-up, administration and quality control of these evaluations and dissemination of evaluation findings are undertaken by EOU staff.

3 World Bank OED criteria for independence of the evaluation function; **Organizational Independence**: The evaluation office reports to a Board of Directors, not management, and needs no management pre-clearance of reports submitted. **Behavioural Independence**: The track record of critiquing organizational work, ability and willingness to speak out and make recommendations for improvement. **Protection from external influence**: For example, by a separate work programme and budget, unrestricted access to all records and staff; control of hiring, promotion and firing of its own staff. **Avoidance of conflict of interest**: including protection against official, financial and familial conflicts.

4 The Chief of the Evaluation and Oversight Unit is a member of Project Approval Group and the Unit also prepares a set of key issues for consideration while formulating biennium programme of work.

13. Besides staff and administration of the office, **the bulk of the resources of the evaluation function are derived from project budgets of substantive divisions.** These are used to conduct project evaluations (both mid-term and terminal), a substantial part (75%) of which is derived from the resources of the GEF and used exclusively to evaluate UNEP's GEF projects. The remaining resources are used to prepare UNEP's Annual Evaluation Report and conduct management studies. In the biennium 2006-2007, **the proportion of the budget which represents the core allocation of UNEP to evaluation is only 0.32%⁵ of the total UNEP fund programme budget.** Over the same period the proportion of project budgets allocated to evaluations are 0.88% (0.56% for activities in the programme of work and 1.03% for GEF projects). Of the \$1.16 million allocated to staff, administration and travel, approximately 27% came from the resources of the GEF and 73% from the UNEP's core budget for the biennium. Only \$150,000 of the core budget is potentially available for; the preparation of the UNEP Annual Evaluation Report, conducting management and/or thematic studies, assessment of uptake, evaluations of impact and special studies such as this. Figure 1 shows the resource utilisation by EOU.

Figure 1: Distribution of evaluation funds in UNEP



5 Calculated from the Environment Fund Budget: Proposed Biennial Programme and Support Budget for 2006-2008, UNEP/GC:23/8

Objectives of the Study

14. This study was intended to help UNEP establish a “demand-driven” orientation to the evaluation function. It aimed to gain a better understanding of what kinds of information are most relevant to the needs of donor/CPR and user (UNEP manager) audiences within the context of improving UNEP’s accountability and in the context of informing their resource allocation decisions. Similarly, it attempted to identify the types of information from evaluation products that are most useful for programme/project managers in improving the efficiency and effectiveness of their work.
15. The report is organised as follows; a methods section describes the data gathering process employed in this study, next, the results section provides details in the trends in responses, overall and by the two key stakeholder groups (CPR members and UNEP managers). Finally, these results are discussed and specific recommendations for the future focus of evaluative studies and organisation of the evaluation function in UNEP are identified.

Methods

16. The survey design for this study was adapted from Raitzer & Winkel (2005). The purpose and scope of the UNEP evaluation demand survey was broadened to include the role, scope, allocation of resources and organisational arrangements pertaining to the evaluation function. The professional staff of UNEP EOU jointly developed the survey instruments which were made available to respondents by use of an online survey service provider (www.surveymonkey.com). The survey instrument contained some questions that were mandatory for all respondents, other questions that were optional and questions that were conditional with respect to certain responses. The survey required respondents to rate or rank the importance of various factors or select statements that best reflected their views from among a number of choices. Opportunities for respondents to make additional comments were provided for each question. The respondents also had the option to fully or partially answer optional questions and, as a result, the number of complete responses varied for several of questions.
17. Survey instruments covering similar issues were prepared and tailored to the role and interests of two key audiences for UNEP evaluation products; CPR members and UNEP staff responsible for project and programme management. Invitations to complete the survey were sent to the 98 members of UNEP’s CPR and 471 UNEP’s professional staff. The EOU received 16 responses to the CPR survey and 42 responses to the UNEP Staff survey. Response rates were 16% from CPR members and 9% from UNEP Professional Staff. The response rates were lower than expected which may have been associated with the busy schedules of CPR members and UNEP staff on travel and/or annual leave during the survey period. Due to the lower response rates, the data collected from the survey were analysed using simple descriptive statistics rather than statistical ‘tests of significance’. The analysis was conducted by examining trends within each respondent group and also by pooling data from the two sources. Common trends and key differences in responses between the CPR and UNEP staff responses have been highlighted.⁶
18. In summary, the study queried the demands of a key accountability audience for evaluation—the Committee of Permanent Representatives (CPR). It also canvassed the opinions of key internal users of evaluation findings often concerned with ‘operational improvement’ issues – project and programme managers within UNEP.

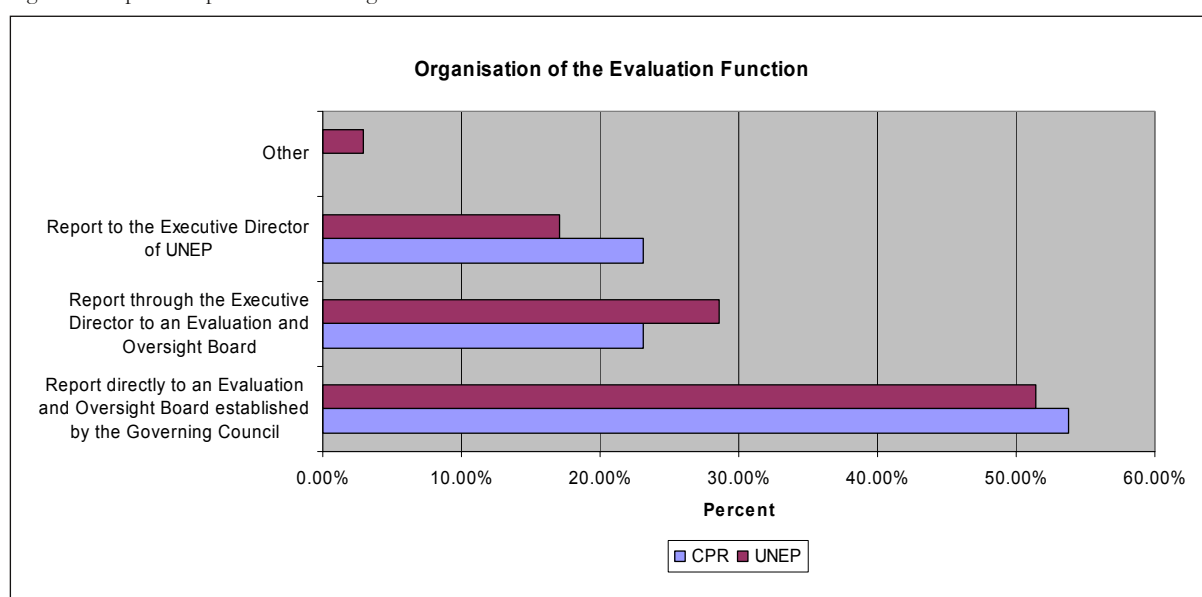
6 The full survey instrument is available from <http://www.unep.org/eou>

Findings and Discussion

Organisation of the evaluation function

19. The survey canvassed opinions from both CPR respondents and staff of UNEP as to what they considered the optimum organisational arrangement for the evaluation function in UNEP. Fifty-one (51%) per cent of UNEP respondents and 54% of CPR respondents stated that they would like to see the Evaluation function report directly to an Evaluation and Oversight Board or Committee established by the Governing Council of the organization (Figure 1).

Figure 1. Respondent preferences for organisation of the evaluation function in UNEP



20. Twenty-nine (29%) UNEP staff respondents and 23% of CPR respondents expressed a preference for the evaluation function to report through the Executive Director to an Evaluation and Oversight Board while 17% of UNEP staff respondents and 23% of CPR respondents preferred the evaluation function to continue the practice of reporting directly to the Executive Director⁷.

Implications

21. The majority of respondents would like to see improved organisational independence for the evaluation function in UNEP. This thinking is consistent not only with the recently adopted *Norms and Standards for evaluation in the United Nations system*⁸ but also with draft evaluation policy of the UN's Chief Executive Board⁹ and the World Bank's criteria for independence of the evaluation function¹⁰.

