Synthesis Report on the UNICEF Evaluation Policy and Function

Final Synthesis Report

By: Universalia Management Group

February 2013
Executive Summary

This synthesis report is the final report among three deliverables in 2012 which contribute to a wider review of the UNICEF evaluation function and the 2008 Evaluation Policy. This report encompasses many of the elements of the two previous reports; however, each is a stand-alone document, with relevant recommendations.

The UNICEF Evaluation Policy aims to establish a common institutional basis for the UNICEF evaluation function and seeks to increase transparency, coherence, and efficiency in generating and using evaluative knowledge for organisational learning, managing for better results and supporting accountability. As highlighted in the Annual Report on the evaluation function and major evaluations in UNICEF (E/ICEF/2012/13), “The organisational context and the wider international settings have seen major changes in recent years, with more changes expected in 2012. UNICEF therefore has undertaken a thorough review of its evaluation function and the Evaluation Policy, with a view to updating the policy for submission to the Executive Board in mid-2013.”

In 2006, a panel of six international experts conducted a peer review of the evaluation function in UNICEF. The findings were presented to senior management of UNICEF and an informal meeting of the Executive Board. The peer review panel concluded that UNICEF could strengthen the evaluation function by: adopting a clear and comprehensive Evaluation Policy document consistent with United Nations Evaluation Group (UNEG) Norms and Standards, using a more predictable budget for evaluation, undertaking additional interventions to strengthen and support field offices, and improving the use of results-based management throughout the organisation. The importance of working with national governments was also a focus of attention at both the Executive Board level and within the aid effectiveness agenda of the international development community.

In 2008, UNICEF’s own Evaluation Policy – developed from the UNEG Norms and Standards as recommended by the Peer Review Panel – was approved by the UNICEF Executive Board. Following approval of the policy, an Executive Directive was sent to all UNICEF offices in March 2009. The policy aimed to establish a common institutional basis for the UNICEF evaluation function in order to increase transparency, coherence, and efficiency in generating and using evaluative knowledge for organisational learning, managing for results, and supporting accountability.

This synthesis report is one of several inputs into an internal UNICEF review which may contribute to an updating or renewal of the Evaluation Policy and function. The main purpose of the synthesis review is to look at experience and performance to date using inputs from various sources and make forward looking recommendations on measures and adaptations required. The synthesis report makes use of primary data produced through a survey of Country Office (CO) Representatives, a set of case studies and a number of key informant interviews; and secondary data in the form of reports and studies produced for the Evaluation Office (EO) previously. In the collection of both primary and secondary data, consultations were limited to UNICEF staff. The following reports acted as key sources and references for this synthesis review:

- 2012 Case studies and key informant interviews (2012 Case Study Report)
- Additional documents produced for the Evaluation Office. These include the final report of the Global Evaluation Report Oversight System (GEROS), the Evaluation Component of the Qualitative Review of Country Office Annual Reports from 2011, the Annual Report on the evaluation function and major evaluations in UNICEF from the past few years.
Using the UNICEF Evaluation Policy as a framework, this report reviews the elements of the experience and performance of the Evaluation function and Policy within UNICEF. Findings of this report will aim to contribute to informing the overall evaluation policy dialogue within the agency, with commensurate suggestions for strengthening and forward-looking guidance. This report’s findings are organised around the 15 key guiding principles of the evaluation function, as highlighted in the Evaluation Policy. The Medium-Term Strategic Plan (MTSP) (2006-2013) also aimed to strengthen six main areas of the evaluation function, and these are presented. Finally, the report uses the UNEG Norms and Standards for Evaluation and its peer review assessment matrix to look at other potential areas for improvement.

Findings

The Evaluation Policy guiding principles as well as the UNICEF MTSP highlight the focus on country-driven programming and developing national capacity in evaluation, both of which are fundamental goals of the Paris Declaration on Aid Effectiveness and its successor declarations. The 2012 Quadrennial Comprehensive Policy Review (QCPR) process has also taken on country-driven programming as one of the priorities for the United Nations as a whole. There is a strong consensus in the 2012 Case Studies that UNICEF’s engagement at the decentralised level in programming and evaluation is a defining characteristic within the agency. This is seen as the best way to respond to two complementary needs: 1) to foster demand and use for different types of evaluations with governments, including for decision making, evidence-based policy advocacy and accountability purposes, and 2) to meet the capacity development and advocacy role at the country level in favour of evaluation, including the support of nationally-led evaluations as well as building a national demand for evaluations. Challenges in meeting high standards set out by the organisation in this regard include: finding the appropriate balance between an M&E staff member’s management of evaluations and the capacity development tasks at the country level; effectively matching the technical demands for national capacity development of governments and national actors with the skills of UNICEF M&E staff, particularly in middle income countries; encouraging communication with national governments on evaluation planning as well as dissemination and use; and encouraging partnerships and joint evaluations at the country level. The decentralised focus of evaluation is strongly entrenched in the organisation. Maintaining the decentralised evaluation function into the future, however, requires a response to some of these challenges, given the resource constraints and declining applicability of a one-size fits all approach for all types of countries and COs.

The guiding principles of the Evaluation Policy set out a number of management and organisational conditions that underline the importance of the function. One of these is that evaluations should serve an explicit management purpose. Respondents of the Case Study Report suggest that the key uses of evaluations are for decision making and for developing evidence-based policy advocacy. Knowledge generation was the evaluation purpose referenced the least often; those who did mention it mostly discussed knowledge generation as a by-product of evaluations rather than as their primary purpose. The 2012 Survey Report asked CO Representatives to provide a relative ranking of the four uses of evaluations; its conclusions were consistent with the Case Study Report: 50 per cent of CO Representatives feel that evidence-based policy advocacy is the most important purpose for carrying out evaluations within UNICEF. The second most important purpose is for decision making (28 per cent), third is for accountability (17 per cent), and knowledge management is viewed as the least important of the four purposes for evaluation. Irrespective of country size and income level, COs that have not carried out evaluations in the last three years appear to be using other types of assessments to generate knowledge, particularly for evidence-based policy advocacy. The most frequently cited assessments that fall outside of the definition of evaluations are reviews, surveys, research studies and situation analyses. The Case Study Report also noted an important level of confusion between these different types of assessments among UNICEF staff.

The emphasis on the **analysis of effectiveness and results** is seen across the organisation, with increased attention on the evaluation function in recent years. Across CO types and sizes, and in all regions, there is a consensus of a continued demand for UNICEF to participate in the development of quality evaluations for accountability purposes. Accountability to donors is also strongly felt at HQ levels. Most COs and senior management staff at the HQ level are driven to carry out evaluations in order to provide direct accountability for project/programme funds to the Executive Board, to donors and to partners. Senior managers are aware that in an extremely competitive and dynamic global funding environment, donors are placing a greater onus on UNICEF to demonstrate the impact and relevance of its programming choices, particularly across the key focus areas of health, education, HIV-AIDS, child protection, and emergency programming.

Another organisational condition outlined in the Evaluation Policy is the **financial resources** allocated for evaluation. The 2012 Survey of Representatives highlighted that financial resources, the leadership of the Representative, donor requests, and mandatory requirements are the key factors driving demand for evaluations at the CO level. At Regional Office (RO) and EO levels, availability of financial resources was also cited as limiting the ability of these offices to provide adequate guidance and support to the decentralised evaluation function.

UNICEF appears to be one of the few UN organisations to attempt to target a level of its expenditure for evaluation, studies and other research within its Evaluation Policy. It currently sets that target at between 3 and 5 per cent of programme expenditures. Its performance since 2006 suggests that it has not been able to meet this target consistently; from 2006 to 2010, expenditures on evaluation, studies and research have oscillated between 2.84 and 3.1 per cent of programme expenditure. The amount of spending on evaluations only is 0.33 per cent, well below the 1 per cent bar implied by the recent **System Wide Evaluation Report**.

The UNICEF Evaluation Policy, in line with the UNEG Norms and Standards for Evaluation, sets out a number of standards for the management of evaluations – from the design phase to implementation, dissemination and use of evaluations. **Ensuring adequate resources at the design stage of evaluations**, the Case Study Report outlined UNICEF staff views on the prerequisites for high quality evaluations and found a great deal of homogeneity across regions, country typologies and levels: a high degree of planning, a well-defined review process, and strong evaluation capacity at the national level, including independent consultants with the right language skills. Linked to planning is the requirement for clearly defined Terms of Reference, another factor commonly cited by respondents at all levels. The absence of any of these factors generally has a detrimental effect on evaluation quality.

The 2011 M&E Survey Report highlighted the overstretched nature of many M&E staff on a range of tasks, particularly regarding evaluation. When asked to calculate how their time is apportioned, evaluation received the lowest percentage, at roughly 14 per cent, compared with 29 per cent for situation monitoring, 24 per cent for planning and performance monitoring, 17 per cent for research, and 16 per cent for “other”. While the majority of M&E staff expressed satisfaction with the availability of human resources for their function, a significant minority (30 per cent) communicated dissatisfaction, especially on evaluation tasks, indicating a sense of overstretch.

RO M&E advisors also find it increasingly challenging to properly manage the number of evaluations being carried out within their regions using internal resources; all have developed some standardised mechanisms to better manage this oversight role, and some engaged external consultants to provide technical support in the drafting of TORs or in reviewing drafts. RO respondents overwhelmingly point to the need for UNICEF to carry out more strategic evaluations at the CO level as a key way to improve overall quality, including the oversight that they can provide.
Another standard for consideration is one which ensures the credibility of evaluations through **professional quality standards**. In most UNICEF COs, the 2012 Case Study Report presented a perception that new initiatives to develop accountability in evaluation quality, including GEROS, provided incentives within the agency to prioritise and improve evaluation quality and ensure professional standards of external evaluators. Respondents have noted a higher profile given to evaluations, a greater understanding of UNEG guidelines, knowledge of GEROS ratings and sensitivity to the independence of evaluations. The Case Study Report also noted this improvement in evaluations of emergencies in recent years. Drawing on 89 reports in 2010, the GEROS report noted a slight improvement in performance year-on-year, with 40 per cent of reports being rated overall as satisfactory according to UNICEF Evaluation Standards (36 per cent in 2010).

**Independence, impartiality and transparency** are also fundamental elements of the management of the design, implementation and dissemination of evaluations. Behind credible and good quality reports is the need for an independent, impartial and transparent process for carrying out evaluations. At UNICEF, the accountabilities outlined in the Evaluation Policy, particularly those around the CO Representative and the RO ensure that evaluation process follows good practice in independence, impartiality and transparency at all stages of the evaluation, from planning and design to implementation and reporting. Reports suggested both some good practices and some challenges in a consistent application of policies on independence, transparency and impartiality at the decentralised level. Accountabilities for M&E staff in particular are less clearly delineated in the current evaluation policy.

**Skills in evaluation, including knowledge of human rights and gender equality together with ethics and integrity** are also important standards outlined in UNEG guidance as well as the UNICEF Evaluation Policy. The 2012 Survey Report found a number of factors to be critical in the production of quality evaluations, including quality of the consultant engaged in the evaluation, Country Office leadership, and internal oversight of the evaluation; 68 per cent of respondents rated the quality of the consultant as “highly important”, compared to “country office leadership” at 49 per cent, and internal oversight of evaluation at 48 per cent. Some ROs have developed rosters of consultants in the region to recommend at the CO level in order to improve this element.

The 2011 M&E Survey Report highlighted the skills of UNICEF M&E staff themselves; it showed that 64 per cent of M&E staff have less than five years of M&E experience with UNICEF. M&E staff at UNICEF are confident of their skills in managing evaluations (developing the Integrated Monitoring and Evaluation Plan and TOR for evaluations) and performance monitoring, and to some extent in national M&E capacity development. They are less confident in evaluating peace-building and humanitarian situations. M&E staff familiarity with ethics and evaluations is low. The majority of respondents (66 per cent) had not received any training on ethical considerations in the design of evaluations, surveys or research.

