

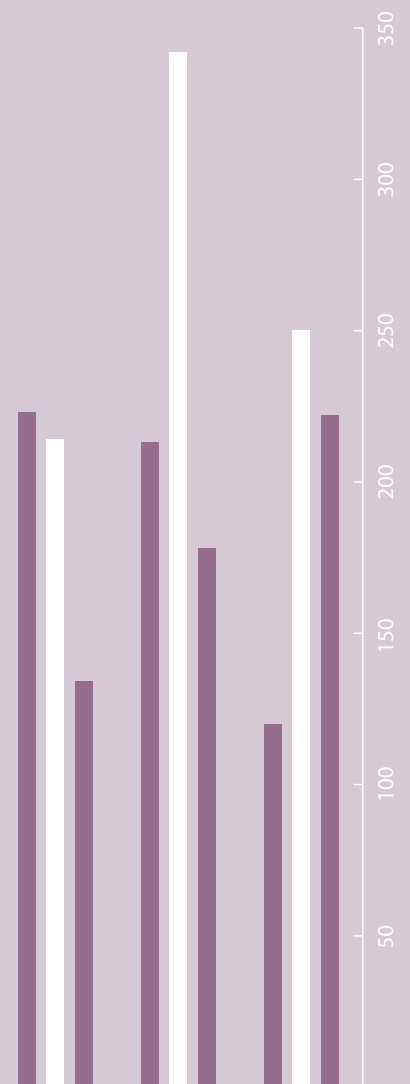
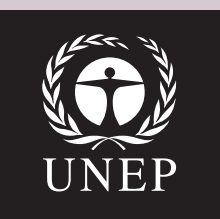
Priority Setting for Evaluation: Developing a strategic evaluation portfolio

M.J Spilsbury, S.Norgbey and C. Battaglino

Special Study Papers No. 3

Evaluation Office
United Nations Environment Programme

January 2010





United Nations Environment Programme

**Priority Setting for Evaluation:
Developing a strategic evaluation portfolio**

M.J Spilsbury, S.Norgbey and C. Battaglino

Special Study Papers No. 3

UNEP Evaluation Office

January 2010

Table of Contents

Acronyms	iv
1- Introduction	1
2- Priority setting Framework	1
3- Priority setting method using ‘evaluation opportunities’	7
4 - Analysis and discussion of the evaluation portfolio for 2008.....	9
4.1 Strategic importance.....	9
4.2 Results Focus of the Evaluation (Distance down the results chain/ impact pathway)	10
4.3 Operational improvement	12
4.4 Resource Mobilization	14
5- Conclusion.....	16

Acronyms

AGM	Annual General Meeting
CGIAR	Consultative Group on International Agricultural Research
CSIRO	Commonwealth Scientific and Research Organization
DTIE	Division of Technology, Industry and Economics
EO	Evaluation Office
GEF	Global Environmental Facility
MSP	Medium Size Project
MTE	Mid-term Evaluation
TE	Terminal Evaluation
UNEP	United Nations Environment Programme

Priority Setting for Evaluation: Developing a strategic evaluation portfolio

1- Introduction

Resources are scarce and best use should be made of them to deliver against the core evaluation purposes of (i) providing evidence of results to meet accountability requirements and (ii) promoting operational improvement and learning. To this end, the Evaluation Office (EO) in the United Nations Environment Programme (UNEP) has developed a method for analyzing and prioritizing potential evaluations to improve the selection of a portfolio of activities that will lead to the greatest pay-off to the organization. The method establishes the relative priority of *'evaluation opportunities'* against criteria that directly relate to the primary purposes of the evaluation function. A range of benefits are expected to accrue to from the application of this thorough a quantitative priority setting process. These include:

- Enhancing the relevance of evaluations to the wider UNEP agenda
- Providing structured justification for allocating funds to evaluation activities
- Promoting more transparent, predictable decision-making
- Enabling realignment of evaluation priorities with changing needs and circumstances
- Improved credibility with stakeholders
- Providing a clear direction and sense of purpose to evaluation staff
- Documenting a clear framework for the development of UNEP's evaluation portfolio.

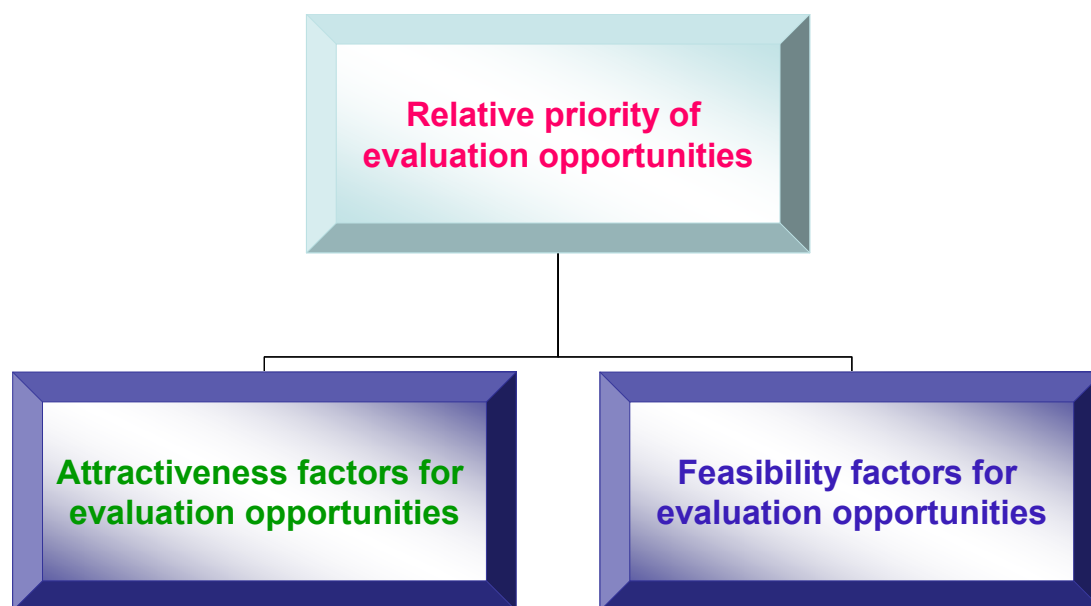
This priority setting method was developed by EO staff in late 2007 and was presented at the Annual General Meeting (AGM) of the UN Evaluation Group in April 2008. Feedback received from UN and Multilateral Development Bank evaluation professionals suggested that it is of general utility to the wider international evaluation community.

The method has been applied in the Evaluation Office's annual work planning for 2008 to help ensure that the extremely scarce evaluation resources, that the office has the freedom to allocate, are used to their best advantage. It shall also ensure, and that the relative importance of the mandatory evaluations, that form the overwhelming majority of the Evaluation Office's work plan, is better understood.

2- Priority setting Framework

Priority setting methods that use a framework similar to the one set out below have been successfully applied at a variety of levels – national (e.g. in New Zealand), institutional, programme and project (e.g. CSIRO Australia and some CGIAR centers). The method requires the definition of a set of criteria and indicators to determine the relative attractiveness and feasibility **for a set of mutually exclusive potential evaluations which we term 'evaluation opportunities'**.