7 Currently, EOU reports to the Executive Director through the Deputy Executive Director.

8 United Nations Evaluation Group, Norms for Evaluation in the UN System, Standards for Evaluation in the UN System, April, 2005

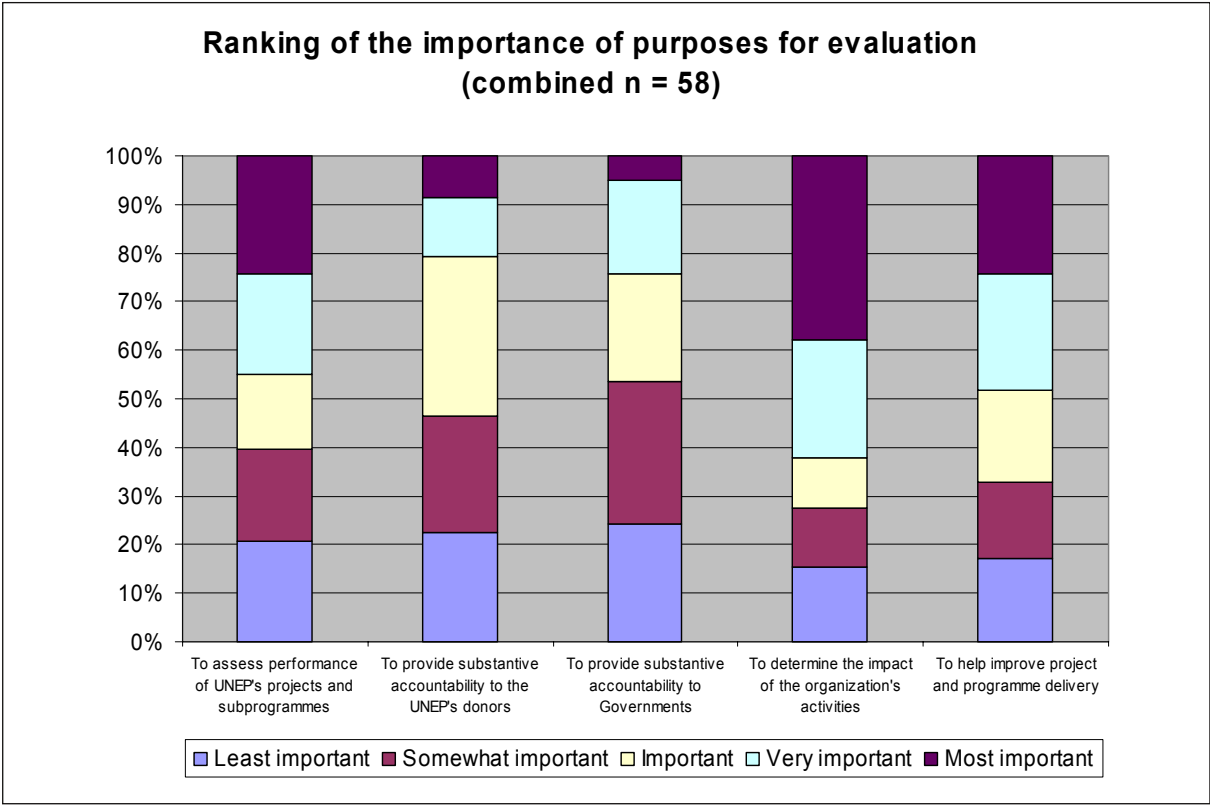
9 Draft proposal for a United Nations Chief Executives Board on Coordination policy statement on "System-wide collaboration in evaluation in the context of operational activities for development, March 2006

10 e.g. IFAD Evaluation Office reports directly to their Executive Board. IFAD Evaluation Policy (2003)

Perceptions about the purpose of evaluation

22. Using a five point scale from “not important” to “most important” respondents were asked to rank the importance of a number of evaluation purposes (Figure 2). The responses from both the internal and external survey were broadly similar. The highest degree of unanimity between UNEP and CPR respondents was found with respect to the importance attached to using evaluations to determine the impact of UNEP’s activities. Seventy-one per cent of UNEP and 75% of CPR respondents, respectively, ranked this category as either ‘important’ or ‘most important’ and 38 % of both groups ranked it as the ‘most important’ purpose for evaluative activities. Two-thirds of the respondents thought that evaluation was important, very important or most important for improving project and programme delivery.
23. Interestingly, ‘accountability to Governments’ was ranked lowest with 53% of all respondents selecting either ‘least’ or ‘somewhat’ important for this purpose. Similarly, ‘accountability to donors’ was ranked as ‘least’ or ‘somewhat’ important by 47% of all respondents - the second lowest ranking. One possible explanation for these responses is that ‘accountability to donors/ governments’ could be subsumed under other evaluation purposes that focus on ‘the nature of the purpose’ (e.g. to determine impact, to enhance project / programme delivery) rather than the intended accountability ‘target audiences’ (governments/donors).

Figure 2. Responses ranking the importance of different evaluation purposes



Implications

24. UNEP evaluations need to increase the focus on demonstrating improved programme/ project delivery and on evaluating the outcomes and impacts of UNEP activities. This focus would largely address other intended purposes of evaluations, for example, accountability to donors.

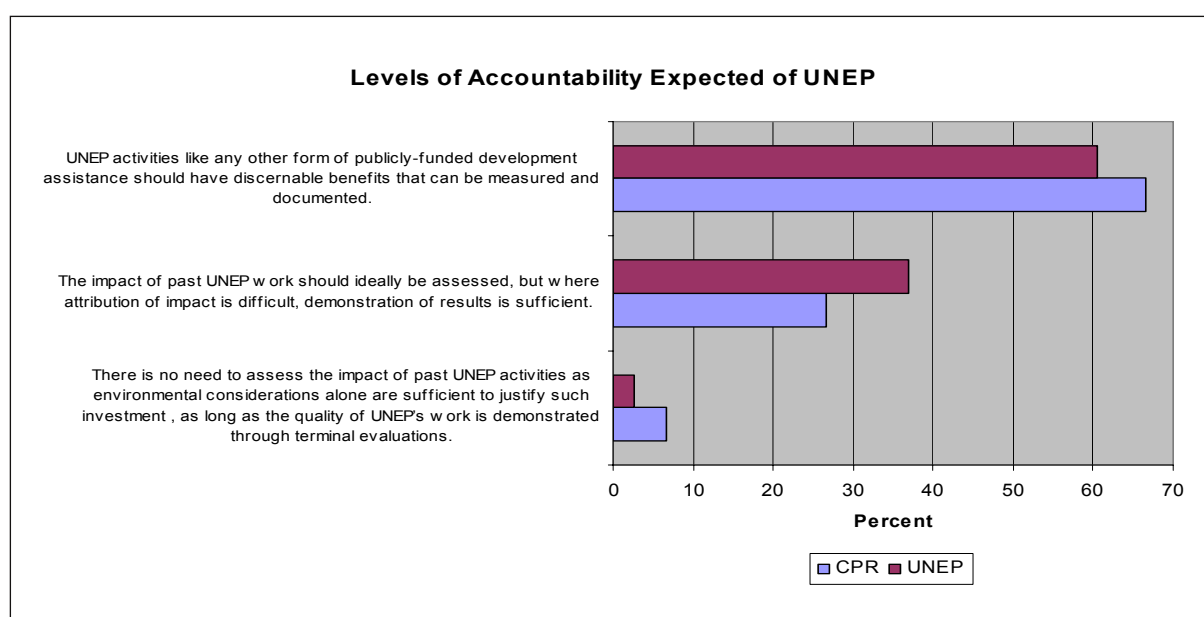
Levels of accountability expected of UNEP

25. To explore perceptions regarding the levels of accountability expected of UNEP, respondents were asked to select the statement that best reflected their view (Figure 3). Three-fifths of UNEP respondents and two-thirds of the CPR agreed that *'UNEP activities like any other forms of publicly funded development assistance should have discernable benefits that should be documented.'* The survey seemed to reject the view that *'environmental considerations alone justify investments in UNEP's activities and therefore there is no need to invest efforts in determining the impacts of these activities.'* Approximately 27% of the CPR and 37% of the UNEP respondents recognized the difficulties in attempting to determine impacts of environmental activities; however, there is an overwhelming consensus on the need to clearly document impact (benefits) stemming from UNEP activities.

Implications

26. Both donors and UNEP programme managers would like to see evaluative evidence regarding the results and benefits stemming from UNEP activities. This demand has not been addressed and more resources need to be allocated for this purpose.