With regard to human rights and gender equality, the GEROS found that “UNICEF has been at the forefront of international efforts to recognise human rights, gender equality, women and children’s empowerment, and socio-economic equity within the evaluation function. This is not being translated into strong performance in these areas.” The 2012 Case Study Report confirmed a more recent focus and interest on the equity approach, and developing improved methods of evaluating the effect of policies and strategies on the most vulnerable populations. Using the Monitoring Results for Equity System (MoRES), respondents cited the need to get government and other partners on board with paying greater attention to the bottom two per cent, or the most vulnerable populations, as they carry out evaluations. The Evaluation Office is leading efforts to better equip UNICEF offices and partners in evaluating pro-equity interventions.
The Evaluation Policy’s guiding principles also encompass the **dissemination of evaluation findings, recommendations and lessons learned at all levels of the organisation.** The UNEG Norms and Standards & Ethical Guidelines rightly emphasises that evaluations are valuable to the extent to which they serve the information and decision-making needs of intended users. Making the results of evaluations public is also part of accountability and transparency. In the 2012 Case Study Report, external communication and dissemination of evaluation findings among national partners was generally reported as strong among case study countries, regardless of the size of the CO or country typology. Most respondents cite the participation of partners as part of reference groups and engaging them in planning as key factors which facilitate dissemination and sharing of evaluation findings within the country. The 2012 Survey Report, however, confirmed that compliance in disseminating findings is good, but that COs have some improvements to make before meeting a standard of full compliance, particularly in how the government is systematically engaged in evaluation. The Annual Report on the Evaluation Function suggests that advocacy and oversight has resulted in improved compliance in the preparation of **management responses** and in uploading completed responses to the tracking system. Weak follow-up on evaluation recommendations directed to accountable stakeholders, however, remains the norm in the organisation.

In a decentralised programming environment, reaching the local population with the results of any evaluation is an important exercise. UNICEF’s evaluation policy is not unique in its aim to **reach local populations, including stakeholders,** although defining who the stakeholders are that should be reached is not made clear in the Evaluation Policy. UNICEF Country Offices are encouraged to translate executive summaries of evaluations into local languages and use other means to publicly inform stakeholders of findings and for learning. The extent to which this guideline has been closely applied is unclear and not reported.

At the senior management level at UNICEF HQ, there is a broad appreciation for the role played by the EO to institute and monitor evaluation quality as a clear indicator that can be tracked. The GEROS results are shared widely at Headquarters and Executive Board levels. Responses during the 2012 Case Study Report suggest that the **relevance and use of evaluation findings by the Executive Board** is strong and well appreciated.

The MTSP focus on **humanitarian evaluations** brings to light the efforts that have been made by the EO to develop tools for these situations. The EO has been lauded for taking steps in improving overall evaluation quality for emergency evaluations, such as developing e-learning tools, standard operating procedures and improved guidance for emergency evaluations in the Programme Policy and Procedures Manual (PPPM).

Strengthening **evaluation of UNICEF focus areas, strategies and operational effectiveness** is another priority set out by the MTSP. UNICEF’s Evaluation Policy is more explicit in its demands for evaluation coverage at the CO level. In order to strengthen the evaluation function, it gives responsibility at all levels (CO, RO and Division Directors) to **Institute measures to ensure that evaluations are strategically selected and that all major programme components are evaluated during the programme cycle, allocating adequate resources.** As noted above, the allocation of resources remains a contentious issue, with respondents at all levels in the Case Study Report noting challenges in securing resources for their strategic evaluations. The Case Study Report also highlighted the evaluation coverage gap that exists in certain UNICEF COs which have not carried out evaluations over the past two years. While they do carry out other types of assessments in the COs, the specific importance of evaluations in UNICEF programming is not reinforced.
Two additional areas of focus for evaluation, as highlighted in the MTSP, are organisational capacity in evaluation and a heightened management attention to the evaluation function. Both are mutually reinforcing: the increased attention brought about by rating tools such as GEROS and the vision dashboard have led to the need, at RO and CO level, to develop the organisational capacity to manage and improve the function as a response. As noted in other findings, RO systems have been developed for quality assurance at the CO level. Heightened coordination between Evaluation Specialists at the CO, RO and EO levels, through network meetings and webinars have helped to share lessons and bring greater professionalism to the function.

UNEG has developed a Normative Framework for the Evaluation Self-assessment in order to allow member organisations to review their own Evaluation Policies and functions. It outlines the key Norms and Standards alongside assessment criteria in order to highlight potential gaps. Many of the areas outlined are covered in the above findings. The 2012 Case Study Report and 2011 M&E Survey Report underline the important role of M&E staff in carrying out the responsibilities and duties outlined in the Evaluation Function. While overall accountability within the CO lies with the Representative, the accountabilities of staff responsible for evaluation are not clearly specified. At the design and implementation stage of evaluations, for example, the Evaluation Policy is silent on who is accountable for the development of independent TORs, the selection of consultants, or the quality review of draft reports.

UNICEF’s Evaluation Policy is silent on its disclosure of evaluation products at all stages. This includes the sharing of an evaluation plan with key stakeholders, sharing of final TORs with stakeholders, and how evaluations should be made publicly available at local levels. The Evaluation Policy is clear on the sharing and dissemination of evaluation results and findings, but in practice, there are some gaps in consistent implementation.

Conclusions

The purpose of this review is to assess the demand for evaluation and the utilisation of evaluation results and, making use of existing documentation, to make an overall assessment of the current status of the evaluation function in UNICEF. The consultancy used the 15 guiding principles outlined in the Evaluation Policy to present findings on UNICEF’s decentralised evaluation function and the governance and management of evaluation activities within UNICEF, including the role of management at all levels in addressing evaluation needs and demands, in ensuring effective utilisation of evaluation results, and in taking account of cross cutting issues.

Upon review of the Evaluation Policy and in particular its guiding principles under the context of demand for evaluation and the quality and use of evaluations, the policy remains relevant and closely aligned to the UNEG Norms and Standards. The decentralised evaluation function championed by the Evaluation Policy aligns well with the operational nature of UNICEF’s decentralised programming environment, given UNICEF’s role to act as an authoritative voice on children’s issues at the country level and the necessity to reflect the differences and particularities of each country and region. The decentralised nature of the evaluation policy is also relevant in the current policy context of the UN. As noted, it is well aligned to the most recent QCPR which highlights a continued focus on decentralised evaluation and national capacity development. The challenge lies in fully implementing and in some cases clarifying the policy, rather than radically altering it.

The Evaluation Policy does have some drawbacks, however. First of all, the policy fails to delink the funding for evaluations from the size of programming expenditures. Under the current policy, the dollar amount of funding for evaluations is proportional to the size of the programme. This will have increasing importance in the coming years as COs in MI countries with fewer resources but greater technical evaluation needs try to remain relevant in the evaluation environment. This will also have relevance in COs that require human and financial resources for both evaluations and national capacity development in evaluation. The Evaluation Policy isn’t clear on whether or not country-led evaluations, rather than
UNICEF-led evaluations, are adequate in the evaluation of “all major programme components” in a cycle. This obligation within the Policy has been interpreted in different ways at the CO level.

The Evaluation Policy lacks a clear set of accountabilities for M&E or evaluation personnel. This may have been the intent of the Policy so as to leave flexibility at the decentralised CO level to set out the resources and accountabilities required. Data suggests that in most cases, COs have indeed committed to staffing for M&E resources. Therefore, the absence of these staff within the policy obscures their critical role within the CO for evaluation as well as research, surveys and for tasks such as performance monitoring or the development of national capacities in evaluation. In fact, the study has shown that staff allocation of time does not necessarily favour evaluation over other tasks, including studies and research, or other monitoring tasks within the organisation. The incentives and skills required for M&E staff to champion national M&E capacity development are similarly undefined.

While the Evaluation Policy highlights the importance of country-led evaluations and joint evaluations, it does not explicitly couch them within the broader Aid Effectiveness agenda. A more clear elucidation of how the aid effectiveness agenda is implicated within the evaluation function would provide a stronger impetus to UNICEF COs to leverage partnerships at the HQ, RO and CO levels, and to advocate for greater country leadership in evaluation.

Finally, new and emerging issues now populate the aid and programmatic agenda and these are not reflected in the Policy. In particular, the Equity Approach within UNICEF, which has received a good amount of attention in recent years by the EO, is not present in the current policy.

In terms of how the Evaluation Policy is implemented, overall the synthesis study suggests that UNICEF’s current evaluation practice holds up well to the guiding principles espoused by the Evaluation Policy. UNICEF staff have noted a continued demand for evaluations both within and outside the organisation for the purposes of evidence-based policy advocacy, decision making and accountability. While overall, financial resources, leadership and donor demand are strong determinants that prioritise evaluations, national priorities for evaluation are also playing an increasing role in determining demand for evaluations, particularly in MI countries. In addition, it is noted that joint evaluations with other UN agencies or with government partners and multi-country evaluations are not well defined or strongly integrated into UNICEF’s current portfolio of evaluations, a problem that is shared in the entire UN system.

In the context in COs where evaluations were not carried out, the CO had not done so for reasons which are often country-specific and which tend to focus on a lack of external demand or priority from the government. This underlines the important role that UNICEF has in building national capacity in evaluation. In some cases, while Senior Management recognised the need for some level of review of performance, this was not done through a formal evaluation but rather through informal monitoring channels, studies or other research. In humanitarian contexts, the demand for evaluation is determined in part by funding thresholds. Carrying out humanitarian evaluations currently reflects an imbalance between demand for such evaluations and the technical know-how to produce and oversee them, and has placed a considerable burden on the EO for management and oversight.

With regards to evaluation quality, the GEROS system and management dashboard have fomented the correct incentives to produce improved quality evaluations at the CO level according to the Evaluation Policy and UNEG guidelines. The EO and ROs have invested significant resources in tracking and supporting evaluation quality with a view to improving UNICEF’s overall evaluation credibility. This effort has not gone unnoticed by Senior Management at all levels of the organisation and has helped to build and sustain a culture of evaluation in the organisation.

This comes, however, with some concerns about the extent to which evaluations should be the only product assessed for quality. Studies, reports, and surveys are other types of assessments which are not included in the GEROS mechanism. Other concerns revolve around the need to balance the quality role at the CO level with the myriad other tasks to which evaluation specialists attend, including monitoring.
situation assessment, and capacity building. This points the way to a greater reliance on partnership in evaluation in order to leverage knowledge and take advantage of skills in other organisations.

Not surprisingly, COs with highly rated evaluations are characterised by strong planning, well developed Terms of References, good national evaluation capacity and an adherence to a quality assurance process that often includes RO involvement. Constraints in terms of having the right mix of staff skills and sufficient financial resources to carry out an evaluation agenda will continue to persist; these constraints suggest that a one-size fits all approach within UNICEF is becoming more and more difficult to administer. For example, certain MI countries require more advanced expertise in capacity development in order to support government with policy evaluations, while other countries require advanced expertise in emergency evaluations. Some COs face challenging political contexts for evaluations, while others face resource constraints and will rely on partnerships and a strong regional presence to carry out evaluations.

Quality is also affected by the skills of M&E staff: M&E staff at UNICEF are confident of their skills in managing evaluations and performance monitoring, and to some extent in national M&E capacity development. They are less confident in evaluating emergency, peace-building and humanitarian situations and their familiarity with ethics for evaluations is low.

Evaluations have no value to UNICEF or its partners if they are not adequately and widely used for their intended purposes. Concerning the use of evaluation results, evaluations appear to be most frequently used for decision-making purposes and evidence-based policy advocacy. Their greatest use is at the country level, where UNICEF uses evaluation results to refocus or advocate for its programming interventions with partners in favour of children’s rights. Evaluation results are often used to formulate management responses but their tracking and implementation has only recently become more systematic at the country level.

External communication and dissemination of evaluation results is seen as generally good at the country level, especially when government partners or other stakeholders are involved in the planning and design of evaluations, or where they participate in an advisory group capacity. However, language barriers and full engagement with governments in some countries has hindered full transparency of evaluations, including issues such as stakeholder presentations and response to evaluation findings, and whether or not the local language was used for evaluation summaries. Communications and dissemination at headquarters is also seen as relatively strong in responding to donor needs, and use of evaluations at headquarters results in stronger technical inputs to programme strategies and the UNICEF MTSP. The processes and dissemination strategies for the internal sharing of evaluation results for knowledge generation, however, are less fully developed within the organisation. As such, knowledge generation and management emerging from evaluations can be improved.

**Recommendations**

There are three main recommendations offered on the Evaluation Policy:

Firstly, the UNICEF Executive Board, through the Evaluation Policy, should commit to spending one per cent of programme resources specifically on evaluation and should seek a way to ensure minimum funding levels for evaluations and capacity development for country-led evaluations. Secondly, the UNICEF Evaluation Policy should more clearly incorporate the role of the M&E staff at the CO level. While accountability for the decentralised evaluation function rests with the Country Representative, this review suggests that in some cases, delineating some accountability at the level of the M&E specialist could be beneficial. Finally, the UNICEF Evaluation Policy should be updated to include greater clarity on the aid effectiveness agenda, the requirement to evaluate every programme at least once per cycle, and the Equity approach.

Five recommendations are given on ways to improve the implementation of the Evaluation Policy:

- The EO should propose guidance and oversight on issues such as sharing of evaluation work programmes, TORs and draft reports with stakeholders; obtaining their views; and managing
dissenting opinions on evaluations. While the UNICEF Evaluation Policy is clear with regard to engagement with country-led programming, national capacity development, and engaging with stakeholders throughout the evaluation process, several findings in this review have shown that improvements need to be made in practice. This points strongly to the transparency criterion outlined in the UNEG guidelines. The EO should propose guidance and oversight on issues such as sharing of evaluation work programmes, TORs and draft reports with stakeholders; obtaining their views; and managing dissenting opinions on evaluations. The EO should also work to clarify in the Evaluation Policy or in other guidance exactly how evaluations should be made public at the CO level, including the language in which reports or summaries of reports should be made available and their dissemination.