Figure 1. The Attractiveness and feasibility framework for priority setting



The Attractiveness criteria address important characteristics of evaluation opportunities that relate directly to the evaluation purposes of providing accountability and promoting operational improvements. The main factors that affect the ‘*Attractiveness*’ of an evaluation opportunity were defined as:

- The consistency of the evaluative topic with the organization’s strategic direction
- The primary results focus of the evaluation (the ‘distance’ along the impact pathway – from activities to impacts .e.g. a mid-term evaluation would focus on a short ‘distance’ along the impact pathway, whereas an impact assessment would focus towards the end of the pathway)
- The magnitude and distribution of environmental benefits from the intervention likely to be assessed at the time of the evaluation
- The potential importance of the evaluation opportunity for resource mobilization
- The number of uptake events required, the directness of the impact pathway and length of lag times for maximum impact to be achieved from an intervention
- The ease of attribution to the project/ programme of the effects evaluated and the existence of adequate baselines/ counterfactuals and/ or performance monitoring information
- The timing of the evaluation opportunity relative to project or programme operations
- The likelihood that operationally relevant findings or lessons will be applicable to other UNEP projects/ programmes
- The potential of a project/ programme to implement adaptive management, or mitigate known risks to project performance

The ‘*Feasibility*’ criteria address issues relevant to the implementation of the evaluation itself, and were defined as:

- The level of technical / evaluative capacity required and the associated availability of consultants
- The level of oversight capacity required for management and quality control of the evaluation opportunity
- The level of effort required in the design of the evaluation approach
- The simplicity of the evaluation methods needed

For each of these 'attractiveness' and 'feasibility' factors a set of verifiers were determined that represent the range of values from 'low' to 'high'. These are shown in Table 1 below.

The overall focus in determining the relative priorities for evaluation is the 'return to UNEP' from investment in evaluation. The overall return is higher, when both 'attractiveness' and 'feasibility' are high. As 'attractiveness' and 'feasibility' both decline, so too does the 'return to UNEP', and so the selectivity in choosing among evaluation opportunities increases.

Figure 2. Selectivity of evaluation investments using attractiveness and feasibility factors

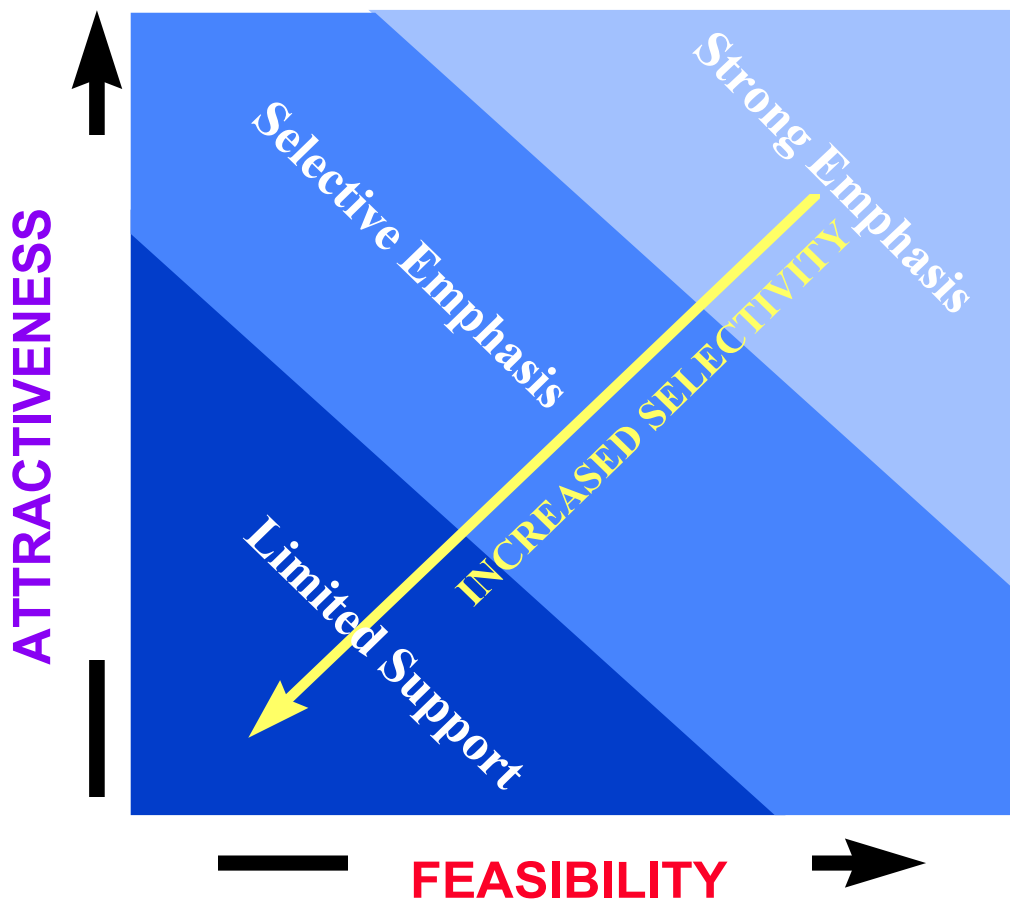


Table 1. Scoring matrix of key factors affecting the ‘attractiveness’ and feasibility’ of evaluations, for use in priority-setting.

It is not intended that all the verifiers in each ‘cell’ will always apply simultaneously. Sometimes verifiers relating to a single evaluation opportunity will fall in differing cells within a row requiring judgments to be made on the score awarded for that indicator.

KEY FACTORS	1	4	7	10	Weight
Strategic importance of evaluative study 20%					
<i>Consistency of the evaluative topic with the organization’s strategic direction.</i>	Poor fit with current strategy; Low future interest to UNEP future resource allocation to topic area is ‘low’.	Moderate fit with current strategy; Important mainly for some UNEP stakeholders; Future resource allocation to topic area is ‘moderate’	Good fit with current strategy; Study important for UNEP’s future, current and future resource allocation to topic area are ‘high’ and ‘high’ respectively.	Strong fit with core elements of strategy; Central to UNEP’s mission and future, addresses a GC directive current and future resource allocation to topic area are ‘low’ and ‘high’ respectively.	20%
Evaluation Attractiveness Criteria 50%					
Results focus 5%					
<i>Primary results focus of the evaluation (distance along impact pathway).</i>	Evaluation focuses on delivery of outputs, efficiency of project or programme operations	Achievement of outcomes overall programme/ project performance (e.g. TE)	Evaluation focuses on overall programme/ project performance and assessment of outcomes, influences and immediate impacts	Assessment of outcomes and influences immediate and long-term impacts	15%
Magnitude and distribution of effects from the intervention 15%					
<i>Magnitude and distribution environmental benefits from intervention likely to be assessed at time of evaluation.</i>	Intervention has largely generated outputs and some outcomes (that are expected to lead to generation of environmental benefits)	Localized significance of repair or avoidance of degradation; Some outcomes likely to be assessed	Low regional/ global significance of repair or avoidance of degradation	High regional/ global significance of repair or avoidance of degradation Multi-regional/ global effect;	10%
Potential for resource mobilization 5%					
<i>Potential importance of evaluation opportunity for resource mobilization.</i>	End of project/ programme evaluation where the approach/ topic is unlikely to secure additional funding. Mid-term evaluations (MTE)	Evaluation where the approach/ topic is of moderate priority to funding agencies and moderately likely help secure additional funding.	Evaluation where the approach/ topic is of moderate priority to funding agencies and likely to help secure additional funding.	End of project/ programme phase where continued funding is likely but depends on evaluation outcome; Pilot project evaluations; Evaluation where the approach/ topic is of high priority to funding agencies and likely to help secure additional funding; Outcome, influence and impact studies.	5%
Ease of Attribution of observed effects to the intervention 10%					