Figure 3. Expectations for UNEP's accountability

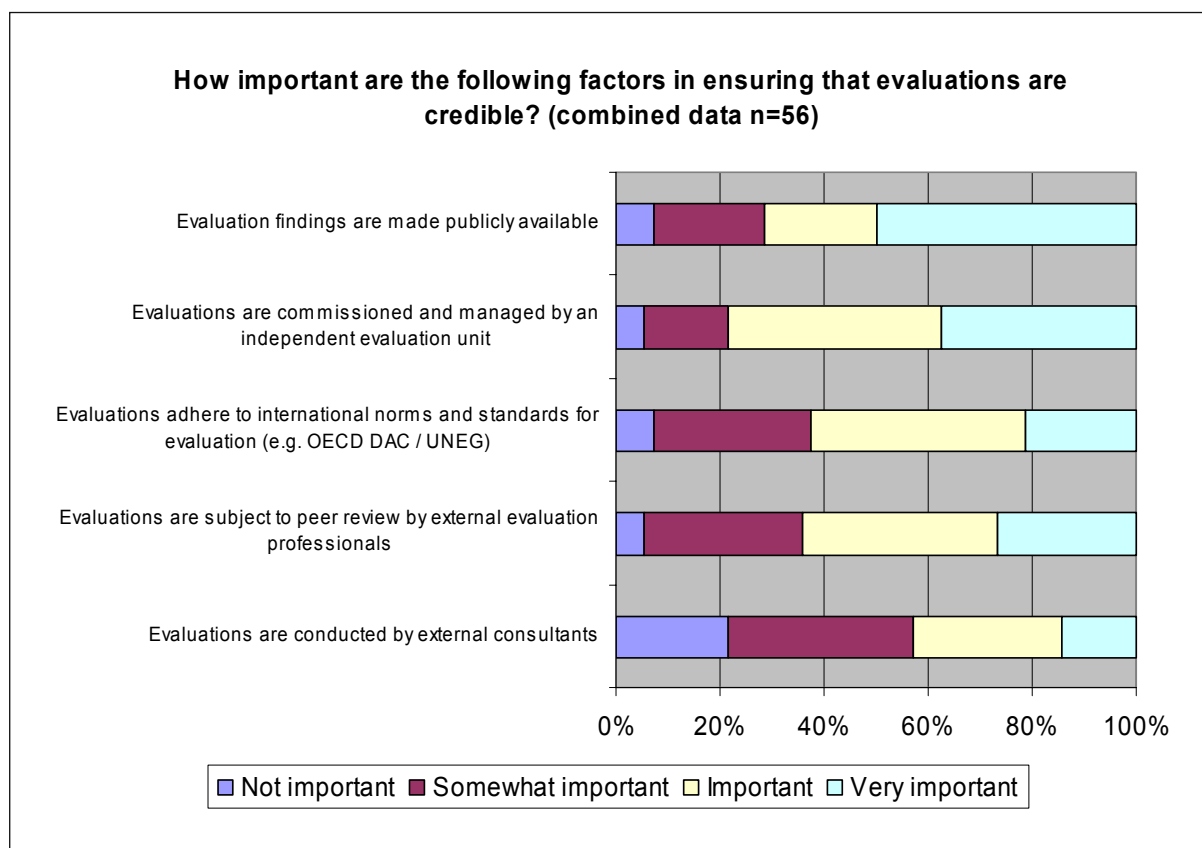


Factors affecting the credibility of evaluations

27. Whether potential users of evaluation findings regard UNEP evaluations as credible is key to ensuring that UNEP's evaluation function fulfils its purpose. Findings from evaluations that lack credibility are unlikely to be accepted, and less likely to be used, undermining the main evaluation purposes of enhancing 'accountability' or 'operational improvement'. Both UNEP and CPR survey respondents were asked to rate each of the factors shown in Figure 4 using a four point scale from "not important" to "very important". Almost 50% and 53% UNEP and the CPR respondents, respectively, regarded public disclosure of evaluation findings as "very important" for ensuring that evaluations are credible and 71% of all respondents regarded this factor as being either 'important' or 'very important'. Only 7% rated this as 'not important'.

28. Across all respondents, over 63% regarded compliance with international evaluation standards as either 'important' or 'very important'. Nearly four-fifths of respondents (78%) attached the same levels of importance, with regard to credibility, for evaluations being commissioned and managed by an independent evaluation unit. Similarly, 64% regarded 'peer review of evaluations' as 'important' or 'very important' for evaluation credibility.
29. However, over 57% of all respondents thought that the use of 'external consultants' was either 'somewhat important' or 'not important' for evaluation credibility. The corresponding percentage among CPR respondents was over 66%.

Figure 4. Factors affecting evaluation credibility



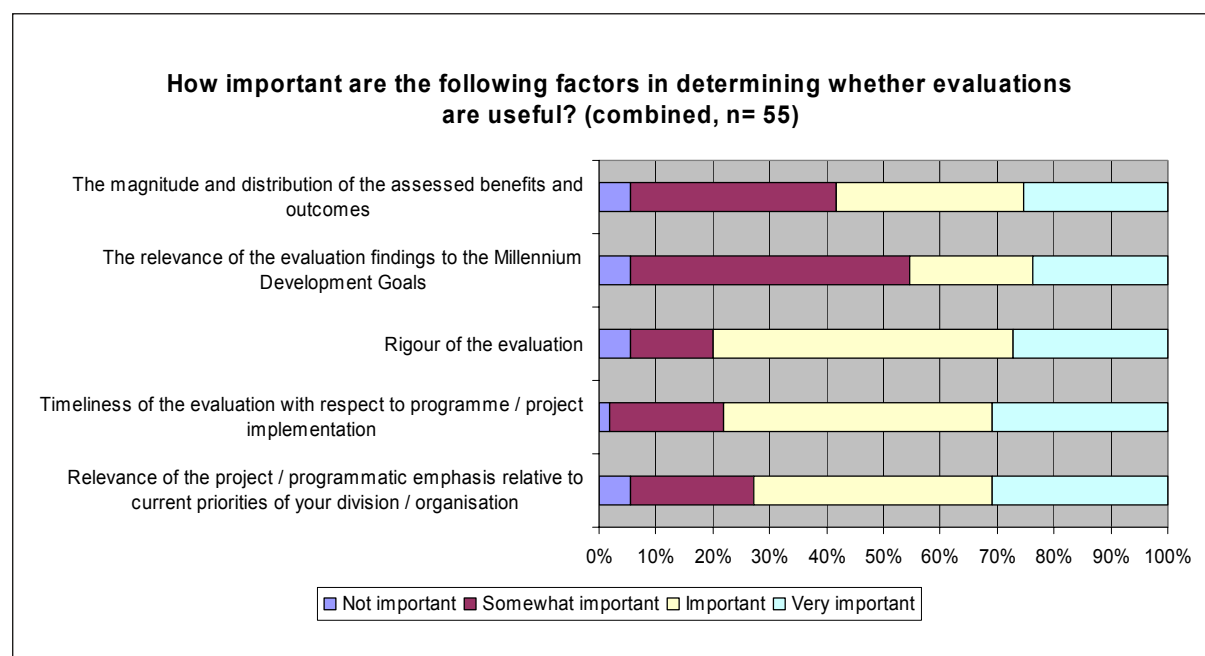
Implications

30. Public disclosure of evaluation findings, application of international norms for evaluation standards, peer review of evaluation products and the independence of the evaluation function appear to be the key elements to ensure the credibility of UNEP evaluations among these stakeholder groups. Furthermore, if the evaluation function is perceived to be independent the need for the use of external evaluation consultants, as a means to enhance credibility, seems to be less important. The underlying assumption here is that an independent evaluation function would be adequately staffed and resourced to conduct evaluations that meet international standards and rigour.

Factors affecting the utility of evaluations

31. Ensuring that evaluation products are useful to their intended audiences is of key importance if the work of EOU is to become truly ‘demand-driven’. To better understand what factors influence the utility of evaluation products, survey recipients were asked to rate the importance of a number of factors (Figure 5).

Figure 5. Factors affecting the utility of evaluation products.