- UNICEF ROs and EO need to provide better oversight and incentives to UNICEF’s role in capacity development of national evaluation systems, with a focus on leveraging partnerships with other actors in order to be more efficient in this regard. M&E staff and RO staff feel that without a more clear focus on national capacity development, this important function will be seen as less important than evaluations. The EO should work to encourage a stronger oversight or incentive mechanism to monitor capacity development steps in countries.

- The EO should clarify the extent to which joint evaluations are encouraged and work with UNEG and other agencies in learning more about the value of joint evaluations in the current programming environment. As highlighted in the 2012 Survey Report, more clarity and oversight is required in the implementation of the Evaluation Policy on the extent to which joint evaluations are encouraged. This recommendation goes beyond a single UN organisation and reflects a wider need of the UN to take stock of the culture gaps in engaging in more joint evaluations.

- The EO should develop KPIs for the design of evaluations to help ensure better compliance with accepted good practices. Several findings in this review highlight the uneven application of good practices in independence and transparency of evaluations at all stages of the evaluation process.

- The EO and ROs should agree on a way to identify COs that have not carried out any evaluations for more than two years and assist them in developing an evaluation strategy. EO and RO should then take steps to ensure that those COs are supported with clear strategies to gather evaluative data for short, medium and long-term planning and decision making.
## Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCC</td>
<td>Core Commitments for Children</td>
</tr>
<tr>
<td>CEDAW</td>
<td>Convention on the Elimination of Discrimination against Women</td>
</tr>
<tr>
<td>CEE/CIS</td>
<td>Central and Eastern Europe/Commonwealth of Independent States</td>
</tr>
<tr>
<td>CMT</td>
<td>Country Management Team</td>
</tr>
<tr>
<td>CO</td>
<td>Country Office</td>
</tr>
<tr>
<td>CPAP</td>
<td>Country Programme Action Plan</td>
</tr>
<tr>
<td>CPD</td>
<td>Country Programme Document</td>
</tr>
<tr>
<td>CRC</td>
<td>Convention on the Rights of the Child</td>
</tr>
<tr>
<td>DfID</td>
<td>Department for International Development, UK</td>
</tr>
<tr>
<td>DHS</td>
<td>Demographic and Health Surveys</td>
</tr>
<tr>
<td>DPP</td>
<td>Division of Policy and Practice</td>
</tr>
<tr>
<td>EAPRO</td>
<td>East Asia and Pacific Regional Office</td>
</tr>
<tr>
<td>EMOPS</td>
<td>UNICEF Emergency Operations</td>
</tr>
<tr>
<td>EO</td>
<td>Evaluation Office</td>
</tr>
<tr>
<td>ESARO</td>
<td>East and Southern Africa Regional Office</td>
</tr>
<tr>
<td>FAO</td>
<td>Food and Agriculture Organisation</td>
</tr>
<tr>
<td>GEROS</td>
<td>Global Evaluation Reports Oversight system</td>
</tr>
<tr>
<td>HQ</td>
<td>Headquarters</td>
</tr>
<tr>
<td>IMEP</td>
<td>Integrated Monitoring and Evaluation Plan</td>
</tr>
<tr>
<td>LI</td>
<td>Low Income</td>
</tr>
<tr>
<td>LMI</td>
<td>Lower Middle Income</td>
</tr>
<tr>
<td>M&amp;E</td>
<td>Monitoring and Evaluation</td>
</tr>
<tr>
<td>MDG</td>
<td>Millennium Development Goal</td>
</tr>
<tr>
<td>MENA</td>
<td>Middle East North Africa</td>
</tr>
<tr>
<td>MI</td>
<td>Middle Income</td>
</tr>
<tr>
<td>MICS</td>
<td>Multiple Indicator Cluster Survey</td>
</tr>
<tr>
<td>MTSP</td>
<td>Medium-Term Strategic Plan</td>
</tr>
<tr>
<td>NGO</td>
<td>Non-Governmental Organisation</td>
</tr>
<tr>
<td>NMECD</td>
<td>National Monitoring and Evaluation Capacity Development</td>
</tr>
<tr>
<td>OCHA</td>
<td>Office for the Coordination of Humanitarian Affairs</td>
</tr>
<tr>
<td>OPT</td>
<td>Occupied Palestinian Territory</td>
</tr>
<tr>
<td>PD</td>
<td>Programme Division</td>
</tr>
<tr>
<td>Acronyms</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>PPPM</td>
<td>Programme Policy and Procedures Manual</td>
</tr>
<tr>
<td>QCPR</td>
<td>Quadrennial comprehensive policy review</td>
</tr>
<tr>
<td>RO</td>
<td>Regional Office</td>
</tr>
<tr>
<td>ROSA</td>
<td>Regional Office for South Asia</td>
</tr>
<tr>
<td>RR</td>
<td>Regular Resources</td>
</tr>
<tr>
<td>RWP</td>
<td>Rolling Work Plan</td>
</tr>
<tr>
<td>SITAN</td>
<td>Situation Analysis</td>
</tr>
<tr>
<td>TACRO</td>
<td>Latin American and Caribbean Regional Office</td>
</tr>
<tr>
<td>TOR</td>
<td>Terms of Reference</td>
</tr>
<tr>
<td>UDHR</td>
<td>Universal Declaration of Human Rights</td>
</tr>
<tr>
<td>UMG</td>
<td>Universalia Management Group</td>
</tr>
<tr>
<td>UMI</td>
<td>Upper Middle Income</td>
</tr>
<tr>
<td>UN</td>
<td>United Nations</td>
</tr>
<tr>
<td>UNCT</td>
<td>United Nations Country Team</td>
</tr>
<tr>
<td>UNDAF</td>
<td>United Nations Development Assistance Framework</td>
</tr>
<tr>
<td>UNDP</td>
<td>United Nations Development Programme</td>
</tr>
<tr>
<td>UNEG</td>
<td>United Nations Evaluation Group</td>
</tr>
<tr>
<td>UNESCO</td>
<td>United Nations Educational Scientific and Cultural Organisation</td>
</tr>
<tr>
<td>UNHCR</td>
<td>United Nations High Commissioner for the refugees</td>
</tr>
<tr>
<td>UNICEF</td>
<td>United Nations Children’s Fund</td>
</tr>
<tr>
<td>UN WOMEN</td>
<td>United Nations Entity for Gender Equality and the Empowerment of Women</td>
</tr>
<tr>
<td>WCARO</td>
<td>West and Central Africa Regional Office</td>
</tr>
</tbody>
</table>
Contents

Executive Summary i

1. Introduction 1

2. Methodology 5

3. Findings on the Evaluation Policy and Function 9
   3.1 Findings on the 15 Evaluation Policy Guidelines 9
   3.2 Findings on MTSP Priorities and Other Areas not Covered by the Current UNICEF Evaluation Policy 38

4. Conclusion and Recommendations 41
   4.1 Conclusions 41
   4.2 Recommendations 43

Exhibits

Exhibit 2.1 Mapping of Reports that fed into the Synthesis Report 7
Exhibit 3.1 Percentage of respondents who rated at “high importance” the following factors which drive demand for evaluation at the CO level (N=52) 18
Exhibit 3.2 Percentage of COs which comply with good practices in their last three evaluations, according to M&E staff (N=86) 26
Exhibit 3.3 Percentage of CO Representatives who rated “highly important” each of the factors which affect the production of quality evaluations (N=55) 28
Exhibit 3.4 Years of UNICEF M&E Experience of M&E staff, in per cent (N=95) 29
Exhibit 3.5 Percentage of M&E staff with more than five years of UNICEF M&E Experience by Size of CO and Country Typology (N=93) 30
Exhibit 3.6 M&E staff Respondents Average Knowledge of M&E Technical Areas (N=95) 31
Exhibit 3.7 Latest participation in UNICEF/UN supported M&E training, in per cent of total responses (N=82) 32
Exhibit 4.1 Framework for Analysis 54
### Appendices

<table>
<thead>
<tr>
<th>Appendix I List of Findings</th>
<th>47</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix II Terms of Reference</td>
<td>50</td>
</tr>
<tr>
<td>Appendix III Framework for Analysis</td>
<td>54</td>
</tr>
<tr>
<td>Appendix IV Survey Questionnaire</td>
<td>60</td>
</tr>
<tr>
<td>Appendix V Case Studies Interview Protocol</td>
<td>70</td>
</tr>
<tr>
<td>Appendix VI Case Studies Interview Protocol (Countries with No Evaluations Conducted)</td>
<td>72</td>
</tr>
<tr>
<td>Appendix VII Key Informants Interview Protocol</td>
<td>74</td>
</tr>
<tr>
<td>Appendix VIII Key Informants Interview Protocol for UNICEF Board Members</td>
<td>76</td>
</tr>
</tbody>
</table>
1. Introduction

1. The current UNICEF Evaluation Policy was approved by the UNICEF Executive Board in September 2008. Following the approval of the policy, an Executive Directive was sent to all UNICEF Offices in March 2009. The policy aims to establish a common institutional basis for the UNICEF evaluation function and seeks to increase transparency, coherence, and efficiency in generating and using evaluative knowledge for organisational learning, managing for better results and supporting accountability.

2. According to the UNICEF Evaluation Policy, an evaluation should provide evidence-based information that is credible, reliable and useful, enabling the timely incorporation of findings, recommendations and lessons into the decision-making processes. UNICEF’s Evaluation Policy is guided by the United Nations Evaluation Group Norms and Standards for Evaluation in the United Nations System, approved by UNEG in April 2005. The UNEG Ethical Guidelines highlight that evaluations are valuable to the extent to which they serve the information and decision-making needs of intended users.

3. As highlighted in the Annual review on the evaluation function and major evaluations in UNICEF (E/ICEF/2012/13), “the organisational context and the wider international settings have seen major changes in recent years, with more changes expected in 2012. UNICEF therefore proposed to undertake a thorough review of its evaluation function and the Evaluation Policy, with a view to updating the policy for submission to the Executive Board in mid-2013.”

4. This synthesis report is one of several inputs into an internal UNICEF review which may contribute to an updating or renewal of the Evaluation Policy and function. The main purpose of the synthesis review is to look at experience and performance of the function to date using inputs from various sources and make forward looking recommendations on measures and adaptations required.

5. This report is divided into four sections. Section 1 presents the background of the Internal Review of UNICEF’s Evaluation Policy and Function. Section 2 presents the methodology used to produce this report. Section 3 presents the main findings emerging from the review of the Evaluation Policy, based on the guiding principles and the priorities outlined in the MTSP. Section 4 comprises the conclusion and recommendations emanating from the review. A list of people interviewed, the original Terms of Reference, and all interview protocols are included in the Appendices.

UNICEF Evaluation Policy

6. In June 2004, the Executive Board reviewed the “Progress report on the evaluation function in UNICEF” and made recommendations to sharpen the strategic focus of evaluation work plans, improve efficiency and raise standards of evaluation work. The Board also re-emphasised that the evaluation function should be carried out in consultation with national authorities. UNICEF was encouraged to further strengthen the evaluation function, highlighting the following: (a) focusing on fewer, high-quality studies with a strong emphasis on analysing effectiveness and results; (b) improving evaluation standards at the country level; (c) accelerating progress towards joint evaluation work; (d) strengthening national evaluation capacity; and (e) identifying best practices and enabling findings to feed systematically into organisational learning.

7. The guidance from UNICEF’s Executive Board has been supplemented by inter-agency collaboration on the evaluation function, beginning with its participation in the UN Inter-agency working group on evaluation and the subsequent United Nations Evaluation Group (UNEG) Norms and Standards for Evaluation in the United Nations System, approved by UNEG in April 2005. At its July 2006 substantive session, the Economic and Social Council took note of the endorsement of the Norms and Standards for Evaluation in the United Nations System.
evaluation by UNEG as constituting a contribution to strengthening evaluation as a United Nations system function.

8. In 2006, a panel of six international experts conducted a peer review of the evaluation function in UNICEF. The findings were presented to senior management of UNICEF and an informal meeting of the Executive Board. The panelists concluded that “the central Evaluation Office demonstrated a high level of independence and produced evaluations which are credible and useful for learning and decision-making within the organisation.” According to the panelists, the decentralised evaluation system was appropriate for the operational nature of the organisation, however its credibility and usefulness were limited by critical gaps in resources and organisational constraints. The peer review panel concluded that UNICEF could strengthen the evaluation function by: adopting a clear and comprehensive Evaluation Policy document consistent with UNEG Norms and Standards, using a more predictable budget for evaluation, undertaking additional interventions to strengthen and support field offices, and improving the use of results-based management throughout the organisation. The importance of working with national governments was also a focus of attention at both the Executive Board level and within the aid effectiveness agenda of the international development community.