KEY FACTORS	1	4	7	10	Weight
<i>Number of uptake events required, the directness of impact pathway and lag times for maximum impact to be achieved.</i>	Multiple uptake events required; Long indirect pathways; Causality poorly understood, very long lag times (>10 yrs).	Several uptake events required; Direct and indirect pathways; Moderate understanding of causality, long lag times >7yrs <10).	Few uptake events required; Indirect pathway; Good understanding of causality, moderate lag times (>4yrs < 7yrs).	Few uptake events for maximum impact; Simple direct impact pathway; Very well understood causality, short lag times (< 4yrs).	
<i>Ease of attribution of effects to project/ programme outcomes.</i>	Difficult to attribute (e.g. changes to policy processes); Measurement metrics are only indirectly linked to the intended effects; Many actors/ similar initiatives or sources of information; Multiple steps in the results chain still exist between outcomes achieved and desired effects; “contribution without attribution”.	Some evidence of causality between outcomes and effects; Multiple steps in the results chain still exist between key outcomes achieved and desired effects.	Some evidence of causality between outcomes and effects. Some specific pathways show stronger causality; Few steps in the results chain still exist between key outcomes achieved and desired effects.	Easy to attribute; Few or no steps in the results chain still exist between outcomes achieved and desired effects.	7%
<i>Existence of adequate baselines/ counterfactuals and/ or performance monitoring information.</i>	No baselines; Baselines cannot be inferred; No results-oriented performance monitoring information available.	No baselines; Baselines can be inferred; Limited results-oriented performance monitoring information available.	Some baseline data exists; Adequate results-oriented performance monitoring information available.	Comprehensive baseline data exists; Comprehensive results-oriented performance monitoring information available.	3%
Potential for operational improvement and learning 15%					
<i>Timing of evaluation opportunity relative to project or programme operations.</i>	Several years after project completion.	Shortly before the completion of project/ programme operations or ‘end of phase’ where continued funding is uncertain.	Formative evaluation of project design (quality at entry).	During project operations e.g. MTE or ‘End of Phase’ evaluation (where continued funding is likely)	5%
<i>Likelihood that operationally relevant findings or lessons will be applicable to other UNEP projects/ programmes.</i>	Project executed by an external organization; Project not representative of a wider portfolio of ongoing or planned projects within a region, theme or sector; ‘One-off’ project; Project partners/ executors unique to the Initiative/ intervention.	UNEP has limited involvement in project execution; Project similar to a limited number of ongoing or planned projects within a region, theme or sector; Project partners/ executors common to a few UNEP Initiatives.	UNEP has substantive involvement in joint project execution; Project partly representative of a wider portfolio of ongoing or planned projects within a region, theme or sector; Project partners/ executors common to some other UNEP Initiatives.	Project solely executed by UNEP; Project fully representative of a wider portfolio of ongoing or planned projects within a region, theme or sector; A pilot project for larger initiative; Project partners/ executors common to many other UNEP Initiatives.	5%

KEY FACTORS	1	4	7	10	Weight
<i>Potential of project/ programme to implement adaptive management, or mitigate known risks to project performance.</i>	“Highly Satisfactory” in the Project Implementation Review (PIR); No problems identified in self evaluation reports; Influence study/ Terminal evaluation (TE).	“Moderately Satisfactory” in the PIR; Few problems identified in self evaluation reports; MTE/ End of Phase.	At risk or “Moderately Unsatisfactory”/ “Unsatisfactory” in the PIR; Several significant problems identified in self evaluation reports; MTE/ End of Phase.	At risk and “Highly Unsatisfactory” in the PIR; Major problems identified in self evaluation reports project at risk of failure; Early MTE/ End of Phase.	5%
Evaluation Feasibility for EO-Criteria 20%					
Evaluation Capacity required for evaluation opportunity 10%					
<i>Level of technical/ evaluative capacity required and availability for evaluation opportunity.</i>	Technical and evaluation expertise required in depth, and very difficult to procure for the evaluation theme/ focus; Team of experts required.	Technical and evaluation expertise required; Difficult to procure for the evaluation theme/ focus.	Limited technical expertise in theme/ focus of the evaluation opportunity is needed; In-depth evaluation experience required fairly easy to procure for the evaluation theme/ focus.	Adequate to use non Technical evaluation expertise which is easy to procure for the evaluation theme/ focus.	5%
<i>Level of oversight capacity required for evaluation opportunity.</i>	Jointly implemented projects with other agencies; Multi-country/ region portfolio evaluations.	Multi-country/ region intervention.	Single region interventions; UNEP sole implementation agent (externally executed); Multiple UNEP divisions/ subprogrammes involved.	UNEP sole implementation and execution agent; Single country involved; Single UNEP division/ subprogramme involved.	10%
Simplicity of evaluation design and methods needed 10%					
<i>Level EO Staff effort required in design and implementation of evaluation approach.</i>	Considerable design work needed in preparation of TORs/ evaluation approach; EO staff leads the evaluation.	Extensive adaptation to standard TORs; EO staff participates substantively in the evaluation.	Limited adaptation to standard TORs; EO staff participates to a limited extent in the evaluation.	Standard TOR template can be applied.	5%
<i>Reliability and precision of evaluation approach.</i>	In depth study; Many sources of triangulation; Extensive use of statistical methods/ Delphi techniques; Clearly articulated assumptions and counterfactuals; Rigorous third party peer review.	In depth study; Many sources of triangulation for findings presented; Detailed field studies; User surveys; Rigorous third party peer review.	In-depth study where existing information is verified by interviews and some reference documents.	Desk study review of documents with some interviews/ simple surveys.	5%

3- Priority setting method using ‘evaluation opportunities’

The matrix shown in Table 1 and the associated priority setting processes were developed by Evaluation Office professionals through an interactive and iterative process of reasoning, debate, and, initially, the use of a ‘dummy’ set of ‘evaluation opportunities’ (recently completed evaluations from the 2007 Evaluation Office workplan) to test the approach.

Assessments for each ‘dummy’ evaluation opportunity, using the scoring matrix, were made independently by UNEP evaluation professionals. The professionals then, together, discussed and reviewed assessments that each made independently. Where there were large variances in the scores awarded, the reasons were examined.

Most commonly these variances were due to differences in understanding of either i) the ‘attractiveness’ or ‘feasibility’ factors used in the scoring matrix or ii) the specific details and context of the ‘evaluation opportunity’ itself.

Clarifications prompted; re-scoring of the ‘evaluation opportunity’ or refinement of the scoring matrix, followed by re-scoring. Once scores for each ‘evaluation opportunity’ were finalized, averages for each factor were calculated and used in the analysis of relative evaluation priorities.