32. CPR respondents regarded the ‘timeliness’ of evaluations as the least important factor affecting their utility among the choices offered. Nevertheless, 60% of these respondents regarded ‘timeliness’ as either ‘important’ or ‘very important’. By contrast, 85% of UNEP respondents regarded ‘timeliness’ as either ‘important’ or ‘very important’ for evaluation utility. One explanation for this difference may be that the CPR respondents are more concerned with accountability-related findings, which are less ephemeral, whereas UNEP respondents are more concerned with the operational utility of evaluation findings which must be made rapidly available to be of operational utility.
33. This explanation is supported, in part, by the fact that CPR respondents regard the ‘*magnitude and distribution of assessed benefits*’, a key accountability issue, as being a more important factor for evaluation utility than do UNEP respondents; 66% and 55%, respectively, rating this factor as ‘important’ or ‘very important’.
34. CPR respondents also attached greater importance to ‘*the relevance of the evaluation findings to Millennium Development Goals (MDGs)*’ (47% ‘very important’) than did the internal respondents (15% ‘very important’) in affecting evaluation utility. Generally, CPR respondents perceived all the specified ‘utility factors’ to be important, and the proportion of such respondents rating these factors ‘important’ or ‘very important’ ranged from 60% - 73%.

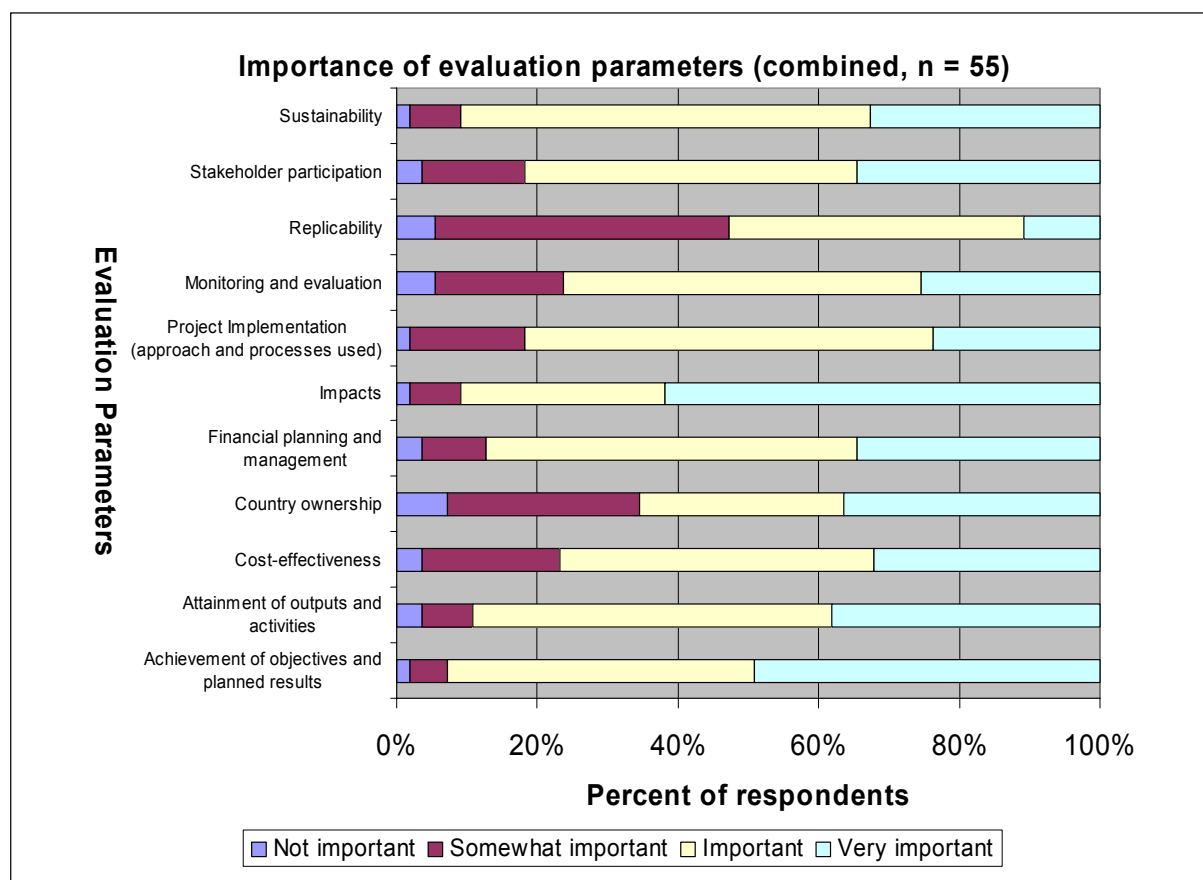
Implications

35. The most important factors affecting the utility of evaluations appear to be *'timeliness', 'rigour' and 'relevance to current organisational priorities'*. EOU needs to ensure that that quality control measures applied to UNEP evaluations address these key issues. For CPR audiences, the relevance of evaluation findings to the MDGs needs to be articulated.

The importance of different evaluative parameters

36. UNEP evaluations commonly address a range of parameters that are designed to capture a wide spectrum of project / programme performance measures. Respondents were asked to rate the importance of a number of evaluation parameters that are regularly considered in UNEP evaluations (Table 1 and Figure 6).
37. Both groups of respondents regarded the evaluation parameters of *'replicability' and 'country ownership'* as the least important. The lower importance afforded to these parameters is perhaps because they were considered by some respondents as subsidiary elements of *'sustainability'* and as such were afforded a lower rating. Somewhat surprisingly, *'cost effectiveness'* was rated as more important among the UNEP rather than the CPR respondents (83% and 60% respectively). However, UNEP respondents collectively regarded *'achievement of objectives', 'impacts', 'sustainability' and 'attainment of outputs'* as the four most important parameters. This showed some consistency with the perceptions expressed by CPR respondents for whom the four most important evaluation parameters were: *'achievement of objectives', 'sustainability', 'financial planning and management' and 'project implementation approach'*.

Figure 6. The relative importance of different evaluation parameters



38. The largest difference in the perceptions of importance of evaluation parameters were associated with *'cost effectiveness'*, *'impacts'* and *'attainment of outputs'*. Surprisingly, *'impact'* (a key accountability issue) was regarded as an important parameter among a greater proportion of UNEP rather than CPR respondents.

Table 1. Percentage of respondents rating evaluation parameters as 'important' or 'very important'

UNEP Evaluation Parameters	UNEP % (n=40)	CPR % (n=15)	Combined % (n=55)
Achievement of objectives and planned results	95	87	93
Attainment of outputs and activities	93	80	89
Cost-effectiveness	83	60	77
Country ownership	65	67	65
Financial planning and management	88	87	87
Impacts/Results	95	80	91
Project Implementation (approach and processes used)	80	87	82
Monitoring and evaluation	75	80	76
Replicability	55	47	53
Stakeholder participation	83	80	82
Sustainability	93	87	91