9. In 2008, UNICEF’s own Evaluation Policy – developed from the UNEG norms and standards as recommended by the Peer Review Panel – was approved by the UNICEF Executive Board. Following approval of the policy, an Executive Directive was sent out to all UNICEF offices in March 2009. The policy aimed to establish a common institutional basis for the UNICEF evaluation function in order to increase transparency, coherence, and efficiency in generating and using evaluative knowledge for organisational learning, managing for results, and supporting accountability.

Roles and Responsibilities

10. UNICEF’s evaluation function is managed through a shared responsibility compact involving the Evaluation Office (EO), regional offices (RO), and country offices (CO), with each assuming distinct roles and responsibilities. The following responsibilities are defined in UNICEF’s Evaluation Policy and the Executive Directive:3

- **Country Offices** are responsible to: (a) Undertake evaluation in accordance with the principles of evaluation in UNICEF; (b) Ensure the evaluable of new programmes and initiatives at the planning stage, including a clear statement of results, the identification of risks, relevant performance indicators, and the setting of baselines at the outset; (c) Enhance the monitoring of programmes to allow UNICEF to more systematically conduct evidence-based performance and impact evaluations; (d) Allocate funds at the start of each year for the Integrated Monitoring and Evaluation Plan; (e) Ensure that all programme staff have a foundational knowledge of evaluation principles and types and their application, and that external evaluators employed by UNICEF have the competencies to ensure that credible evaluations are undertaken; (f) Make sure that existing staff engaged in evaluation have access to professional development opportunities, and ensure that new appointments to such posts are made against the evaluation competencies agreed by UNEG, while seeking technical clearance from the Regional Office and Evaluation Office, as appropriate; (g) Institute measures to ensure that evaluations are strategically selected and that all major programme components are evaluated during the programme cycle, allocating adequate resources; (h) Maintain a level of independence in evaluation by increasing joint evaluation with partners, encouraging country-led evaluations, and ensuring that managers of UNICEF programmes under evaluation do not have decision-making responsibility for evaluation terms of reference, the selection of evaluation consultants, or the acceptance of evaluation reports; (i) Ensure that evaluation recommendations are fully considered with concerned partners, that accepted

---

3 Taken from Appendix A of the Terms of Reference.
recommendations are acted on, and that annual reports include a statement on the status of evaluation follow-up; (j) Submit completed evaluations to the UNICEF on-line evaluation database, within three months of their completion.

- **Regional Offices** are responsible to: (a) Provide support and technical advice to country offices to assist them in fulfilling their obligations under the Evaluation Policy; (b) Monitor and report annually on the quality of evaluation in the region, the appropriateness of staffing and financial resources dedicated to the evaluation in the region, and country arrangements for ensuring independence, transparency and impartiality in country-level evaluation; (c) Maintain a tracking system for the implementation of any regional evaluations and country evaluations with implications for regional strategies; (d) Maintain a roster of pre-qualified evaluation consultants and institutes from within the region.

- The **Evaluation Office** is responsible to: (a) Develop corporate strategies for the improvement of the evaluation function and issue and update corporate guidance on evaluation practices; (b) Manage the evaluation quality rating process and rate all evaluations within six months of submission; (c) Report annually to the Executive Board on evaluation function, and to the Evaluation Committee on evaluation needs and expenditures across UNICEF; (d) Provide technical assessment of candidates applying for positions related to evaluation in UNICEF; (e) Ensure that evaluation staff are aware of relevant UNICEF, UNEG and third-party evaluation training materials and courses; (f) Maintain a global roster of pre-qualified evaluation consultants and institutes; (g) Maintain a network of communication and exchange with evaluation staff, providing them with updates on evaluations findings, events and methodologies; (h) Track the implementation of accepted evaluation recommendations from global thematic and institutional evaluations, reporting annually on the status of follow-up to the Evaluation Committee.

11. In summary, the Evaluation Policy and Executive Directive assign responsibilities for evaluation largely at the level of the CO, with accountability resting primarily with the Representative. Strategy, oversight, technical assistance, capacity development and reporting are the key responsibilities of the ROs and the EO.

**Programme Policy and Procedures Manual**

12. While the Evaluation Policy and Executive Directive set out clear roles and responsibilities and management measures to strengthen the evaluation function in the organisation, the Programme Policy and Procedures Manual (PPPM) provides institutional guidance to support both monitoring and evaluation in Chapter 5, outlining several layers of involvement and responsibility for M&E. These are summarised below.

- **The M&E officer:** The PPPM defines the M&E officer as the principal technical advisor for the office, especially during the TOR design phase and implementation phase. Where necessary, technical support from the regional level, UNICEF HQ, or external sources may be sought. It also notes that many COs create an internal team to support the M&E function.

- **Programme/project officers:** The PPPM notes that programme/project officers have primary responsibility for identifying M&E and initial research needs, for reading and considering the findings of evaluations and for ensuring a response to the results.

---

4 In practice, the monitoring function is managed through a shared responsibility involving the CO, RO, EO and the Statistics and Monitoring Section (SMS).

5 While this assessment uses the term M&E specialist throughout, staff targeted for the survey were those who were principally responsible for the M&E functions within the CO. They carry a range of titles such as M&E specialist, M&E officer, Planning Monitoring and Evaluation (PM&E) Officer, M&E staff, and so on.
• **Government and other partners in M&E planning**: While noting that M&E planning has been a co-responsibility of UNICEF and governments, the PPPM highlights that the UNICEF Integrated Monitoring and Evaluation Plan (IMEP) must be consistent with the United Nations Development Assistance Framework (UNDAF) M&E plan – which means that other UN agencies are stakeholders in the UNICEF IMEP. It also notes the role of the ROs, in particular, in the review of the IMEP.

• **Country Management Team (CMT)**: The CMT has responsibility in the entire M&E process, and should actively engage at several points, including reviewing the comprehensiveness of the IMEP, allocating resources, setting priorities, and developing a management response.

• **The Representative**: In line with the Evaluation Policy, the Representative\(^6\) carries key responsibilities with regard to M&E, ensuring:
  - the overall programme team is meeting its IMEP accountabilities;
  - the M&E officer or focal point has access and the support needed to ensure a good technical product;
  - the overall quality of UNICEF-supported research, monitoring and evaluations;
  - that adequate resources are dedicated to M&E activities (along with Regional Director);
  - that results of evaluations are discussed and follow-up actions recorded in a meeting of the CMT; and
  - systematic follow-up to an evaluation or monitoring report.

---

\(^6\) The PPPM sometimes refers to the Senior Programme Officer, which for the purposes of this survey was defined as the Deputy Representative. Together they are defined in this report as the Senior Management Team.
2. Methodology

13. Terms of Reference for this evaluation were agreed upon in August 2012 with the Evaluation Office (see Appendix II). An inception report, outlining the proposed review methodology and the key instruments for data gathering, was then drafted and shared with both the EO and RO Evaluation specialists. Adjustments to the approach and methodology proposed were identified and the final inception report was revised and finalised. The Evaluation Framework is presented in Appendix III.

14. The synthesis report makes use of primary data, produced through a survey of CO Representatives, a set of case studies and a number of key informant interviews, and secondary data in the form of reports and studies produced for the EO previously. Consultations in both the primary and secondary data collected were limited to UNICEF staff. The following reports acted as key sources and references for this synthesis review:


15. UNICEF Country Office Representatives were asked to participate in an on-line survey covering three key areas of the evaluation function, namely:

- Assessing the demand for evaluations both by UNICEF and its partners;
- Assessing the quality of evaluation within UNICEF and contributory factors that lead to good quality, and;
- Assessing the use of evaluation results and the importance of management responses and disseminations strategies both within UNICEF and externally.

16. The survey report summarised and disaggregated data at several levels, including CO size, country typology and UNICEF Region. In total, 67 CO Representatives responded to the survey for a response rate of 51.5 per cent. Four CO Representatives did not answer the survey due to their lack of knowledge of evaluation in their current CO. Of the remaining survey responses, all observations, including partially completed surveys, were included. In total, data was available for 63 COs. Three principle limitations were noted in the survey: the low response rate which limits any broad generalisations by region difficult; the limitation of asking new CO Representatives about the evaluation function in their country, which was

---

7 UNICEF Region: There are 7 UNICEF Regional Offices: Central and Eastern Europe/Commonwealth of Independent States (CEE/CIS), East Asia and Pacific Region (EAPRO), Eastern and Southern Africa Region (ESARO), Latin America and Caribbean (TACRO), Middle East and North Africa (MENA), South Asia (ROSA), and West and Central Africa (WCARO).

Size of CO: As defined by the CO overall budget. Budget is determined based upon total expenditure tables for end of year 2010, including Regular Resources, Other Resources and Support Budget.

- Small CO: Less than $5 million in total expenditures in 2010;
- Medium CO: Between $5 million and $20 million in total expenditures in 2010;
- Large CO: Greater than $20 million in total expenditures in 2010.

Typology of country: As defined by the World Bank Atlas, the three groups are:

- Low income (LI) $1,005 or less in annual per capita income;
- Lower middle income (LMI): $1,006 to $3,975 in annual per capita income.

Upper middle income (UMI), which includes per capita income of $3,976 to $12,275 and also includes the few UNICEF countries which are defined by high income per capita income of more than $12,275.
reflected in the ability to opt-out and; the respondent bias of certain questions which asked CO Representatives to respond to their own leadership.

2012 Case studies and key informant interviews (2012 Case Study Report)

17. For the purpose of better understanding evaluation demand, quality and use in UNICEF, 47 interviews were carried out with senior managers at Headquarters, Executive Board Members, Regional Directors and Regional M&E/ Evaluation Advisors as well as with UNICEF staff in twelve country case studies. Both the case studies and the key informant interviews were completed through phone interviews based on interview protocols. In the country case studies, up-to four staff members in the Country Office (CO) were targeted: the Representative, Deputy Representative, a Programme Officer, and the Evaluation Specialist. Case study countries were selected to cover variations in CO size, country typology, and emergency context. An additional criterion for selection was the quality of evaluation submissions in the past three years; based on the Global Evaluation Reports Oversight System (GEROS) and evaluation information collected through Country Office Annual Reports, five COs assessed with high quality evaluations, four COs assessed with low quality evaluations and three COs which had not reported completing evaluations in recent years, were included as case studies.


18. This report focuses on the status of the UNICEF decentralised monitoring and evaluation functions; identification of strengths and weaknesses, including good practices and systemic constraints, within those functions; and areas within them that may require improvement. Data collection was carried out in the form of a survey of decentralised UNICEF M&E staff at the country office level. Particular attention is given to the evaluation function; one objective of the survey is to act as a progress report on the UNICEF Evaluation Policy and Executive Directive. The survey was conducted online over a five-week period between 17 June and 7 July, 2011. In total, data was collected for 96 COs, which gave a response rate of 74 per cent in terms of the number of country offices that responded relative to the total number of country offices in UNICEF (130).

Other Documents used for this review (GEROS, Qualitative Review of Annual Reports, etc.)

19. Several additional documents were reviewed for the purposes of this synthesis report. These include the Global Evaluation Reports Oversight system (GEROS), the Global evaluations dashboard, and the Qualitative Review of Annual Reports – Evaluation component (Qualitative Review), which provides an analysis of the country office annual reports. Regional reports on the evaluation function were also reviewed, such as the TACRO review of the evaluation function and the MENA M&E Scorecard Report. The GEROS has been developed in response to a need to provide decision makers in UNICEF with information about evaluation reports that better support using and improving the knowledge generated by the evaluation function. The quality criteria and scoring system available on the GEROS will be used in addition to the data collected through case studies, key informant interview and survey. It is the relationship between the perceptions of the informants and the objective information contained on the GEROS that will allow the evaluation team to identify trends on the evaluation function and policy. The Qualitative Review of Annual Reports – Evaluation component, was another product used to inform the synthesis report. The review added to some of the summative information on evaluations carried out the CO level. Once again, the relationship between the information contained in these reports and data generated through case studies and key information interviews help to establish perspectives on the Evaluation Policy. Exhibit 2.1 illustrates the key products and how they contributed to this final synthesis report.
Bringing the products together into a final synthesis

20. Using the UNICEF Evaluation Policy as a framework, this report reviews the elements of the experience and performance of the Evaluation function and Policy within UNICEF. Findings of this report will aim to contribute to informing the overall evaluation policy dialogue within the agency, with commensurate suggestions for strengthening and forward-looking guidance.