The scoring and analysis process was then repeated using ‘evaluation opportunities’ selected from the Evaluation Office’s 2008 workplan. The process can be summarized as follows:

- Each ‘evaluation opportunity’ was independently scored by Evaluation Office professionals for each of the various ‘attractiveness’ and ‘feasibility’ factors using the scoring matrix;
- A modified Delphi approach was used – and the reasons for differences in scores between professionals and of each criterion for each ‘evaluation opportunity’ were discussed;
- The evaluation opportunities were then re-scored, and;
- Ultimately, the averages for each evaluation opportunity provide the basis for the analyses that highlight the relative priorities for ‘evaluation opportunities’.

The 2008 Evaluation Office workplan lists approximately 60 planned evaluations. Rather than assessing the relative priorities for all evaluations, a sampling approach was used. The aim was to apply the priority setting framework to a sample of planned evaluations that was broadly representative of the entire work plan. All six planned management studies/thematic evaluations were included in the sample.

The remaining project/portfolio evaluations were then categorized by thematic area according to UNEP’s strategic framework. Project/portfolio evaluations were then sampled from within each thematic area to give, as far as possible, a balance between mid-term and terminal evaluations and single country, multi-country/regional project approaches.

The selections also aimed to achieve a balance across geographic areas. The average scores for each evaluation opportunity provided the basis for the ‘attractiveness-feasibility’ analyses.

Table 2. Planned project /portfolio evaluations for 2008, and numbers sampled for priority setting classified by UNEP thematic areas

UNEP Thematic area	N° of evaluations in the 2008 work plan	N° evaluations sampled
Climate Change	9	3
Disaster and Conflict	1	1
Ecosystem Management	20	5
Environmental Governance	6	1
Harmful Substances	6	2
Resource Efficiency	12	3
Total	54	15

The 'evaluation opportunities' considered in the priority setting exercise were:

1. UNEP influence study on the Global Environment Policy Agenda (Policy Influence)
2. UNEP study of its Civil Society Programme (Civ. Soc)
3. Terminal evaluation of the project "New Arrangements for the Agreement on Small Cetaceans of the Baltic and North Seas Secretariat" (ASCOBANS)
4. Terminal evaluation of the Belgium-UNEP partnership (Belg. Partn)
5. Evaluation of the UNEP Division of Technology, Industry and Economics (DTIE)
6. Evaluation of the Quality of Project Supervision in the Division of Global Environment Facility Coordination (DGEF Supervision)
7. UNEP terminal evaluation of the project "Achieving the Johannesburg Plan of Target of IWRM and Efficiency Plans by 2005" (IWRM)
8. Global Environment Facility (GEF) terminal evaluation of the medium size project (MSP) of the Ozone Portfolio (Ozone Port.)
9. GEF terminal evaluation of the full size project (FSP) "Solar and Wind Energy Resource Assessment" (SWERA)
10. UNEP terminal evaluation of the project "Using Carbon Finance to Promote Sustainable Energy Services in Africa" (C-Finance)
11. GEF terminal evaluation of the MSP "Sustainable land use planning for disaster preparedness in the lower Limpopo Basin" (Limpopo)
12. GEF terminal evaluation of the FSP "Reversing Environmental Degradation Trends in the South China Sea and Gulf of Thailand" (S. China Sea)
13. GEF mid-term evaluation of the FSP "Guinea Current LME- Combating Living Resource Depletion and Coastal Area Degradation" (GCLME)
14. GEF mid-term evaluation of the project "African/Eurasian Flyways - Enhancing Conservation of the Critical Network of Sites of Wetlands Required by Migratory Waterbirds" (Flyways)
15. GEF terminal evaluation of the MSP "Cedar Forests in the Mediterranean Region- Development of an Action Plan for Integrated Management of Forests and Assessment of Insect Infestation" (Cedars)
16. GEF mid-term evaluation of the MSP "Building the Partnership to deliver the Global 2010 Indicators" (2010)
17. GEF terminal evaluation of the FSP "Development of National Implementation Plans for the Management of Persistent Organic Pollutants" (NIPs)
18. GEF terminal evaluation of the MSP "Assessment of Existing Capacity and Capacity Building Needs to Analyse POPs in Developing Countries" (POPs Labs)
19. UNEP evaluation of the implementation of a Regional Programme of Sustainable Production and Consumption in Latin America and the Caribbean (LAC SCP)
20. UNEP China Rural Energy Enterprise Development (CREED)
21. GEF mid-term evaluation of the project "Conservation and Sustainable Use of Biodiversity through Sound Tourism Development in Biosphere Reserves in Central and Eastern Europe" (Tourism)

4 - Analysis and discussion of the evaluation portfolio for 2008

The combined scores for each 'evaluation opportunity' across all factors, provides one measure for assessing the relative priorities associated with Evaluation Offices' 2008 work plan in terms of the likely benefit accruing to UNEP from their successful completion.