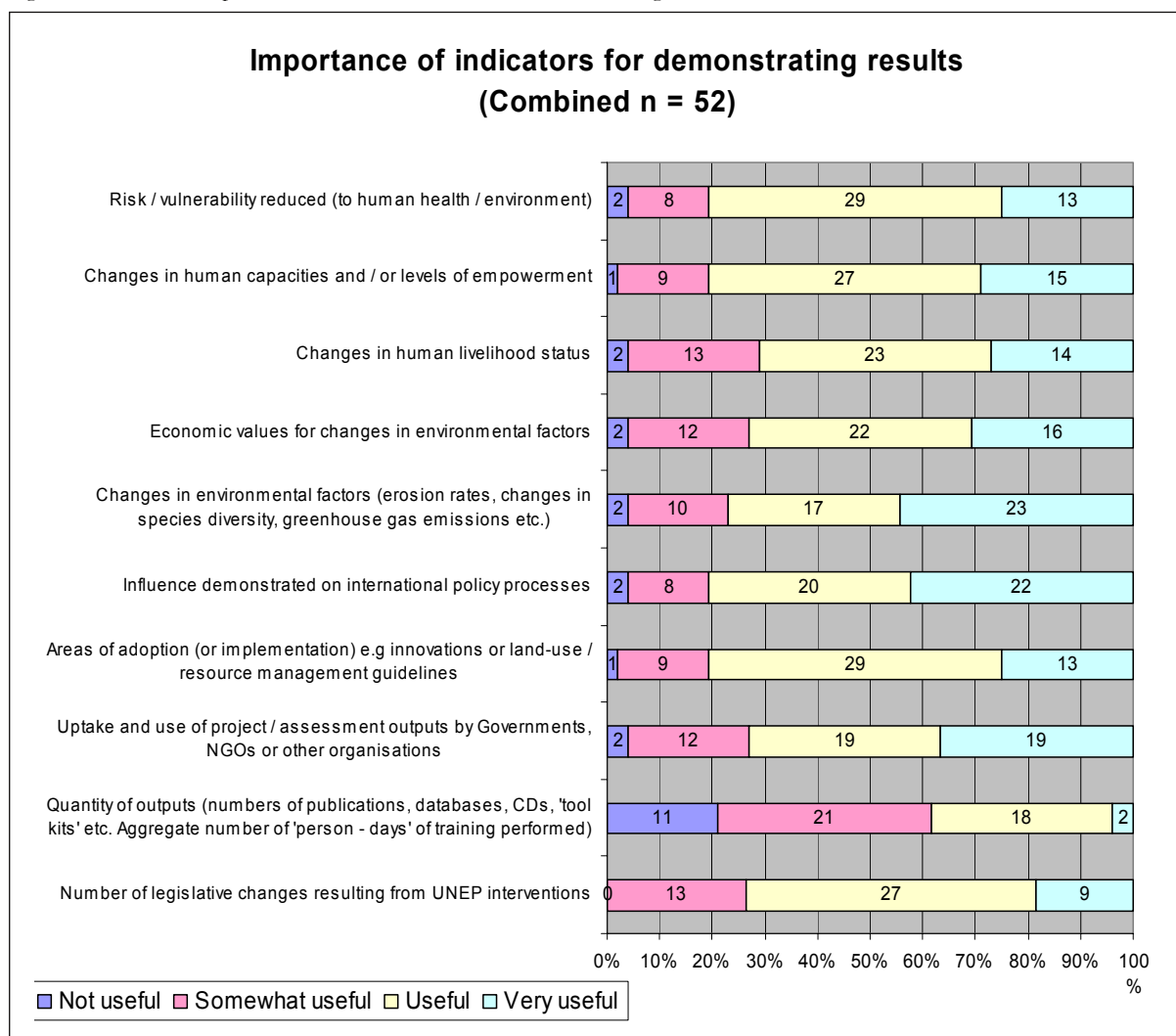
Implications

39. The importance of evaluation parameters used in UNEP evaluations have been reaffirmed by both groups of respondents (CPR members and UNEP staff) and hence these parameters should continue to form the basis for future UNEP evaluations. Although there was relatively less recognition of the country ownership and replicability parameters by the respondents, both of these parameters are key elements in ensuring sustainability (a parameter recognised to be important by 91 per cent of the respondents).

The importance of indicators for demonstrating the results of UNEP activities

40. The respondents were asked to rate the importance of ten selected indicators for demonstrating the results of UNEP activities. The indicators were chosen based on an analysis of UNEP's mission and goals including standard evaluative practices. The responses were recorded as 'not useful', 'somewhat useful', 'useful' or 'very useful'. Figure 7 summarizes the responses. The largest proportion of respondents considered the following indicators as 'useful' or 'very useful' for assessing the impact of UNEP activities (a) *'areas of adoption (or implementation) of innovations e.g. land-use / resource management guidelines'*; (b) *'Influence demonstrated on international policy processes'*; (c) *'changes in human capacities and/or level of empowerment'*; and (d) *'Risk / vulnerability reduced (to human health / environment)'*. These were closely followed by *'changes in environmental factors (erosion rates changes in species diversity greenhouse gas emissions etc.)'* and *'uptake and use of project / assessment outputs by Governments NGOs or other organizations'*. On the other hand, only 38 per cent of the respondents thought that *'quantity of outputs (numbers of publications databases CDs 'tool kits' etc. aggregate number of 'person - days' of training performed)'* was a 'useful' or 'very useful' indicator to assess the results stemming from UNEP activities.

Figure 7. The relative importance of different indicators for demonstrating results



41. The relative importance of the indicators was consistent between the two groups of respondents (CPR and UNEP professional staff) with a few exceptions (Table 2). For example, *'the number of legislative changes'* was considered to be a 'useful' or 'very useful' indicator by 78 per cent of UNEP staff respondents in contrast with only 47 per cent of CPR respondents. With the exception of *'quantity of outputs'*, at three-quarters or more, UNEP respondents considered the stated indicators 'useful or very useful'. On the other hand, the CPR respondents viewed *'reduced risk or vulnerability to human health and environment'* and *'influence demonstrated on international policy processes'* to be leading indicators as 'useful' or 'very useful', closely followed by *'changes in human capacities and/or levels of empowerment'* and *'economic values for changes in environmental factors'*.

Table 2: Percent of respondents expressing result indicators as ‘important’ or ‘very important’

Result Indicator	CPR respondents	UNEP respondents
Number of legislative changes	47%	78%
Quantity of outputs	47%	35%
Uptake and use of project / assessment outputs	60%	78%
Areas of adoption (or implementation) of innovations	67%	86% (1 st)
Influence demonstrated on international policy processes	80% (2 nd)	81% (3 rd)
Changes in environmental factors	67%	81% (3 rd)
Economic values for changes in environmental factors	73% (3 rd)	73%
Changes in human livelihood status	60%	76%
Changes in human capacities and/or levels of empowerment	73% (3 rd)	84% (2 nd)
Risk / vulnerability reduced (to human health / environment)	87% (1 st)	78%

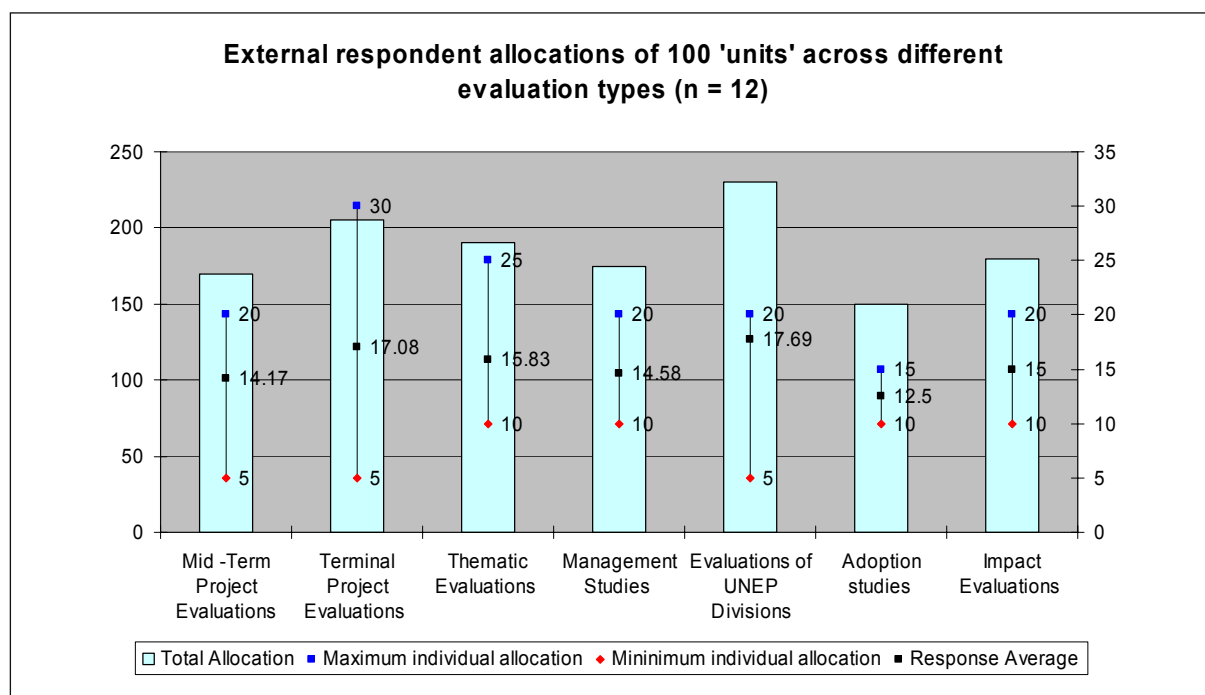
Implications:

42. The results suggest that UNEP needs to place greater emphasis on impact indicators such as ‘*reduced risk and vulnerability*’, ‘*influence on international environmental policy processes*’, ‘*changes in human capacities and/or levels of empowerment*’, ‘*uptake and use of project/assessment outputs*’ and ‘*economic valuation of changes in environmental factors*’. Indicators relating to production of quantitative outputs should be regarded as being of lower importance in assessing the performance of UNEP projects / programmes. The findings from the survey are consistent with the notion that, in addition to technical parameters, project results need to take into account environmental, social and economic dimensions. Furthermore, results-based indicators need to be incorporated in UNEP programme and project evaluations. For this to happen, it is important that such indicators be clearly identified for monitoring purposes and associated baseline data be established early in programmed activities.