21. This report’s findings are organised by the 15 key guiding principles of the evaluation function, as highlighted in the Evaluation Policy.¹ These principles are as follows:

1) Evaluation must respect the principle of universality and country-driven programming;
2) UNICEF is to support programme countries in evaluating their own programmes and to contribute to the strengthening of evaluation capacity in these countries;
3) Whenever possible, evaluations must be undertaken in partnership with national authorities, with the United Nations system and with interested partners;
4) Evaluations at all levels (whether of strategic governance, global, regional or country programmes, or projects) must serve an explicit management purpose;
5) It is important to preserve the decentralised nature of the evaluation system in UNICEF;
6) An amount totalling 3 to 5 per cent of programme expenditures will be dedicated to evaluation, studies and research;

7) Evaluation concerns must be addressed at the design stage of any intervention, with adequate resources set aside;

8) Evaluation must emphasise the analysis of effectiveness and results;

9) Evaluation must be credible by meeting professional quality standards;

10) Evaluation must be conducted in an independent, impartial and transparent manner;

11) Evaluators must have skills in evaluation, together with personal and professional ethics and integrity, and basic skills in human rights and gender equality analysis;

12) Evaluation findings, recommendations and lessons are made public and disseminated to all stakeholders concerned;

13) Evaluation must be duly considered, with management responses and action plans made public;

14) At minimum, the executive summary of the evaluation must be translated into relevant local language(s) and made available to stakeholders;

15) Evaluation findings of relevance to the Executive Board are to be brought to its attention.

22. The report also covers the six main areas of strengthening the evaluation function covered under the MTSP (2006-2013):

1) National capacity-building and strengthened national leadership in country-level evaluations;

2) Strengthened evaluation within the United Nations system and with other partners;

3) Evaluation in humanitarian crises;

4) Evaluations related to MTSP focus areas, strategies and operational effectiveness;

5) Strengthened organisational capacity in evaluation;

6) Heightened management attention to the evaluation function.

23. Finally, the synthesis report includes some elements that may appear in other Evaluation Policies of other organisations but are not currently covered in the UNICEF Evaluation Policy.
3. Findings on the Evaluation Policy and Function

3.1 Findings on the 15 Evaluation Policy Guidelines

Guideline 1: Evaluation must respect the principle of universality and country-driven programming.

Finding 1: UNICEF’s decentralised evaluation function has its grounding in the country-driven programming of the organisation.

24. The recent focus on country-driven programming lies within the Paris Declaration on Aid Effectiveness and its successor declarations. The recent QCPR process has also taken on country-driven programming as one of the priorities for the United Nations as a whole. While UNICEF uses the term country-driven programming in its Evaluation Policy, the evaluation policies of UNDP and UN Women highlight the term “national ownership.” For example, UNDP has defined national ownership as follows within their revised 2011 Evaluation Policy:

National Ownership: Evaluation should be guided by national priorities and concerns and should be conducted in alignment with national systems. It should be inclusive and take into account diverse national interests and values. Evaluation should strengthen partnerships with governments and key stakeholders. It should build the capacity of national institutions to implement, monitor and evaluate.9

25. There is a strong consensus in the 2012 Case Studies that UNICEF’s engagement at the decentralised level in programming and evaluation is a defining characteristic within the agency. A number of key informants in the 2012 Case Study Report suggest that use of evaluations is greatest when they are closely linked to country level decision making or evidence based policy advocacy in support of country level partners. Country-driven programming implies a certain level of responsiveness to the needs of national governments and the utility of evaluation results, which was also highlighted in the 2012 Case Study Report. It can also be tied to UNICEF’s support of country-led evaluations, which is addressed in subsequent findings in this report.

26. Across all countries and regions and within the various types of country case studies presented, a strong case was made in the 2012 Case Study Report for a continuation of the evaluation function at a decentralised level. This was seen as the best way to respond to two complimentary needs: 1) to meet the demand and use requirements for the different types of evaluations, including for accountability, decision making, and evidence-based policy advocacy purposes, and 2) to meet the capacity development and advocacy role at the country level in favour of evaluation, including the support of nationally-led evaluations as well as building a national demand for evaluations in complex political contexts. In both cases, demand and use of evaluations are best served by a strong decentralised evaluation function.

Finding 2: UNICEF evaluations are seen as relevant across all levels of the organisation, although knowledge generation in general is seen as weak.

27. In decision 2002/9, the Executive Board reiterated, in the context of the evaluation function, the principles of universality and country-driven programming for the activities of UNICEF. The Board emphasised the importance of preserving the decentralised nature of the evaluation system in UNICEF.

---

given the diversity of situations that exist in different countries, and recognised that it is necessary to take into consideration this diversity while drawing lessons at the global level for policy development from evaluation exercises. One interpretation of the term “universality” is that it refers to drawing lessons at the country level which can be applied globally, and thereby contributes to knowledge generation. In other words, evaluations should have universal relevance within the agency, and be useful in a broader regional or global sense. Another interpretation is the link with the universality of UNICEF’s mandate and therefore the need for an evaluation function to be developed accordingly: any review of the evaluation policy should consider the role of UNICEF evaluation function beyond evaluation of traditional UNICEF projects and programmes especially in countries where funding is no longer the main commodity. Similarly, regional and global evaluations should have universal relevance and applicability at the country level. This definition of universality has some resonance across other development organisations. UNEG Norms for Evaluation in the UN system refers to “universality and neutrality in its preamble. In a review of other Evaluation Policies, most explicitly refer to “knowledge generation” or “knowledge management” within their policies. UNICEF is the only agency which highlights universality specifically in its Evaluation Policy. The UNICEF Evaluation Policy does not provide additional guidance on the term “universality” and the reference to the Executive Board decision 2002/9 is not very instructive in guiding Evaluation Specialists in understanding this principle, particularly where it may be confused with the human rights principle of universality.

28. Overall, the guideline on universality, taken together with country-driven programming suggests that the two are mutually reinforcing. Lessons learned at the country level may have application at regional and national levels, and UNICEF is often seen as well-positioned to generate and share lessons across regions and through global fora. At the same time, it is able to bring global lessons and best practices down to the country level through its programming revisions and reviews. Evaluation, as one Case Study report respondent put it, “…helps to validate the operational strategy in key sectors and allow programme staff to make important revisions and adjustments.” UNICEF programme staff, by acting as an extended team of technical experts from COs, ROs and HQ, are able to synergise findings and recommendations of evaluations and make policy and programmatic improvements and develop new guidance where needed.

29. Internal communication and sharing of evaluation findings within the CO, at higher levels within the agency and in the wider development community was seen less favourably by some respondents of the Case Study Report. These respondents note that research, studies and evaluations often get requested despite the existence of relevant information from previous evaluations. Others suggest that the large number of evaluations and assessments conducted makes it particularly challenging for senior management to stay abreast of all evaluation findings and keep track of the most informative evaluation products. At the country level, Country Management Teams (CMTs) or Programme Management Team meetings often do not fully disseminate evaluation findings and recommendations to the entire office for learning purposes, with the exception of mid-term and final country programme evaluations, which are generally disseminated and used adequately within COs for programme improvement purposes.

30. Regional Management Team (RMT) meetings, regional programme meetings, DROPS meetings technical consultations and webinars were all cited as ways in which evaluation results are communicated and disseminated. However, none seems to be fully satisfactory in generating knowledge and learning within the organisation in the eyes of most respondents. Many interviewees of the Case Study report did not cite evaluation sections within COARs, or the summary products that emerge from them, as ways in

---

which good practices and lessons learned on evaluations within the organisation are shared and disseminated.

31. The 2011 GEROS report underscores a challenge related to the correct identification of useful lessons learned. It would appear that this is still largely due to misunderstanding what lessons learned are. The 2011 GEROS report suggests that as a knowledge-centred organisation, *it has to be of concern to UNICEF that only 15 per cent of reports adequately demonstrate an understanding of what the purpose and value of lessons learned is. Lesson learned are mostly misinterpreted as object-specific management changes.*

**Guideline 2: UNICEF is to support programme countries in evaluating their own programmes and to contribute to the strengthening of evaluation capacity in these countries.**

**Finding 3:** UNICEF M&E staff at COs, ROs and the EO have integrated national capacity development into their day-to-day work in response to this noted priority.

32. UNICEF has enshrined capacity development in national evaluation systems into its evaluation function. Not only is it a guideline in the Evaluation Policy itself, but it was also highlighted as a priority for the evaluation function in the MTSP (2006-2013).

33. It has also been highlighted as a critical role for the United Nations and development agencies more generally. Several examples can be taken from other development partners: WFP’s Evaluation Policy provides a definition of country-led evaluation. DFID’s policy highlights country-led/owned evaluation and building a “culture of learning.” A 2010 UNDP Review of its Evaluation Policy suggests that a niche role for UNDP may be in supporting regional evaluation networks. UNFPA’s Executive Board, in a decision in its 2010 Annual Session, underlined the need to enhance the ownership and leadership of programme countries in evaluation activities, and, in this regard, requested UNFPA, to collaborate with other United Nations institutions to optimise the capacity-building of implementing partners. The QCPR Review noted that country-level evaluations require functioning evaluation units in government and evaluation capacity in civil society organisations. This was also recognised by the 2007 TCPR. Building national capacity becomes increasingly important as governments pursue country-led evaluations or joint evaluations. The Joint Inspection Unit (JIU) raised the issue of national evaluation capacity in a 2008 report, stressing that monitoring and evaluation of nationally executed projects in a number of instances were not government-led. Therefore, in the most recent QCPR Recommendations, the Secretary General,

> “Called upon members of the United Nations development system to support strengthened national ownership and leadership of evaluation of operational activities for development at the country level. In this regard, the General Assembly may wish to request the United Nations entities to intensify efforts to build national capacity in programme countries for evaluation of operational activities for development. Also, in this regard, the General Assembly may wish to request UNEG and the UNDG to develop guidelines for further strengthening of national evaluation capacities for operational activities for development of the United Nations system;”

---

11 Also one of six areas for strengthening in the MTSP: National capacity-building and strengthened national leadership in country-level evaluations.


13 Quadrennial comprehensive policy review of operational activities for development of the United Nations system.
34. UNICEF’s own review of national M&E capacity, as part of the 2011 Survey, suggests that only 23 per cent of UNICEF staff were satisfied (“largely satisfied” or “somewhat satisfied”) with national M&E systems in their respective countries. Most respondents (58 per cent) believed that the capacity of national M&E systems were either largely unsatisfactory (23 per cent) or somewhat unsatisfactory (35 per cent). Significant scope for improvement exists, which provides some justification for the attention that UNICEF is giving to this area. By country typology, respondents expressed the greatest dissatisfaction with national M&E capacity in large COs, low income (LI) and lower middle income (LMI) countries.

35. Given a list of key interventions, UNICEF respondents in the 2011 M&E Survey Report identified the top three priorities of their COs in developing capacity of national M&E systems in their respective countries:

- Supporting implementation and strengthening of M&E national units (53 per cent of respondents);
- Strengthening policy-makers’ demand for M&E systems and data (41 per cent of respondents); and
- Supporting data collection and dissemination (35 per cent of respondents).

36. UNICEF staff are justifiably proud of the leadership role within the UN that they play in building capacity in evaluation at the country level. At the CO level, both the 2011 M&E Survey Report and 2012 Case Study Report highlighted that UNICEF staff at the country-level see the organisation as a leader in supporting country-led evaluation and developing national capacity for evaluation. At the RO level, a number of Regional Evaluation specialists have prioritised national capacity development in M&E, putting significant time and resources into regional initiatives and helping to inculcate a culture of evaluation in the region.

37. In 2010, UNICEF EO developed guidance materials on national capacity development for country-led evaluation systems. After one year, 71 per cent of M&E staff were familiar with these guidance materials, according to the 2011 M&E Survey Report. This high level of knowledge of these guidance materials after two years reflects positively on UNICEF EO leadership in this area.

38. Evidence of integration of this guideline within UNICEF’s programming at the CO level has been collected through the COARs. The 2011 Qualitative Review grouped activities into three categories:

1) Activities related to enabling the national M&E environment
2) Activities related to institutional M&E support
3) M&E support provided at individual level

39. In total, more than 400 different national monitoring and evaluation capacity development (NMECD) interventions were carried by 120 COs in 2011. Most of NMECD interventions were at institutional level, and the COAR reports highlight different examples of the types of activities and some best practices.

Finding 4: While UNICEF efforts in national capacity development have improved the culture for evaluation at the country level, some challenges have been noted in how UNICEF prioritises national capacity development (e.g., in terms of providing adequate evaluation expertise, balancing capacity development with monitoring evaluation quality, and the RO role).

40. Key informant interviews and the 2012 Case Studies helped to bring more clarity to the issues of capacity development of national M&E systems. The conclusions of this report were threefold, as discussed below.