Figure 3. Combined 'feasibility' and 'attractiveness' scores

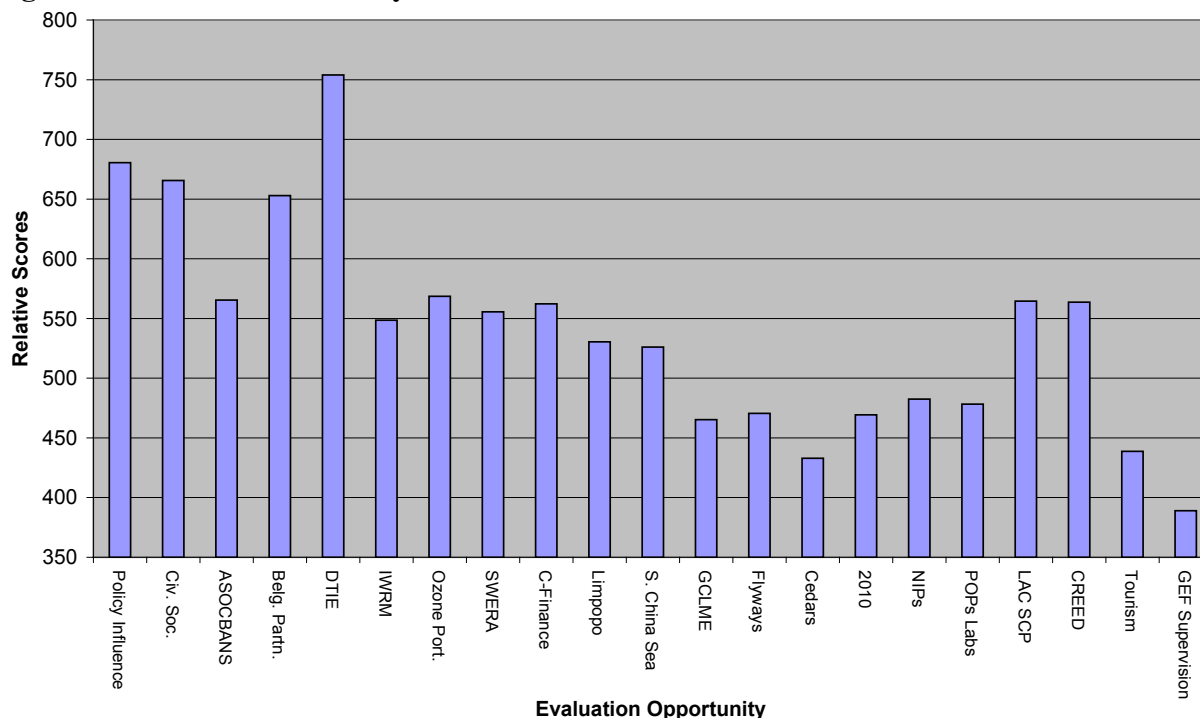
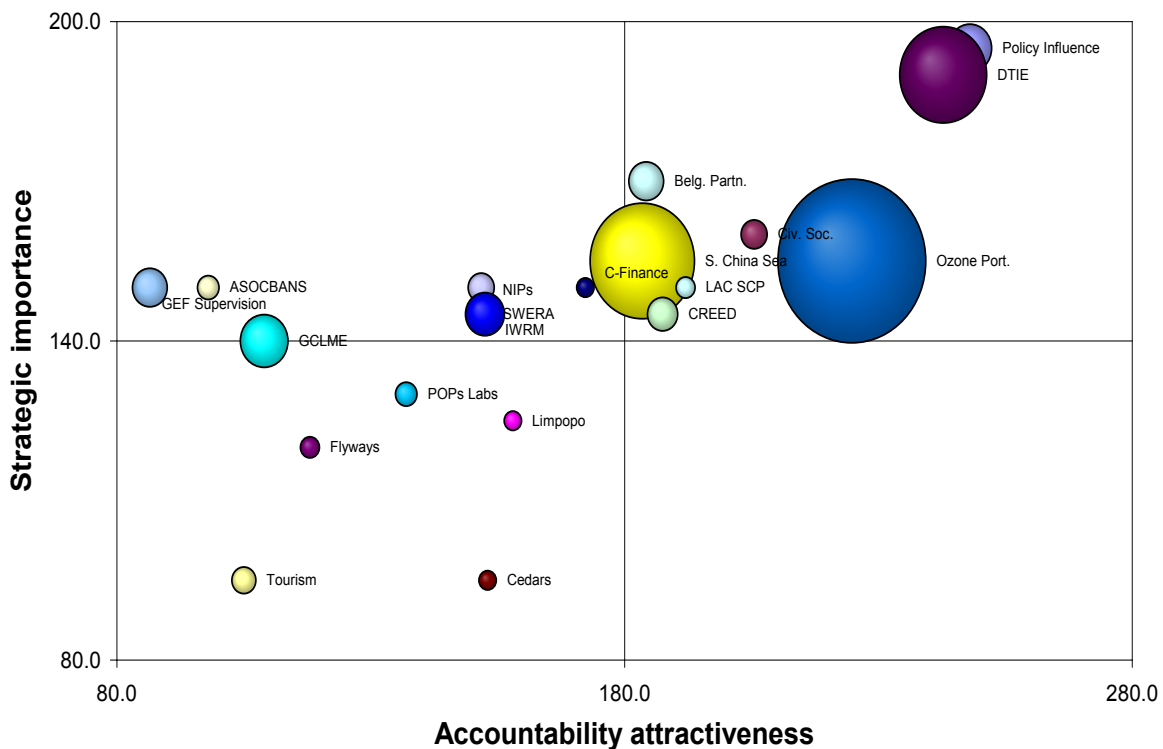


Figure 3 shows that the highest ranking overall evaluation priorities in the 2008 work plan are the thematic and management studies. These aggregated scores yield overall priorities but do not allow an interpretation of those priorities with respect to the different purposes evaluations can serve. To gain an understanding of these issues, different attributes of the 'evaluation opportunities' are examined in more detail.

4.1 Strategic importance

Some project or programmatic activities in the UNEP work programme have greater relevance to the future strategy of the organization as compared to others. Evaluations of projects/programmes of high strategic importance will be of greater potential utility to the organization to improve future operational performance as compared to projects or programmes that are unlikely to feature in the future work of the organization. Providing evidence in terms of results (outcomes, influences and impacts) achieved in the areas of future strategic importance to UNEP enhances accountability by providing credible evidence to support the organization's 'track record'. Figure 4 shows the relative priorities for evaluation opportunities with respect to their 'strategic importance' — how well evaluation opportunities are aligned with UNEP's strategic focus and their 'accountability attractiveness'.

Figure 4. Strategic importance versus accountability attractiveness of ‘evaluation opportunities’



Note: The diameter of the plotted points is proportional to the estimated cost for each evaluation opportunity.

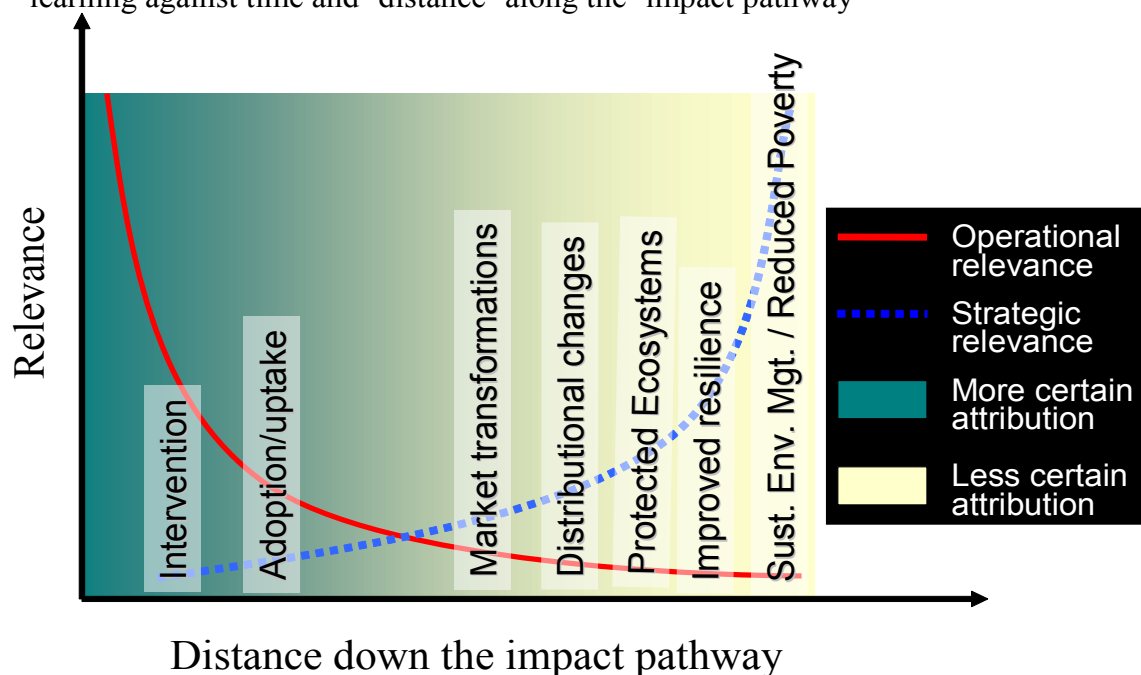
The graph highlights, in the top right hand quadrant, the evaluations that are most likely to deliver findings that assess outcomes, influences and impacts of UNEP work in areas of strategic significance. These represent the priorities with respect to the accountability purpose of UNEP’s evaluation function. The top priorities are the proposed evaluation of the ‘Policy Influence’ project of DTIE’s *Chemicals* subprogramme element and the overall subprogramme evaluation of DTIE. In general, few project-level terminal evaluations are ranked as high priority. Exceptions include the terminal evaluations of the GEF “Ozone Portfolio” and the “South China Seas” Project.