Allocation of resources across different types of evaluation

43. Evaluation takes place within the context of limited resources, which means that it is not feasible to comprehensively evaluate every activity of UNEP. In addition, certain types of evaluations require greater resources than others but are associated with different levels of accountability. In this context, respondents were asked to allocate one hundred notional ‘resource units’ across the different types of evaluative activity that could be undertaken by UNEP EOU with the aim of reflecting their relative importance.
44. Aggregate resource allocation patterns within both CPR and UNEP respondent groups were fairly similar. CPR respondents allocated resources across all evaluation types, whereas UNEP respondents tended to be more selective, preferring to allocate resources to some evaluation types to the exclusion of others (26%). UNEP respondents more frequently allocated resources to ‘Terminal Evaluations’ and ‘Impact Assessments’ than did the CPR respondents.

Figure 8. Aggregate CPR allocation patterns for 100 notional 'resource units' across different evaluation activities.



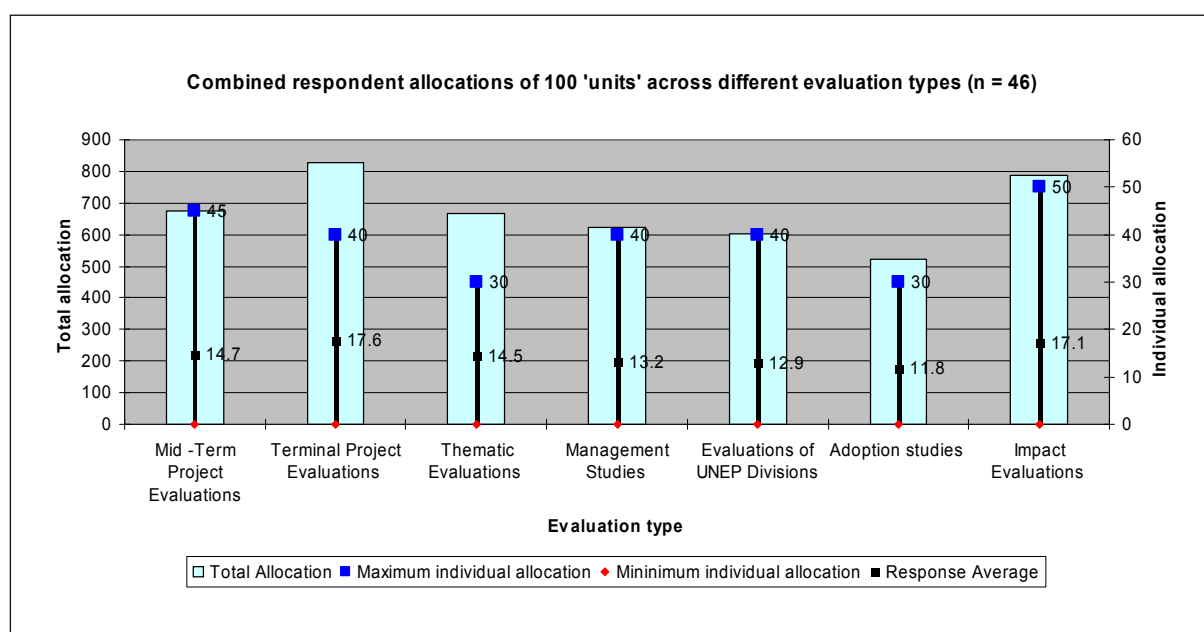
45. CPR respondents allocated the largest aggregate share of resources to Divisional (UNEP sub-programme) evaluations, but no individual respondent allocated more than one fifth, or less than one tenth, of the available resources. Terminal evaluations received the second largest aggregate share of resources from CPR respondents, with individual respondents allocating a maximum of one third, and a minimum of one twentieth, of the available resources. Average allocations for both of these evaluation types were very similar.

Implications

46. The survey findings show that both key stakeholder groups express a considerable demand for studies that, due to current resource constraints, are not conducted by UNEP EOU; namely, 'adoption studies'¹¹ and 'impact evaluations'. However, there is also continued high demand for the types of evaluation that EOU has conducted to date. To fully satisfy the expressed demands would require a considerable expansion of evaluation activities supported by necessary resources, including staffing at a satisfactory level.

¹¹ 'adoption studies' establish the uptake and use of new knowledge, proven technologies and / or management practices among key target audiences.

Figure 9. Aggregate combined (CPR & UNEP) allocation patterns for 100 notional ‘resource units’ across different evaluation activities.



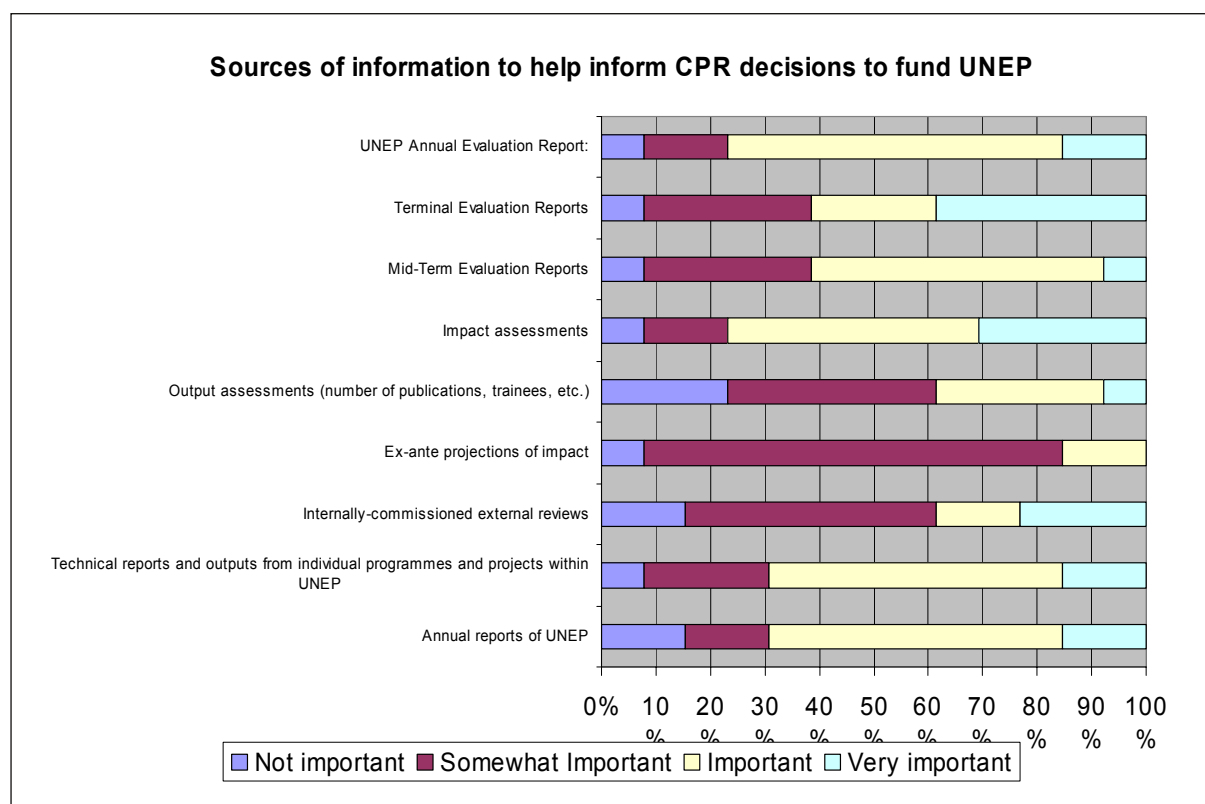
Perceptions on utility of UNEP evaluative products for resource allocation decisions

47. CPR members (who represent governments) were asked to rate the relative importance of different sources of information in terms of their influence on (or utility for informing) decisions for funding UNEP (Figure 10). The responses suggest that the least useful sources of information include ‘*ex-ante impact projections*’, ‘*output assessments (numbers of publications trainees etc.)*’ and ‘*internally commissioned external reviews.*’ Respondents selected the ‘*UNEP Annual Evaluation reports*’, and ‘*impact evaluations*’ as the two most useful sources of information with 77% of respondents regarding these as either ‘important’ or ‘very important’. ‘*Technical reports for specific projects and programme’s*’, along with the ‘*UNEP Annual Report*’ were both rated by 69% of respondents as being ‘important’ or very important’ for informing funding decisions.
48. However, the nature of the ‘use’ of these products in funding decisions remains unclear. Additional comments provided by respondents suggest that ‘influence’¹² or ‘symbolic use’¹³ may be the most frequent means by which evaluations affect UNEP funding decisions

12 The term “influence” recognises that evaluation can incrementally contribute to the body of knowledge surrounding a programme or policy and gradually inform the context under which programmatic decisions are taken and there are often indirect pathways through which evaluation results affect the actions of decision-makers.