41. Efforts made in national capacity development have shown results at the CO level. Small UNICEF COs in middle income (MI) countries and large COs in low income (LI) countries both play an important role in advocating for the importance of evaluations with national partners. In LI countries, this is done through an engagement with specific government ministries and by backstopping local evaluation societies.
and associations. These efforts create a stronger evaluative environment which generates a rise in demand for evaluations in these countries. COs that do not invest in national evaluation capacity strengthening have found it harder to engage with national partners on issues of evaluation. In some case study countries, capacity development was therefore seen as an essential part of an M&E specialist’s role.

42. The second conclusion underscores a unique challenge for UNICEF in MI countries to meet the technical rigor for evaluation required in middle income countries. In some MI countries, donors and government have increased their demand and interest for nationally-led evaluations. These MI governments are increasingly engaging in evaluations at the country level, both to demonstrate the value of UNICEF’s interventions and also to justify the government’s own financial investment in such interventions. Respondents, however, note that the technical rigor of UNICEF staff in advanced evaluation techniques is sometimes inadequate for the emerging needs in MI countries. For example, certain MI countries require more advanced expertise in capacity development in order to support government with policy evaluations, impact or equity evaluations. At the same time, the 2011 M&E Survey Report suggests that compared to their peers in LI countries, the M&E staff in MI countries tend to have less experience. As such, many respondents note that while UNICEF remains a relevant actor in evaluations at the CO level, it is not due to its expertise, but rather due to its field presence and credibility with government partners. This merits a re-examination about whether UNICEF can engage the right type of staff members in the right place. If not, it must think about how it can leverage partnerships with other development organisations or research institutes to meet the evaluation demands of MI countries.

43. A final conclusion of the 2012 Case Study Report is the fact that capacity development of national M&E evaluations is not highly incentivised by UNICEF in performance reviews or at a management level. Close monitoring of the internal quality dimension of evaluations and management responses through GEROS and the VISION dashboard detracts attention from the capacity development work carried out by many M&E staff at the CO and RO levels. The 2011 M&E Survey Report clearly highlighted that M&E staff are responsible for a range of tasks at the CO level, greatly limiting the amount of time they can dedicate to the evaluation function. In some instances, the GEROS ratings and VISION dashboard raised questions as to the role of the M&E specialist in the country: is it to build capacity of partners, or to focus on leading and managing high quality evaluations? In other words, the rating systems in place for evaluation oversight shifted priorities away from capacity development in some countries and regions. Respondents felt that the rating system on evaluation quality implies a relatively high importance of the evaluation function compared to capacity development, as well as other tasks, such as performance or situational monitoring.

Guideline 3: Whenever possible, evaluations must be undertaken in partnership with national authorities, with the United Nations system and with interested partners.14

Finding 5: UNICEF engages in evaluations in partnership with national governments, UN agencies, and other partners to some extent, but this guideline is not well-defined and the results of these efforts are not consistent.

44. A cursory review of the Evaluation Policies of development organisations brings up the word “partnership” in several; IFAD, DFID and UNDP mention partnership in one way or another, while UN Women speaks of “UN System Coordination”. UNICEF’s own MTSP reinforces this guideline with a view to further strengthen UN system coordination in evaluation and greater collaboration with other

14 This guideline is consistent with the MTSP focus on “Strengthened evaluation within the United Nations system and with other partners”
partners. The QCPR Report noted weakness in joint-evaluations with other UN agencies, specifically with regard to monitoring and evaluation of the UNDAF. It requested the UN agencies, including UNEG, to “…intensify the use of UNDAF and system-wide evaluations in lieu of agency-specific evaluations of operational activities for development, wherever possible, particularly in programme countries adopting the delivering-as-one approach.”

45. UNICEF’s strong presence at the decentralised level and its engagement with national governments often leads to the assumption that a priori, UNICEF is a partner with government on all evaluations. The evidence suggests that this may not always be the case: UNICEF’s engagement with government on the Annual IMEP, stated as 55 per cent of the time in the 2011 M&E Survey Report, suggests that this assumption should be treated with some degree of caution. UNICEF carries out a number of mandated evaluations, such as country programme evaluations, mid-term evaluations and reviews. In general, UNICEF staff insist that they do partner with government on most evaluations. The level of partnership with government in these evaluations, for example, may vary from full engagement to very little engagement. At the same time, there seems to be a lack of clarity in what is meant by partnership with national authorities, and whether it implies government approval of the IMEP, shared costs with government, participation in steering committees, or something else.

46. Similarly, the level of commitment to joint evaluations at UNICEF appears to be mixed. On one hand, data suggests that it is committed to UN system coordination. According to the 2011 M&E Survey Report, most respondents (87 per cent) believed that UNICEF currently invests a level of effort in the UNCT/UNDAF monitoring and evaluation work that is either about right (52 per cent) or too much (35 per cent). Only a few respondents (4 per cent) think that too little effort is invested for that purpose. In the case study interviews, UNICEF see themselves as leaders in harmonised efforts in favour of M&E at the country level.

47. However, while some of the case study countries reflected the importance of conducting joint evaluations, this was not consistently expressed across all countries. Small COs and those in MI countries appear to be more likely to engage in joint evaluations. Two reasons are cited for this: the first is that there is a greater imperative to use limited resources efficiently; the second is that MI countries are requesting increased joint evaluations because several organisations may contribute to a policy or intervention. The aid effectiveness agenda (i.e., the 2005 Paris Declaration on Aid Effectiveness, the 2008 High Level Forum in Accra, the 2011 High Level Forum in Busan, and the OECD’s work on aid effectiveness) is another factor which contributes to an increased demand for joint evaluations in some countries. Interviewees have expressed that greater clarity needs to be provided linking UNICEF’s Evaluation Policy to the aid effectiveness agenda.

48. For those detractors of joint evaluations in the Case Study Report, several concerns were raised: joint evaluations make it harder to assign contribution and more difficult to respond to donor requests for accountability. Moreover, one senior manager pointed out that in order to have joint evaluations, there needs to be joint programming, and this is not often the case at the CO level. The Annual Report of the Evaluation function and major evaluations in UNICEF confirms the lack of engagement in joint-evaluations: The number of evaluations co-managed with other United Nations agencies increased from 1 per cent in 2009 to 6 per cent in 2010.

49. According to GEROS, of the evaluations whose management arrangements could be determined from reports, those conducted jointly with other development partners produced the highest number (three) of satisfactory reports, but almost the highest percentage (over 70 per cent) of not confident ratings. Jointly managed evaluations with UN agencies and country-led evaluations were both small in number (five each), but produced proportionately higher rates of satisfactory reports (both 20 per cent) and lower rates of not confident ratings (20 per cent and 40 per cent respectively).
Guideline 4: Evaluations at all levels (whether of strategic governance, global, regional or country programmes, or projects) must serve an explicit management purpose.

Finding 6: Evaluations are most often used for evidence-based policy advocacy and decision-making purposes.

50. UNICEF’s Evaluation Policy mirrors that of most other development agencies with regard to the focus of evaluation on serving an explicit management purpose. UNEG Standards, for example, refer to ensuring that: “…evaluation is part of the organisation’s governance and management functions,” and that “Evaluation makes an essential contribution to managing for results.” This language is pervasive across the development agencies. UNDP’s Evaluation Policy, for example, “…seeks to increase transparency, coherence and efficiency in generating and using evaluative knowledge for organisational learning and effective management for results, and to support accountability.”

51. UNICEF Evaluation Office has been broadly defining four principle uses of evaluations as accountability, decision making, evidence-based policy advocacy, and knowledge generation. The various studies and surveys in the past years confirm that UNICEF is increasingly conscious of defining the purpose of evaluations at their design stage. However, it is also clear that most senior managers cannot easily define one single “explicit” management purpose. Instead, evaluations often have multiple management purposes, with one being dominant. For example, respondents of the Case Study Report suggest that the key uses of evaluations are for decision making and for developing evidence-based policy advocacy. Knowledge generation was the evaluation purpose referenced the least often. Those who did mention it mostly discussed knowledge generation as a by-product of evaluations rather than as their primary purpose.

52. The 2012 Survey report asked CO Representatives to provide a relative ranking of the four uses of evaluations; its conclusions were consistent with the Case Study Report: that 50 per cent of CO Representatives feel that evidence based policy advocacy is the most important purpose for carrying out evaluations within UNICEF. The second most important purpose is for decision-making (26 per cent) and third is for accountability (17 per cent). Knowledge generation is viewed as the least important of the four purposes for evaluation. As in the case studies, it was noted that all evaluations generate knowledge even if it is not the primary objective, and furthermore, knowledge generation can be carried out through studies, research and surveys as well as through evaluation.

53. At the HQ level, evaluations for decision-making purposes are seen as the barometer which informs technical experts on how well the programme strategies are translated into practice. In other words, they are demanded by senior managers at HQ who rely on an uptake of information from CO evaluations to update global UNICEF programme guidance within a certain programmatic area. This drives UNICEF-wide decisions on high level policies and practices which are then translated back to the CO level in the form of policy and programmatic guidance.

54. Both of the key uses of evaluations in UNICEF imply strong country level engagement - a pillar of the decentralised evaluation function. In other words, use of evaluations is greatest when they are closely linked to country level decision making or evidence-based policy advocacy with country level partners. In difficult political contexts, this use might involve demonstrating value for money and areas of improvement on a politically sensitive programme that a government ministry could not carry out without UNICEF leadership. It may involve evaluating a pilot programme for potential scale-up. Or, it may involve generating the leveraging power to argue for new policies or programmatic direction. The greater the reflection on the use of evaluations at the planning stage, the better the ultimate use of evaluations.
55. Evaluations in UNICEF that respond to the accountability uses of donors, the Executive Board or partners may be more implicitly understood by interviewees. Once produced, these evaluations are shared with donors and government partners and respond to this accountability purpose. However, the broader use of these evaluations at the country level more often includes decision making and evidence-based policy advocacy. Senior managers insist that the key to any evaluation is the link to country-level decision making. Therefore, they see a more limited use of multi-country or global evaluations because these result in findings and recommendations that may be inapplicable for certain countries. Evaluations which are most relevant are those which provide insight that can be used to improve country-level performance.

Guideline 5: It is important to preserve the decentralised nature of the evaluation system in UNICEF.

Finding 7: While UNICEF’s evaluation system is decentralised, there are challenges in ensuring sufficient human and financial resources for this structure to meet expectations.

56. The decentralised focus of evaluation is strongly entrenched in the organisation. It is also present in evaluation policies of several other agencies, including UNDP and DFID, WFP and UN Women, all of which separate, and sometimes define corporate evaluations as compared to decentralised evaluations for the purposes of the Evaluation Policies.

57. For example, UNDP has put increased emphasis on the decentralised evaluation function in recent years. Its 2010 Executive Board Report “note(d) with concern the continuing and serious challenge of the coverage, compliance, quality and use of decentralised evaluations... and request(ed) UNDP, as a matter of priority, to strengthen decentralised evaluation capacity and increase its use in line with the independent review of the UNDP evaluation policy and the management response.” In addition, the Executive Board made note of the decline in the total number of decentralised evaluations, and requested UNDP to strengthen decentralised evaluation capacity to monitor and evaluate programmes at the country level and to increase the use of these evaluations as the basis for decision-making for future improvements, “taking into account the need to support regions and focus areas with lower evaluation coverage and higher resource levels.” In so doing, UNDP has set out both a strong rationale and justification for an increased focus on the decentralised evaluation function.

58. Within UNICEF, approximately 95 per cent of evaluations supported by UNICEF are managed at the decentralised level, according to the Qualitative Review report. Case Study respondents agree that the key objective of the decentralised evaluation function is to produce quality evaluations which are relevant at the country level. A strong decentralised function is also the best way to provide strong capacity development of national partners. M&E staff play a key role in quality assurance at all phases in the evaluation process; from planning, to crafting TORs to working with ROs and programme staff in ensuring a robust methodology for the evaluation. M&E staff also contribute to certifying the use of evaluations, including dissemination and follow-up on the management response. With regard to funding for evaluations, senior managers point to a continued need to ensure adequate resources are available to produce these quality evaluations.

59. Case studies and key informant interviews offer different perspectives on the necessary human and financial resources based on the size of the CO and the level of development of the country or region. Larger COs confirm that they are more likely to have experienced M&E staff and the required funding to

---

ensure quality evaluations. The challenges they face are those of building national capacity in evaluation, managing several concurrent evaluations and retaining high quality consultants.

60. In contrast, medium and small COs have fewer resources and are less likely to have experienced M&E staff available to ensure quality evaluations, and while the dedication with which they support decentralised evaluation is in line with the Evaluation Policy, the reality is that fewer human and financial resources in some regions will mean carrying out more joint and multi-country evaluations.