4.2 Results Focus of the Evaluation (Distance down the *results chain/ impact pathway*)

There is a tradeoff between the level at which an evaluation is focused in terms of results (outputs, outcomes, influences or impacts), and its utility for ‘accountability’ or ‘operational efficiency/learning’ purposes.

For example, an evaluation that aims to assess long-term impacts will usually only take place long after a project or programme has ended. Whilst such an evaluation may have a high utility for accountability purposes, it will be of little relevance for the purposes of improving the ‘operational efficiency’ of that particular project or programme (because it has ended). However such evaluations may generate some lessons for future programme/projects of a related nature.

Figure 5. The tradeoff between accountability and operational improvement / learning against time and ‘distance’ along the ‘impact pathway’



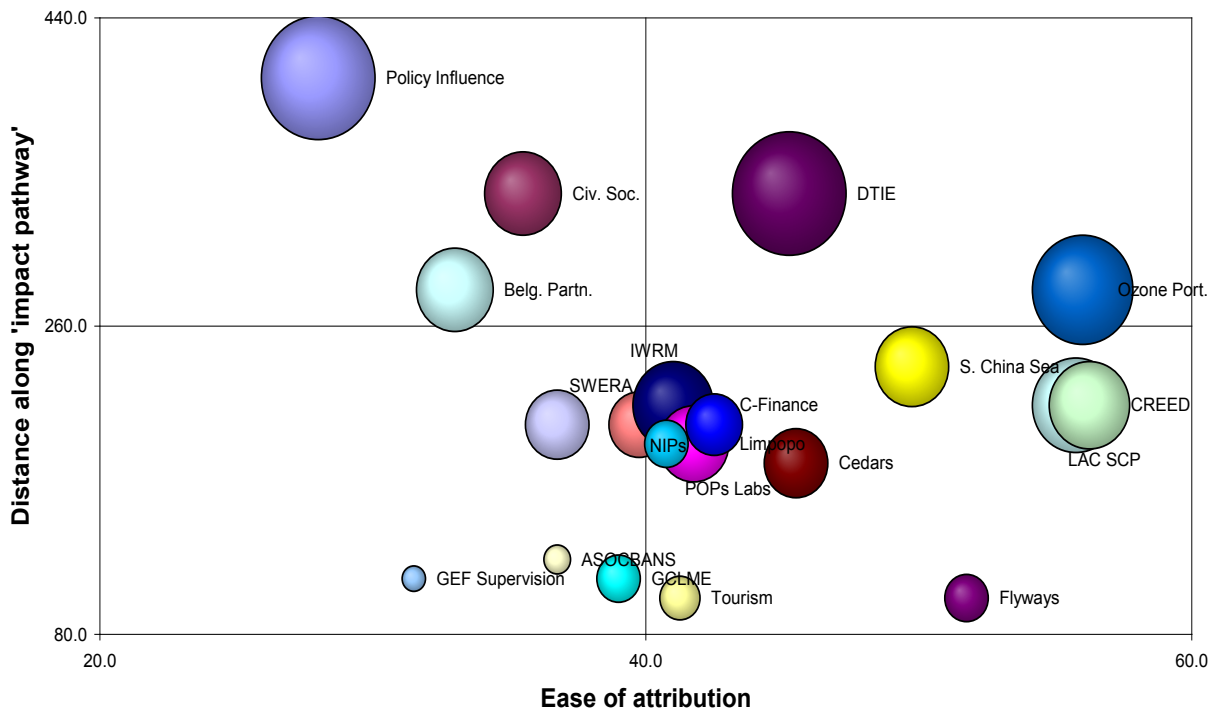
Conversely, an evaluation that takes place during project/ programme implementation will be likely to have high utility for the purposes of ‘operational improvement’ (adaptive management) of that project/programme but is likely to be too soon in the life of the intervention to assess the intended impacts. There is, therefore, a tradeoff between the time/ speed of delivery of evaluation findings and their operational relevance to the entity being evaluated.

For example, a mid-term evaluation of a project or programme will typically yield larger benefits in terms of suggestions to improve project implementation (course-correction/ adaptive management) than would an evaluation of the same project at its completion.

The ‘results focus’ of an evaluation also includes the nested tradeoffs of ‘attributive ease’ versus ‘lack of impact knowledge’ (see Figure 5 above). Assessing outcomes shortly after completion of an initiative might be easier in terms of establishing attribution of any effects to the intervention, but real impacts will often be yet to accrue. Conversely, an impact assessment some time after project/ programme completion might be better able to measure changes in impact-related metrics but attribution to the original intervention may be more challenging due to the passage of time and the lack of adequate baselines/ counterfactuals to mitigate the difficulty of separating out the effects of other events and actors.

It is also instructive to examine the ‘attractiveness for accountability’ (Figure 6) by plotting the scores for evaluation opportunities in terms of their “distance down the impact pathway”, against the corresponding “ease of attribution”, for the causal effects evaluated. The score for *the magnitude and scale of effect likely to be assessed at the time of the evaluation* is represented by the diameter of the points plotted. This graph offers the prospect of identifying any ‘low hanging fruits’ – evaluations that are likely to demonstrate environmental results that are more readily attributed to UNEP interventions. There are no such evaluations within the sample of projects selected from within the 2008 work plan.

Figure 6. Attractiveness of ‘evaluation opportunities’ for accountability purposes.



Note: The diameter of the plotted points is proportional to the scores for the magnitude and distribution of benefits likely to be assessed at the time of the evaluation

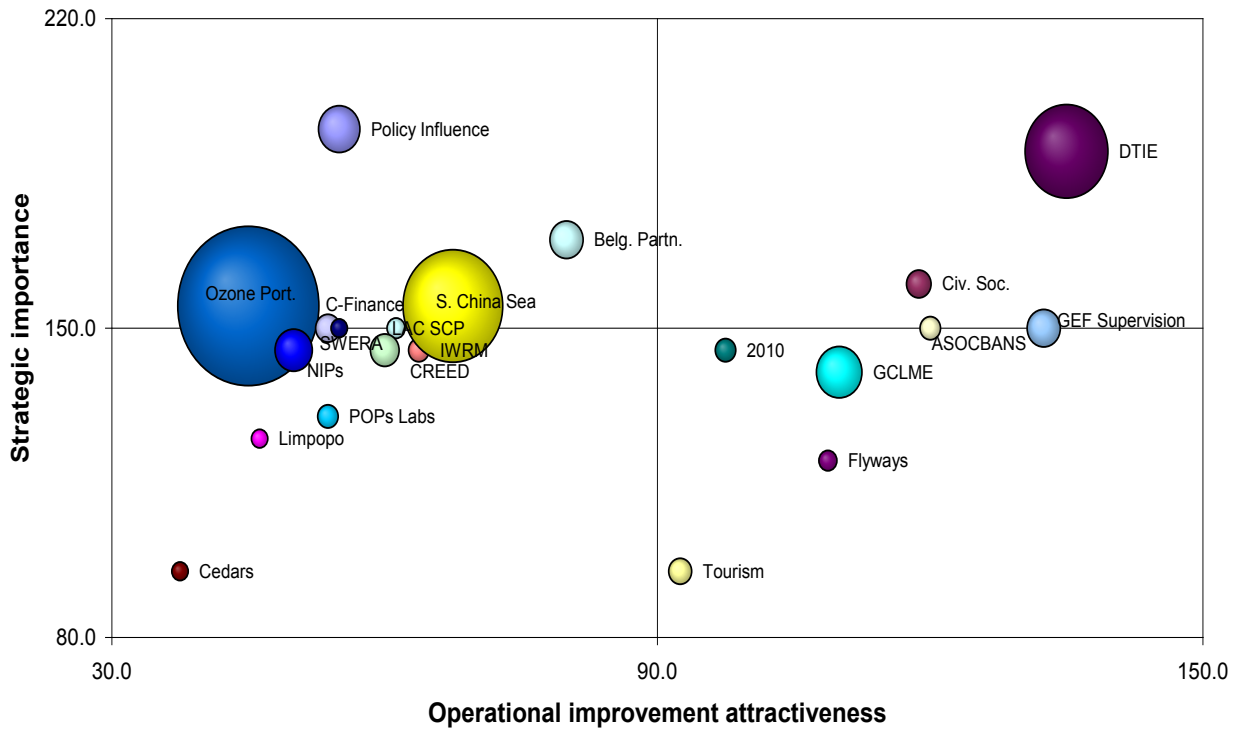
The “Policy Influence” evaluation, the DTIE subprogramme evaluation, and the GEF “Ozone portfolio” evaluation all scored similarly for the magnitude of benefits likely to be assessed. However, whilst the “Policy Influence” evaluation of DTIE’s *Chemicals* subprogramme element is regarded as the most attractive in terms of its ‘results focus’, it also presents the greatest challenges in terms of attribution.