13 Symbolic use refers to selective application of evaluation findings, so as to support a preordained policy preference (Leviton and Hughes, 1981). Hence, such use actually makes little or no difference to programmatic implementation, and only serves to make prior policy preferences appear as the result of rational deliberative processes.

Figure 10. Sources of information to help inform funding decisions.



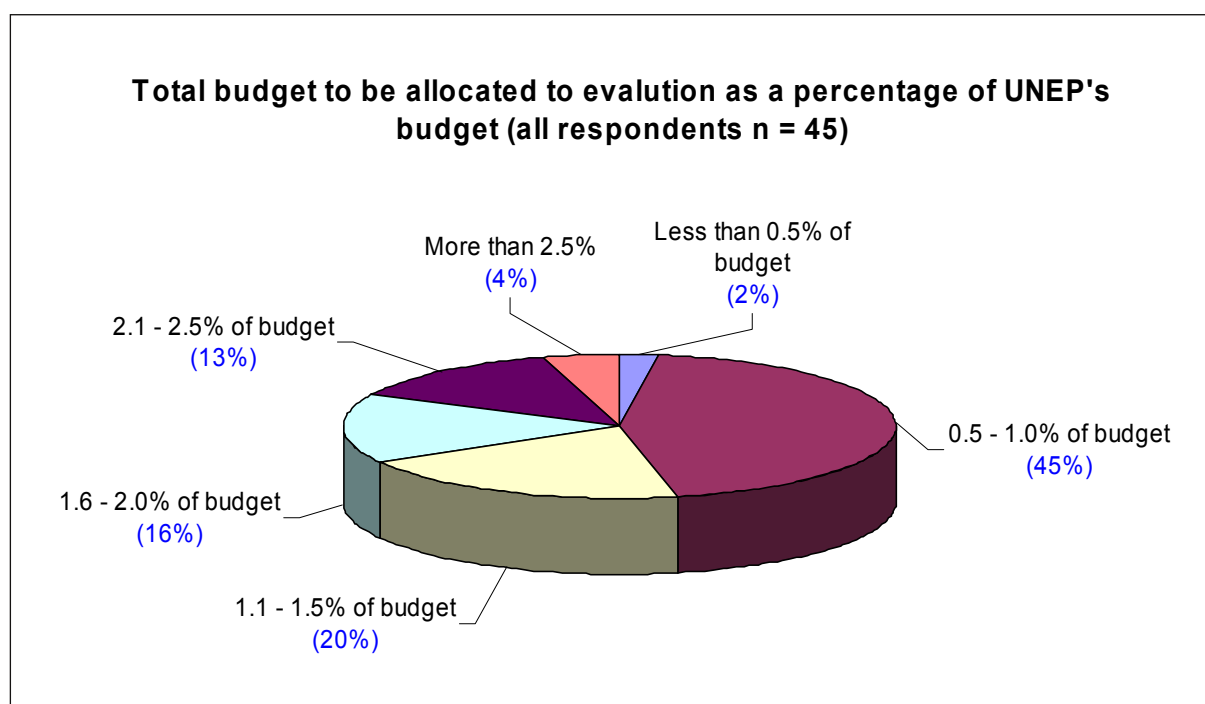
Implications

49. The survey reinforces the demand for evaluation activities currently undertaken by the evaluation function. In addition, it also suggests that both UNEP Staff/management as well as members of the Governing bodies would like to see evidence-based evaluations which determine impacts or benefits of UNEP activities no matter how difficult the determination of such impacts or benefit streams would be.

Perceptions about resourcing evaluation and oversight activities

50. The importance of evaluation and oversight activities in relation to other project and programme activities are either often misunderstood or underestimated. As a result, the evaluation and oversight activities tend to be of limited nature due to lack of a clear buy-in by the project and programme partners and the management. While there is no clear benchmark as to what should be the optimal level of resources allocated for evaluation and oversight activities, it is important to understand perceptions reflecting the willingness to support such activities. The survey specifically asked the respondents to reveal their opinions (which we assume are based on their perceptions of the importance of evaluation and oversight functions) as to what should be the organisation's evaluation and oversight budget in relation to (a) organisation's budget and (b) project budget. Based on the categorical responses, the results are presented in Figures 11 and 12, respectively.
51. Forty-five per cent of respondents suggested that UNEP should allocate between 0.5 to 1.0 percent of the organisation's budget (Figure 11). One in five respondents considered that resources for evaluation and oversight should account for between 1.1 and 1.5 per cent of UNEP budget and one in six expressed the figure should be between 1.6 and 2.0 per cent.

Figure 11. Preferences for budgetary allocation to evaluation and oversight as percentage of the total UNEP budget.



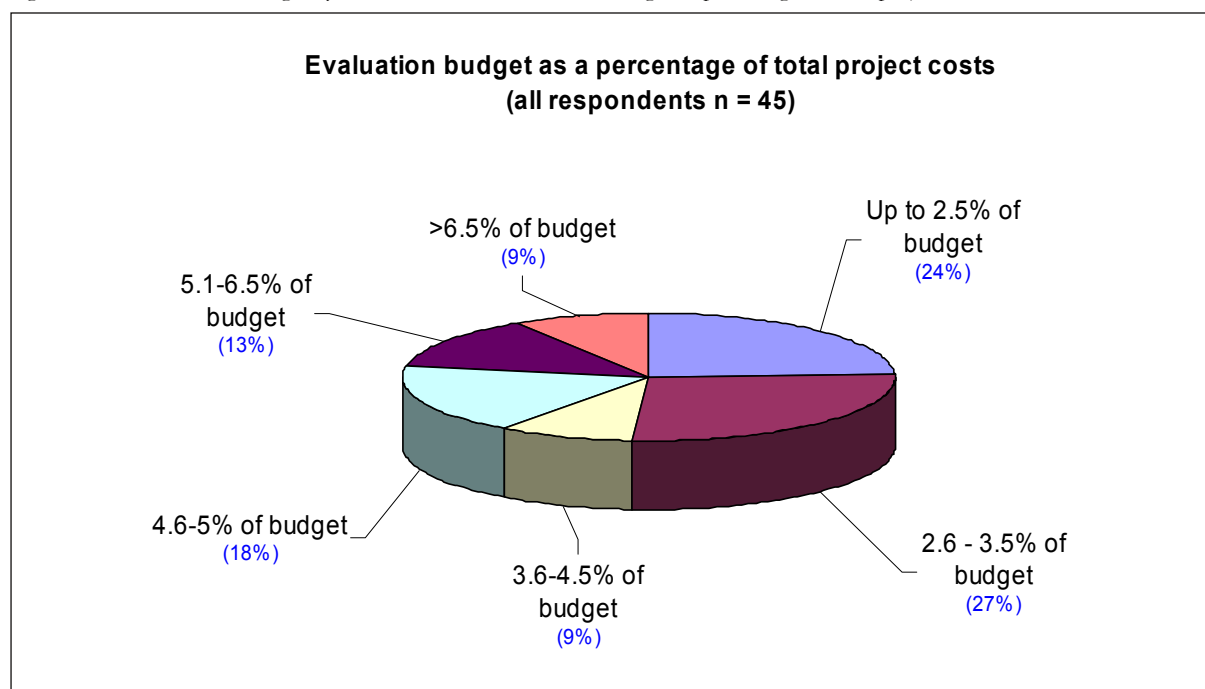
52. There was, however, a remarkable difference in the responses between the CPR and UNEP respondents (Table 3). Overall, the CPR respondents indicated that more resources should be allocated to the evaluation function than did the UNEP staff. Thirty-six per cent of CPR respondents indicated that the evaluation budget should account for between 1.6 to 2 per cent of UNEPs' total budget while half of the UNEP respondents considered the level of funding for evaluation between 0.5 and 1.0 per cent of UNEP budget.