Guideline 6: An amount totalling 3 to 5 per cent of programme expenditures will be dedicated to evaluation, studies and research.

Finding 8: Although UNICEF is meeting its minimum expenditure target of 3 per cent for evaluation, studies and research, its expenditures on evaluation alone are significantly below one per cent.

61. Most evaluation policies contain some mention of budgets for evaluations. Some specify the types or sizes of programmes which require evaluation. For example, all programmes that receive regular budget resources in the FAO require evaluations. In UN Women’s case, summative evaluations are mandatory for budgets over $1 million while summative and formative evaluations are required for budgets over $3 million.

62. UNICEF appears to be one of the few organisations to attempt to target a level of its expenditure for evaluation, studies and other research within its Evaluation Policy. It currently sets that bar at between 3 and 5 per cent of programme expenditures. Its performance since 2006 suggests that it has not been able to meet this target consistently; from 2006 to 2010, expenditures on evaluation, studies and research have oscillated between 2.84 per cent and 3.1 per cent percentage of programme expenditure. In 2010, the latest year where comparable data was available, the total amount spent on evaluation, studies and research was roughly $99 million, of which $11 million was on evaluation only, compared to an overall programme budget of $3.354 billion. Year-on-year change was flat relative to overall programme spending.

63. Focusing on the amount of spending on evaluations only, citing the Lusthaus and Bester report on System Wide Evaluation, the recent QCPR report stated that “…very few entities in the system have evaluation expenditures that exceed 1 per cent of their total expenditures. The financial and human resources for most evaluation units in the United Nations system are limited to their day-to-day work and participation in system-wide, joint or multi-agency evaluations often means working without extra resources.” The amount spent by UNICEF on evaluation only in 2010 is only 0.33 per cent, well below the 1 per cent bar implied by the System Wide Evaluation Report.

---


Finding 9: The demand for and quality of evaluations at the CO level are affected by the availability of financial resources, the leadership of the Representative, and donor requests. The Evaluation Policy also ensures evaluations are carried out for all programmes at least once per cycle.

64. The 2012 Survey of Representatives highlighted the key factors driving demand for evaluations at the CO level are financial resources, the leadership of the Representative, donor requests, and mandatory requirements.

65. According to the Annual Report of the EO to the Executive Board, “In spite of UNICEF’s commendable performance in evaluation quality, country offices are concerned about the slow growth of the budget line for evaluation.” One of the key objectives of the 2012 survey was to better ascertain the different factors that help dictate demand for evaluations. CO Representatives were asked to rate on a scale from 1 to 5 (with 1 being “low importance” and 5 being “high importance”) each of a number of pre-identified factors which were seen to drive the demand for evaluations at the CO level. These include both internal factors (i.e., leadership of the Representative, Regional Office or Evaluation Office), and external factors (i.e., donor requests, political context). The intention was to elucidate any trends that may emerge overall, and within the various categories of analysis.

66. Overall, the survey results suggest that four key factors drive demand for evaluations: financial resources (32 per cent of respondents rated “high importance”), leadership of the Representative (30 per cent), donor requests (29 per cent) and mandatory requirements (26 per cent). Exhibit 3.1 presents the results in full where factors seen by COs as “high importance” are sorted.

Exhibit 3.1 Percentage of respondents who rated at “high importance” the following factors which drive demand for evaluation at the CO level (N=52)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial resources in your country office for evaluation</td>
<td>32%</td>
</tr>
<tr>
<td>Leadership of Representative</td>
<td>30%</td>
</tr>
<tr>
<td>Donor requests</td>
<td>29%</td>
</tr>
<tr>
<td>Mandatory requirements</td>
<td>26%</td>
</tr>
<tr>
<td>Evaluation capacity (HR) of the country office</td>
<td>19%</td>
</tr>
<tr>
<td>Political context</td>
<td>17%</td>
</tr>
<tr>
<td>Government requests</td>
<td>17%</td>
</tr>
<tr>
<td>Leadership of Evaluation Office</td>
<td>8%</td>
</tr>
<tr>
<td>Leadership of Regional Office</td>
<td>6%</td>
</tr>
<tr>
<td>Number of evaluations carried out in your country office in the previous years (tradition of evaluation)</td>
<td>0%</td>
</tr>
</tbody>
</table>

---

67. Financial resources for evaluation in the country office are viewed as the most important factor affecting demand for evaluations. Lack of financial resources has several implications for the demand to carry out evaluations. For example, it may mean fewer evaluations are planned in the IMEP or that the CO limits resources for evaluations in favour of other uses. It may cause COs to seek more opportunities to carry out joint or multi-country evaluations. When broken down by CO size, the data shows, a greater percentage (42 per cent) of respondents from small COs felt that financial resources was a factor that decreased their demand for evaluations, confirming that financial resources are more likely to be a negative factor on demand in small COs. A similar trend is noted by country typology; UMI country respondents are more likely than other typologies to note that financial resources decrease demand for evaluations.

68. At RO and EO levels, the ability of these offices to provide adequate guidance and support to the decentralised evaluation function was limited by resource levels. All ROs have implemented quality assurance mechanisms to assure a certain level of quality across evaluations in the region, and some have engaged external consultants to provide technical support in the drafting of TORs or in reviewing draft reports. Some ROs responded by encouraging more strategic evaluations in their regions in order to provide adequate oversight to all evaluations in the region. However, some case study informants expressed confusion as to how fewer evaluations would conform to the Evaluation Policy, which states that the CO Representative is accountable to “institute measures to ensure that evaluations are strategically selected and that all major programme components are evaluated during the programme cycle, allocating adequate resources.” This part of the Policy is unclear to many, who feel that these must be distinct evaluations, rather than one evaluation that covers multiple programmes (such as a policy evaluation, or a country programme evaluation).

69. An example of insufficient resources at the EO level is the demand to evaluate Level 3 emergencies, which currently outweighs the resources available to properly support COs and ROs to do so, for both for accountability and for learning and decision making purposes. This demand is particularly notable at the level of EO, which takes a much more active role in evaluating Level 3 responses as compared to Levels 1 or 2.

Finding 10: COs that have not carried out evaluations in the past few years have cited lack of resources and lack of country-level demand for evaluation as reasons.

70. Several COs have not conducted evaluation during the last three years for reasons which are often country-specific and which tend to focus on a lack of external demand. Occasionally, lack of resources (e.g., funding, in-house evaluation specialists or external evaluation consultants) have prevented COs from conducting formal evaluations. One CO suggested that government was not requesting any evaluations, and since it relies heavily on in-country resources, it must also respond to the needs of the country. This comment underlines two points: first is UNICEF’s role of capacity development of national evaluation systems, as highlighted in Finding 4. The second is the reality that UNICEF’s role in MI countries will increasingly be to encourage the government to carry out their own evaluations and providing timely technical assistance to that end. This would be a government led evaluation, and therefore would not be included in the GEROS database. It is unclear whether this meets the criteria set out in the Evaluation Policy.

71. Irrespective of country size and income level, COs that have not carried out evaluations in the last three years also appear to be using other types of assessments to generate knowledge, particularly for evidenced based policy advocacy. The most frequently cited assessments that fall outside of the definition of evaluations are reviews, surveys, research studies and situation analyses. Other forms of assessments include needs assessments, appraisals, evaluability assessments, inspections, investigations, audits and surveys. These other forms of assessment are defined in the UNICEF PPPM.
use of evaluations may come at the expense of the quality of other assessments that are not subject to such rigorous quality control mechanisms. This is further reinforced by the fact that products from other assessments are neither uploaded nor taken into account in the GEROS process. Other respondents simply had not sufficiently considered the pros and cons of each type of assessment method and the specific role of an evaluation as opposed to other forms of assessment.

Guideline 7: Evaluation concerns must be addressed at the design stage of any intervention, with adequate resources set aside.

Finding 11: UNICEF COs with highly rated evaluations are characterised by strong planning and Terms of Reference, good national evaluation capacity, and an adherence to a quality assurance process that often includes RO involvement.

72. As expected, good planning is a fundamental building block of any evaluation, and also a cornerstone of any Evaluation Policy. Coupled with planning is evaluability, which is explicitly highlighted by the Evaluation Policies of UNODC, USAID and WFP. In 2009, UNDP’s Executive Board emphasised the need for UNDP to improve programme planning and results-based management in order to improve the design and methodology of decentralised evaluations. In particular, the Executive Board commented on the evaluability of UNDP programmes as a prerequisite for effective programme design, continuous monitoring with national partners and evaluation. In the UNICEF survey of Representatives, the evaluability of projects/programmes was also cited as one of the factor affecting the quality of evaluations.

73. The Case Study Report outlined UNICEF staff views on the prerequisites for high quality evaluations and found a great deal of homogeneity across regions, country typologies and levels: a high degree of planning, a well-defined review process, and strong evaluation capacity at the national level, including independent consultants with the right language skills. Linked to planning is the requirement for a clearly defined TOR, another commonly cited factor by respondents at all levels. The absence of any of these factors generally has a detrimental effect on evaluation quality.

74. Almost all respondents pointed to the need for a high level of planning for evaluations, taking into consideration the ultimate use at the design and planning stage so that government buy-in could be secured, as well as buy-in and support from the Country Management Team, RO, and other partners. Most respondents agreed that insufficient attention was allocated to the proper planning of evaluations, that evaluations were not properly integrated into IMEPs, or that referencing in IMEPs was sufficient to ensure proper evaluation planning.

75. More broadly, each of these elements point to the need for a clear quality control processes (e.g., planning as part of IMEP, advisory groups, feedback processes, application of norms and standards) which should exist within COs. The process should be carried out by M&E staff and programme staff, be enforced by senior management and be guided by the RO M&E advisors, programme staff and management. Existence and adherence to an agreed workflow is another important factor influencing the quality of evaluations.
76. COs with highly rated evaluations generally operate in countries with strong national evaluation capacity and an openness to carrying out evaluation. They also have access to a wide selection of experienced evaluation consultants and are less hindered by challenges in finding experienced consultants with the right language skills. They often have developed good monitoring mechanisms and have access to baseline data. Finally, good quality evaluations are generally those that take advantage of the RO expertise for the design and planning of good evaluation TORs, as well as during methodological discussions and during the review of draft evaluation reports.

**Finding 12:** Ensuring adequate human and financial resources for monitoring and evaluation at the country level is necessary to strengthen the decentralised evaluation function; however, this has different implications across regions and COs.

77. All respondents agree that the key objective of the decentralised evaluation function is to produce quality evaluations that are relevant at the country level. A strong decentralised function is also the best way to provide strong capacity development of national partners.

78. M&E staff play a key role in quality assurance at all phases in the evaluation process; from planning, to drafting TORs to working with ROs and programme staff in ensuring a robust methodology for the evaluation. M&E staff also contribute to ensuring the use of evaluations, including dissemination and follow-up on the management response. With regard to funding for evaluations, senior managers point to a continued need to ensure adequate resources are available to produce these quality evaluations.

79. Through an open question, CO Representatives were offered the opportunity to provide recommendations on ways to improve the quality of evaluations. Responses were grouped under relevant headings for analysis purposes. The most common responses to the open question were as follows: to have mandatory regional mechanisms to review TORs and draft reports (12 per cent), hiring good quality consultants (12 per cent), and providing systematic training of key programme and M&E staff (12 per cent). Other frequent responses included setting mandatory funding aside for evaluations (11 per cent) and providing clear guidance on evaluation standards (11 per cent).

---

20 It should be noted that such a mechanism exists in all ROs to some extent. This response may suggest that not all CO Representatives are familiar with all the mechanisms available to them through the RO.
80. The 2011 M&E Survey Report highlighted the overstretched nature of many M&E staff on a range of tasks, including planning and performance monitoring, situation monitoring, research and evaluation. When asked to calculate how their time is apportioned, evaluation received the lowest percentage, at roughly 14 per cent, compared with 29 per cent for situation monitoring, 24 per cent for planning and performance monitoring, 17 per cent for research and 16 per cent for “other”. While the majority of M&E staff expressed satisfaction with the availability of human resources for their function, a significant minority (30 per cent) communicated dissatisfaction, especially on evaluation tasks, indicating a sense of overstretch.

Guideline 8: Evaluation must emphasise the analysis of effectiveness and results.

Finding 13: Although the Evaluation Policy does not set out clear definitions of results, the GEROS review accounts for output, outcome and impact evaluations.

81. Evaluation is generally accepted within the United Nations as an essential tool for demonstrating the effectiveness of the United Nations development system. With the greater focus on results, there is also an increasing demand for evaluative evidence. The UNEG Standards state simply that any institutional framework should support “…the evaluation function's key role in contributing to the effectiveness of the Organisation,” a view shared across the Evaluation Policies of most development agencies.