In contrast, the “Ozone portfolio” evaluation has the most favorable attribution score because of the linkage between reduction and elimination of Ozone Depleting Substances (a performance measure of projects in this GEF portfolio) and environmental benefits of avoided damage to the Ozone layer. The GEF “Ozone portfolio” evaluation is the highest priority evaluation when assessed against these criteria.

4.3 Operational improvement

If we examine evaluation priorities in terms of the likely contributions to operational improvement in UNEP, a different pattern emerges. Mid-term evaluations or evaluations of ongoing projects and programmes tend to score higher because evaluation findings are likely to feed back directly into management and implementation. The DTIE subprogramme evaluation emerges as the highest priority evaluation against the criteria of ‘*attractiveness for operational improvement*’ combined with ‘*strategic importance*’.

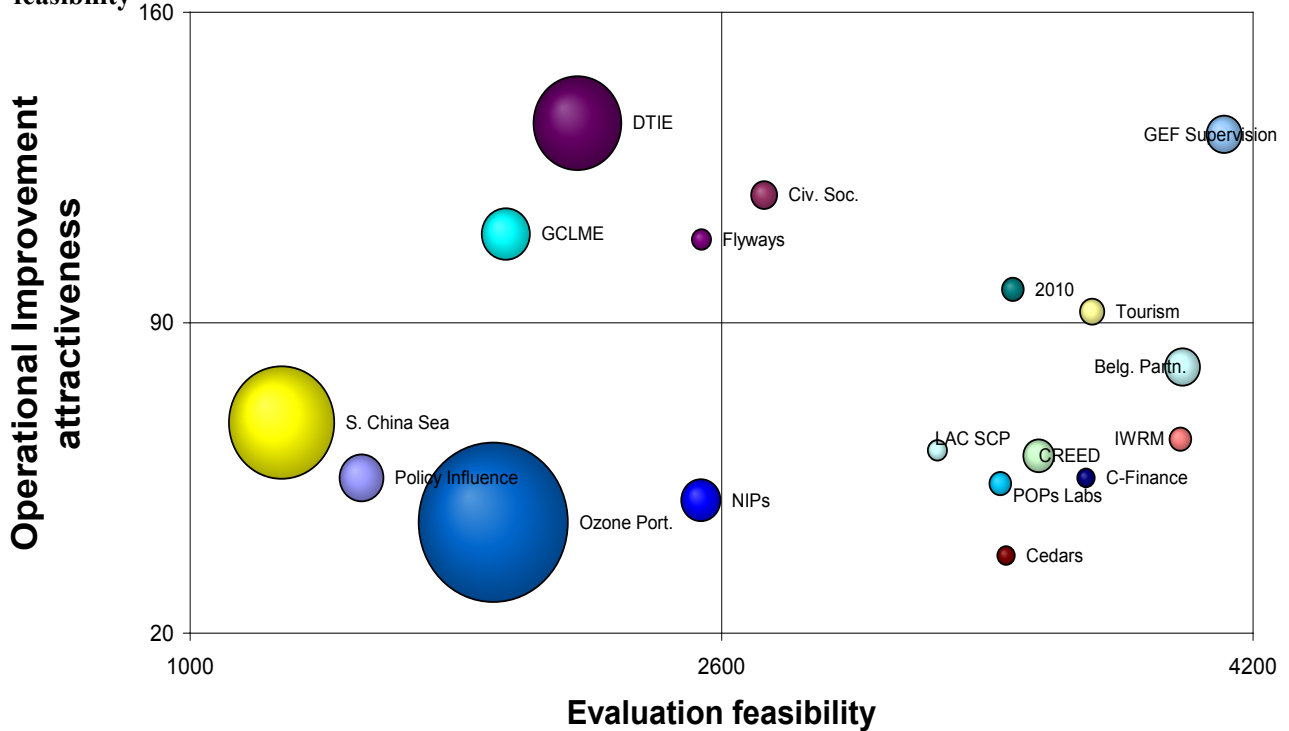
Figure 7. Strategic Importance & ‘attractiveness for operational improvement’



However, if the Evaluation Office were to place the emphasis on securing ‘quick wins’ in terms of learning and operational improvement, then we would look at the priorities in terms of ‘*Operational improvement return*’ and plot the combined scores for ‘*attractiveness for operational improvement*’ against the scores for the factors that relate to ‘*evaluation feasibility*’. The plot shows that evaluation of ongoing projects/programmes is more ‘*attractive*’ than those that have ended or are nearing completion. The evaluation of GEF project supervision being the most ‘feasible’ and the most likely to yield significant ‘operational improvement’ benefits.

Figure 8 also shows that those projects that are executed by UNEP score more highly than those projects that are executed by a third party (as is the case for several of the GEF evaluations), because in this latter case, operational improvements within such a project are not likely to be directly captured by UNEP – but by the executing agency. However, the ‘*attractiveness for operational improvement*’ may still be significant if the project is similar in nature and focus to others that feature in UNEP’s work programme.

Figure 8. ‘Attractiveness for operational improvement’ versus ‘evaluation feasibility’



Note: The diameter of the plotted points is proportional to the estimated cost for each evaluation opportunity.