Table 3: Respondent preferences for the level of the evaluation and oversight budget in proportion to UNEP budget (Per cent respondents)

Evaluation budget (% of UNEP budget)	CPR respondents	UNEP respondents
Less than 0.5%	0%	3%
0.5 - 1.0%	27%	50%
1.1 - 1.5%	18%	21%
1.6 - 2.0%	36%	9%
2.1 - 2.5%	18%	12%
More than 2.5%	0%	6%

53. Responses to the question regarding perceptions about the size of evaluation budget in proportion to the project costs are summarized in Figure 13. Slightly over one-quarter (27 per cent) of the respondents indicated that between 2.5 to 3.5 per cent of project budget should be allocated for evaluation activities whilst nearly one-quarter (24 per cent) suggested that the maximum allocation should be up to 2.5 per cent of the project budget. Twenty-two per cent of the respondents, on the other hand, supported more than 5 per cent of project costs for evaluation activities.

Figure 12. Preferences for budgetary allocation to evaluation and oversight as percentage of total project costs.



54. The two groups of respondents differed in their opinion in response to the question on evaluation budget in relation to project costs. Overall, the CPR respondents were supportive of higher level of funding for project evaluation – 54 per cent thought that evaluation budget should be between 3.6 and 5 per cent. On the other hand, the UNEP respondents were relatively more conservative and one-fourth considered evaluation budget up to 2.5 per cent of project costs and another 32 per cent of respondents supported between 2.6 to 3.5 per cent of project costs. Limited support was found for evaluation budget beyond 5 per cent of the project costs. The range of opinion reflected in responses to this question may reflect the size and variation in the complexity of project activities – which, in turn, implies a range of evaluation costs.

Table 4: Perceived evaluation budget in proportion to UNEP budget (Per cent respondents)

Evaluation budget (% of project costs)	CPR respondents	UNEP respondents
Up to 2.5%	18%	26%
2.6 - 3.5%	9%	32%
3.6-4.5%	27%	3%
4.6-5%	27%	15%
5.1-6.5%	18%	12%
>6.5%	0%	12%

Implications:

55. There is strong support for dedicated budget for evaluation both at the organizational and project level. Moreover, the relative importance of the evaluation function in UNEP is duly recognized by both external (CPR) and internal (UNEP project managers) stakeholders and is reflected by an overall support for evaluation budgets that exceed current resource allocation patterns. The expectations of both external and internal stakeholders for evaluation products that meet international standards means that the evaluation function needs to be adequately resourced to fulfil its mandate.

Readership of evaluation products

56. Eighty-two per cent of UNEP respondents and 90 per cent of CPR respondents indicated that they had read a UNEP evaluation product. Those respondents that had read an evaluation product were asked to provide a rough estimate of the number of UNEP evaluation studies read over the last 2 years.
57. For the CPR respondents the survey captures a number of potential anomalies. These anomalies may have been as a result of different understanding and/or interpretation of various evaluation products. For example, according to these respondents, the average number of impact evaluations read was 3 while UNEP has, in fact, only produced one impact evaluation since 1997. Also, the average number of annual and sub-programme evaluations that were read was 3 which, if correct, would suggest that there is a 100 per cent readership of reports produced in the specified timeframe. However, sub-programme evaluations are normally only shared externally upon request so this is unlikely to be the case. Finally, UNEP has not produced any thematic evaluations, however according to the external respondents the average number read was two.
58. The responses from UNEP staff show more consistency with what has actually been produced by EOU over the past two years. The average number of annual and sub-programme evaluation reports read was two, which matches the number produced over the past biennium suggesting a 100% readership. The average number of terminal evaluations that respondents reported having read was 5.
59. Sixty-two percent of CPR respondents reported accessing EOU evaluation reports as printed copies and 38% accessed them through email. None made use of EOU's web-site. This finding is not surprising because the EOU website link does not appear on the main UNEP web page. Within UNEP, evaluation reports were mainly accessed by email (46%) closely followed by printed copies 42 % but some, albeit few, were accessed through EOU's website 12%. Relatively low use of the EOU website may be associated with inadequate familiarity.

Perceptions on the quality of UNEP evaluation reports

60. Survey respondents were asked to rate the quality of UNEP evaluation products in terms of their '*credibility*' / *rigour*', '*comprehensiveness of coverage*' and their '*relevance and utility for your organisation's needs / interests*'. Overall, 54% of respondents rated the credibility/rigour of UNEP evaluation reports as 'good'. Eleven percent of CPR respondents rated the 'credibility / rigour' as 'poor' (against zero in UNEP) whilst 36 % of UNEP staff rated it as 'fair'. Even though broadly positive these results show that there is still room for improvement in the rigour of EOU evaluations.
61. There also seemed to be unanimity among CPR and UNEP respondents regarding the '*comprehensiveness of coverage*' of UNEP evaluations with 77% and 72 % respectively rating the coverage as either 'good' or 'excellent'. Again, there were no UNEP respondents rating the coverage as 'poor', 22% of CPR respondents though rated the '*comprehensiveness of coverage*' of UNEP evaluations as either 'poor' or 'fair'. Correspondingly, 28 % of UNEP respondents thought that the coverage was 'fair'.
62. The largest differences in perceptions between survey groups regarding EOU report quality were in relation to the '*relevance and utility of evaluations to your organization's needs/interests*'. While 22% of the CPR respondents thought that evaluations reports were of 'poor' relevance to their organizations needs 77% of them thought that it was between 'good' an 'excellent' which is a large difference and maybe related to the different organizations (or types of organization) responding. Within UNEP, none of the respondents rated the relevance as 'poor' but 32 % rated it as 'fair' and 48 % as either 'good' or 'excellent'.

Sharing of UNEP evaluation products

63. All respondents appear to share evaluation reports with others. Among CPR respondents, half reported having shared evaluation documents with a few other staff and the other half were split between sharing documents with more than 5 staff and with senior management. Within UNEP, respondents reported that a combined 77% of evaluation products are shared within their respective divisions, terminal evaluations being the most frequently circulated to professional staff in the divisions (80%). The sub-programme evaluations are not frequently shared outside of UNEP whereas mid-term and terminal evaluations are fairly well distributed outside the organization. This is because the sub-programme evaluations are neither widely publicised nor uploaded on the UNEP/EOU website.

Implications

64. EOU needs to make an extra effort to sensitise CPR members about the different types of evaluation products it produces so that they are understood and interpreted appropriately. Even though CPR and UNEP staff perceptions of the quality of EOU evaluations are broadly positive, there is still room for improvement in the rigour of EOU's work. The survey findings also highlight low levels of usage of the EOU website suggesting a need raise awareness of its existence. The evaluation website should appear on the main UNEP page for easier access.

Conclusion

65. The study draws several conclusions from findings of the online survey of key stakeholders and the results reinforce positions articulated in the Evaluation Policy. First, evaluation has an important role to play in responding to increased demands for accountability by providing information on programme results and the impact of UNEP's activities. There is broad recognition that evaluation can also help identify where improvements can be made to project and programme delivery. Secondly, in order to gain or maintain credibility, the evaluation function must be perceived to be functionally independent of the organisation's operational divisions. To that end, the study demonstrated that a strong link between the evaluation function and the organizations Governing Body is required. Thirdly, the scope of UNEP's evaluation activities should expand to cover evaluation of discernable benefits from the implementation of its activities; this is regarded as useful information for informed funding decisions. Fourth, UNEP evaluations need to apply international norms and standards for evaluation, to enhance their credibility and legitimacy. Fifth, evaluations must be relevant and produced on a timely basis to inform decision-making and aid the development and implementation of programme activities. Finally, the study highlights that resources currently allocated to the evaluation function are not sufficient to meet the demands for evaluative products and information expressed by survey respondents; and hence calls for additional measures to strengthen the function.

REFERENCES

- Wenar, L. (2006). Accountability in International Development Aid. *Ethics & International Affairs* (1), 1-23.
- Raitzer, D. and Winkel, K. (2005). Donor Demands and Uses for Evidence Of Research Impact – The Case of the Consultative Group on International Agricultural Research (Standing Panel on Impact Assessment, CGIAR, Science Council Secretariat, Food and Agriculture Organization of The United Nations, Rome, Italy 56p.