4.4 Resource Mobilization

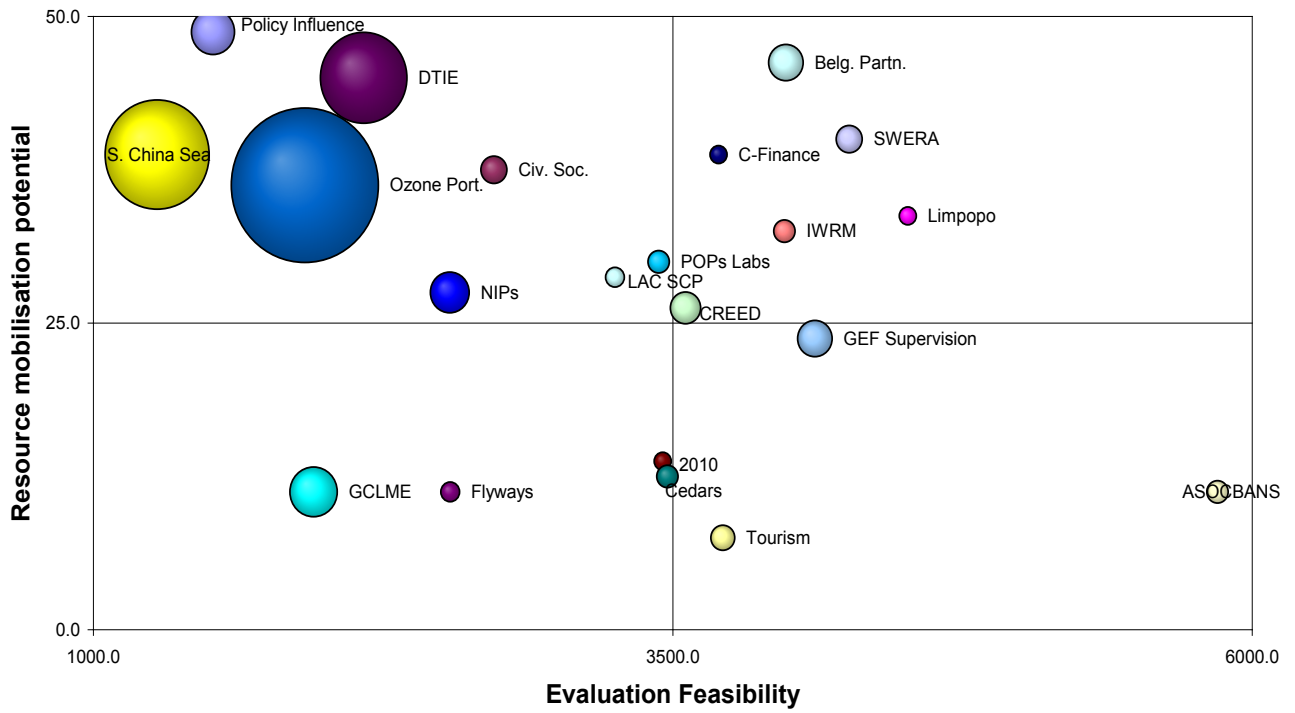
Evaluation findings can play an important role in helping to secure financial resources for UNEP. If we examine the 2008 Evaluation Office workplan in terms of the potential of the ‘evaluation opportunities’ to contribute to resource mobilization and with respect to the ease with which the evaluation can be conducted – the combined ‘feasibility’ factors – several patterns can be observed.

Project evaluations that focus on course correction and operational improvement often have a more limited potential for use in resource mobilization efforts – as the resources still remain in the project and the major achievements have often yet to be realized.

By contrast, evaluations that are likely to provide evidence of effective UNEP performance in achieving important outcomes and influences that have (or are likely) to generate tangible environmental benefits can contribute greatly to resource mobilization efforts. This is especially so when the focus of the evaluation coincides with the funding priorities of important donors.

In the case of GEF projects, completion of a high quality terminal evaluation that suggests that continued investment in the project, or its follow-up phases, is merited is essential for securing continued GEF financial support.

Figure 9. Attractiveness for resource mobilization versus the feasibility of evaluation opportunities



Note: The diameter of the plotted points is proportional to the estimated cost for each evaluation opportunity.

5- Conclusion

Whilst applying a rigorous priority setting approach helps considerably with the development of an evaluation portfolio that can maximize the returns to the organization (with respect to the primary purposes of the function) **tactical investments in some evaluation activities** will still be necessary. Such investments may waive the normal requirements for ‘*evaluation opportunities*’ to be of high strategic importance, or to make significant contributions to the organization’s accountability or operational efficiency.

For example, they may be evaluations undertaken to meet specific external requirements, to fulfill a specific request from management, or to enhance the professional credibility of UNEP’s evaluation function. Evaluation managers should use the priority setting method as a tool to analyze the ‘portfolio’ of evaluation investments, and not to unilaterally ‘filter out’ evaluation opportunities with lower potential.

For the key evaluative purpose of demonstrating accountability, and in the context of ‘results-based management’, there are obvious advantages to preferentially selecting UNEP ‘success stories’ when investing scarce resources in an evaluation portfolio. Focusing accountability-oriented evaluations entirely on cases where causal linkages to outcomes and impacts are readily established (high attributive ease) has pragmatic appeal but would tend to limit evaluations at UNEP either to studies that focus ‘early’ in the impact pathway (on outputs and immediate outcomes) or restrict accountability-oriented evaluations to a minority of cases where attribution issues are straightforward.

Since much of UNEP’s work seeks impacts via indirect impact pathways that are often linked to policy change processes, this would yield an unbalanced evaluation portfolio and enshrine a tendency to ‘shy away’ from the greatest challenges in results-based evaluation and impact assessment. It would also result in much of UNEP’s investment in evaluation lacking sufficient focus on initiatives aimed at generating environmental benefits. Whilst the evaluation function will strive to link UNEP activities to quantification of environmental benefits, this is only likely to be possible (due to the considerable attribution difficulties) for a small proportion of activities in UNEP’s work programme.

It should, however, be noted that an earlier study on “evaluation demand” in UNEP showed that the demand for accountability-oriented evaluations is high and currently not well-matched, either with the capacity or the level of resources required by UNEP’s evaluation function to deliver them.

‘In-house’ evaluation functions often face criticisms that their evaluations are heavily biased by self-serving motives. This acts to increase the level of skepticism with which evaluation findings are regarded by external audiences unless; the evaluation function has a high level of independence in the organization, a track record of disclosing negative as well as positive evaluation findings is established, and the organization subjects its evaluations to rigorous peer review processes.

In contrast, operational improvements and institutional learning (feedback) that result from evaluation findings are (by definition) also ‘self serving’. This is frequently perceived favorably by donors and other external stakeholders and can help to foster perceptions, and the reality, of UNEP being a ‘learning organization’. Operational improvement and learning are important reasons for investing in evaluation and should therefore be a consideration in the selection of a balanced portfolio of evaluation studies.

Whilst it is very informative to understand the relative priorities of different evaluations on the Evaluation Office work programme, in practice, the overwhelming majority of evaluations the Evaluation Office currently undertakes are mandatory and their costs are not fungible – the evaluation budgets reside within the project or subprogramme budgets and not with the Evaluation Office. For such **priority setting efforts to be meaningful there has to be a strong link to resource allocation decisions.**

The findings from this study indicate that the greatest pay-offs to UNEP will come from evaluations that focus on the assessment of outcomes, influences and impact of ongoing programmatic activities that are closely aligned with UNEP’s strategic priorities.

Evaluation Office

United Nations Environment Programme
P.O. Box 30552
Nairobi, Kenya
Tel: +(254-20) 7623387
Email: Segbedzi Norgbey@unep.org
Email: eou@unep.org
URL: <http://www.unep.org/eou>

www.unep.org

United Nations Environment Programme
P.O. Box 30552 Nairobi, Kenya
Tel: ++254-(0)20-762 1234
Fax: ++254-(0)20-762 3927
Email: unep@unep.org
web: www.unep.org