82. Many Evaluation Policies provide definitions of results, including differences between outputs, outcomes and impact. This allows for some development organisations to be more specific on “effectiveness,” which usually refers to results at the outcome level\(^{21}\). The UNDP Executive Board in 2009 noted UNDP’s decline in country programme compliance with undertaking outcome evaluations, and requested UNDP to improve compliance through the establishment and appropriate resourcing of achievable country programme evaluation plans, including the resourcing of decentralised outcome evaluations from programme resources.

83. According to the 2011 GEROS Report, there was fairly even distribution of evaluation reports across the levels of result—output, outcome, and impact. Furthermore, there was a very clear trend in ratings: quality of output evaluations by UNICEF was by far the least satisfactory (less than 20 per cent.

---

\(^{21}\) These practical definitions of output and outcome are not shared in the UNICEF Evaluation Policy, but rather clarified within the PPP Manual. While UNICEF’s use of language such as programme component result (PCR) and Intermediate Result (IR) in recent years has been cited as improving a focus on results, it has also hindered cross-agency harmonization of language and terminology.
Finding 14: The overall demand for evaluation to serve accountability purposes has been increasing in recent years, due in part to an increased focus on results by donors.

84. Across Country Office types and sizes, and in all regions, there is a consensus of a continued demand for UNICEF to participate in the development of quality evaluations for accountability purposes. Most COs and senior management staff at the HQ level are driven to carry out evaluations in order to provide direct accountability for project/programme funds to the Executive Board, to donors and to partners.

85. The UNICEF Executive Board has embedded the requirement to report back periodically on UNICEF’s field level application of its programmes and strategies across the organisation, as well as how policy and pilot interventions are being implemented. UNICEF has well-established policies in place to encourage transparent reporting of its results with government partners and donors at the country level. In all cases and at all levels, there is a strong importance placed on demonstrating continued relevance, effectiveness, and efficiency of programmes and interventions.

86. Donor accountability is noted most strongly in LI countries and large COs, where the range of donor funds is greater and UNICEF’s involvement in piloting of programmes or new initiatives require evaluative techniques to demonstrate their effectiveness. Large COs and LI countries receive significantly higher donor resources, with attendant demands to ensure appropriate and transparent use of funds. With regard to evaluations of UNICEF’s emergency response, accountability is greatest for Level 3 emergencies for the same reason; these emergencies receive significant donor resources for which UNICEF emergency programming is held accountable.

87. Pressure for greater donor accountability also exists in MI countries. In these cases, it is often one or two key donors who make requests for evaluations in response to specific programme funding. This is coupled with increasing pressure from some MI governments for evaluations to report on the results of UNICEF supported government interventions and to provide justification for the types of interventions carried out with UNICEF support. In some cases, these MI governments have also contributed their own funds to UNICEF-backed initiatives, and have their own interests in reviewing evaluative results.

88. Accountability from donors is also strongly felt at HQ levels. Senior managers are aware that in an extremely competitive and dynamic global funding environment, donors are placing a greater onus on UNICEF to demonstrate the impact and relevance of its programming choices, particularly across the key focus areas of health, education, HIV-AIDS, child protection as well as emergency programming. It also reflects a global trend of wider and faster dissemination of information, resulting in an increased pressure to provide transparent and timely information.

Guideline 9: Evaluation must be credible by meeting professional quality standards.

Finding 15: UNICEF’s recent focus on evaluation quality through GEROS is an incentive for COs to improve evaluation quality.

89. Quality is paramount in evaluations. UNICEF’s current Evaluation Policy contains several guidelines which contribute to producing quality and credible evaluations, both in terms of process and in reporting. Like UNICEF, most other agencies have also integrated credibility into their Evaluation Policies. The rationale for doing so is clear; as the 2012 Survey Report indicates, quality is the most important determinant of evaluation use. Fully 63 per cent of respondents of the survey rated the quality of evaluation
as being the most important factor affecting internal use of evaluations, and this figure rose to 74 per cent when respondents rated this for external use of evaluations.

90. The UNDP review of its Evaluation Policy\textsuperscript{22} asked the following questions on credibility: “Do EO’s evaluations meet the quality criteria as stipulated in the UNEG Norms and Standards? Do EO evaluations have meaningful and transparent consultation with stakeholders? Are EO evaluations conducted with ethical considerations as expressed in the policy?” These questions highlight the importance of meeting the credibility standards during the process of the carrying out the evaluation, as well as the standards of the evaluation report itself.

91. In most UNICEF COs, the 2012 Case Study Report presented a perception that new initiatives to develop accountability in evaluation quality, including GEROS, provided the correct incentives within the agency to prioritise and improve evaluation quality. Respondents have noted a higher profile given to evaluations, a greater understanding of UNEG guidelines, knowledge of GEROS ratings and sensitivity to the independence of evaluations. The Case Study Report also noted this improvement in evaluations of emergencies in recent years.

92. The importance of the independence of evaluations in particular was raised by a number of stakeholders in the 2012 Case Study Report. This showcases the fact that technical quality criteria for evaluation are now well understood and that processes to improve evaluation quality are paying dividends.

93. At the regional level, evaluation specialists and Regional Directors have noted that their own incentives and accountabilities are being clarified and better defined as a result of the GEROS system and the VISION dashboard. This has led most ROs to develop their own quality assurance systems and guidance tools to engage in and track the evaluations taking place on a country-by-country basis in their respective regions. However, as mentioned in Finding 4, resource constraints to fulfil oversight accountabilities at the RO have put into question just how far ROs should be held accountable for GEROS results in their region.

94. At the senior management level, there is a broad appreciation for the role played by the EO to institute and monitor evaluation quality as a clear indicator that can be tracked. The GEROS results are shared widely at Headquarters and Executive Board levels. The EO has also been lauded for taking steps in improving overall evaluation quality standards for emergency evaluations, such as developing e-learning tools, standard operating procedures and improved guidance for emergency evaluations in the PPPM.

95. Despite an overwhelming approval for the process of defining and measuring quality, there were some concerns raised about the GEROS system and its focus on evaluation quality and the accountability for improving quality. As highlighted in Finding 5, were those who suggested that this tool placed too much emphasis on evaluation quality without also addressing the quality of other research products and studies that might also be carried out at the country level. Others suggested that the GEROS system distracted from capacity development work in the country and contributes to the perception that the evaluation function strictly entails conducting evaluations.

Finding 16: UNICEF’s GEROS system is the primary tool for assessing the quality and credibility of evaluation reports. Results are mixed, with some improvements between 2009 and 2010.

96. The 2011 GEROS report outlines UNICEF’s evaluation report standards, upon which all submitted evaluations are assessed. It should be noted that the GEROS only assesses reporting, and not the evaluations behind them, according to its guidelines, which state that An evaluation report is assessed as satisfactory when it is a credible report that addresses the evaluation purpose and objectives based on evidence, and therefore can be used with confidence.

\textsuperscript{22} UNDP (2010). Independent review of the UNDP Evaluation Policy.
97. Drawing on 89 reports in 2010, the GEROS report noted a slight improvement in performance year-on-year, with 40 per cent of reports being rated overall as satisfactory according to UNICEF Evaluation Standards (36 per cent in 2010). Of the remaining reports, 30 per cent could be brought up to a satisfactory standard with a little more work. Of some concern, however, was the 30 per cent of reports found to have substantive weaknesses: 14 per cent were in the previous year. At the top end of the scale, three reports were rated as outstanding – although eight reports had at least one individual section rated as outstanding.

The GEROS qualitative analysis suggests that “…there is a general trend of reports rated unsatisfactory as being highly descriptive in nature; as compared to better performing reports that used a much more analytical approach. This belies the fact that the differentiating factors between stronger and weaker reports are still fundamental evaluation issues: criteria, methodology, conclusions and so on.”

Guideline 10: Evaluation must be conducted in an independent, impartial and transparent manner.

Finding 17: The overall picture regarding compliance with good practices for evaluations carried out at the decentralised level is mixed.

98. As highlighted in Finding 15 and 16, credibility through good quality reports is a fundamental objective of GEROS. Behind credible and good quality reports is the need for an independent, impartial and transparent process for carrying out evaluations. At UNICEF, the accountabilities outlined in the Evaluation Policy, particularly those around the CO Representative and the RO, help to ensure that the evaluation process follows good practice in independence, impartiality and transparency at all stages of the evaluation, from planning and design to implementation and reporting.

99. The 2011 M&E Survey Report asked M&E staff about compliance at the CO level based on nine standards for evaluations. It should be noted that the survey only applies to evaluations carried out at the CO level. The relatively high number of “sometimes” responses for many of these questions may be an indication of the challenges facing UNICEF in ensuring the consistent application of policies on independence, transparency and impartiality across the organisation. The low level of compliance of M&E staff highlighted in some of these standards is also of particular concern given that this role is seen as a critical function of M&E staff. A partial explanation for this deficiency may be found in Finding 12 in the appropriation of the M&E specialist’s time across numerous priorities. On a more positive note, the low number of “never” responses in standards for evaluation is a positive sign that the institutional framework for evaluation is well-grounded at the CO level.

100. The standards for evaluation in the survey focused on the planning for evaluations (TORs, selection of consultant) and the completion of evaluations (review of draft report, approval, and management response). In particular, respondents were given a list of good practices in the management of evaluations at the decentralised level and had to indicate the frequency of compliance (always, sometimes, never) for each of the good practices. The strongest compliance was found in the following standards:

- Final Terms of Reference approval by Senior Management (66 per cent stated “always”)
- Final Selection of the consultant is approved by Senior Management (65 per cent stated “always”)
- The M&E staff provides technical assistance in development of the TORs (54 per cent stated “always”)

101. On the other hand, low levels of compliance are linked to:

- Involvement of the RO in quality review of draft evaluation reports (19 per cent stated “always”)
• The development of the management response by Senior Management (30 per cent stated “always”)
• M&E staff technical assistance in the selection of the consultant used for an evaluation (30 per cent stated “always”)

102. Another area of concern outlined in the 2011 M&E Survey Report is the lack of independence of some M&E staff. Fully 70 per cent of survey respondents indicated that they report to either the Representative (27 per cent) or the Deputy Representative (43 per cent). However, 30 per cent noted that they report to someone other than senior management, primarily Social Policy or Planning Chiefs.

Exhibit 3.2 Percentage of COs which comply with good practices in their last three evaluations, according to M&E staff (N=86)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Never</th>
<th>Sometimes</th>
<th>Always</th>
<th>Don't Know/ NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>The M&amp;E focal point provides technical assistance in the development of the ToR used for the evaluation</td>
<td>2%</td>
<td>33%</td>
<td>54%</td>
<td>11%</td>
</tr>
<tr>
<td>The draft ToR is sent to the RO for quality review</td>
<td>6%</td>
<td>45%</td>
<td>37%</td>
<td>12%</td>
</tr>
<tr>
<td>Final ToR is approved by the Rep/ Dep Rep or CMT/ Programme Meeting</td>
<td>10%</td>
<td>9%</td>
<td>66%</td>
<td>15%</td>
</tr>
<tr>
<td>The M&amp;E focal point provides technical assistance in the selection of the consultant used for the evaluation</td>
<td>8%</td>
<td>48%</td>
<td>30%</td>
<td>15%</td>
</tr>
<tr>
<td>The final selection of the consultant is approved by the Rep/ Dep Rep or CMT/ Programme Meeting</td>
<td>10%</td>
<td>11%</td>
<td>65%</td>
<td>14%</td>
</tr>
<tr>
<td>The M&amp;E focal point provides technical assistance in ensuring the quality of the draft evaluation report</td>
<td>5%</td>
<td>36%</td>
<td>45%</td>
<td>14%</td>
</tr>
<tr>
<td>The draft evaluation report is sent to the RO for quality review</td>
<td>11%</td>
<td>51%</td>
<td>19%</td>
<td>19%</td>
</tr>
<tr>
<td>The final report is approved by the Rep/ Dep Rep or CMT/ Programme Meeting</td>
<td>13%</td>
<td>24%</td>
<td>42%</td>
<td>21%</td>
</tr>
<tr>
<td>The Representative/Dep Rep leads the development of the management response of the final evaluation</td>
<td>19%</td>
<td>30%</td>
<td>30%</td>
<td>21%</td>
</tr>
</tbody>
</table>

103. The 2011 M&E Survey Report also highlighted some differences across CO size and country typology, implying differences in how M&E staff operate in different contexts: for example, small COs and COs in upper middle income countries tend to engage with ROs to a greater degree than do large COs on the evaluation function.
Guideline 11: Evaluators must have skills in evaluation, together with personal and professional ethics and integrity, and basic skills in human rights and gender equality analysis.

Finding 18: The quality of the evaluator is recognised as one of the key indicators of quality evaluations.

104. This guideline in the Evaluation Policy is another directly contributing to credible evaluations at UNICEF. This guideline covers several elements, including the skills of the evaluator and of the evaluation manager, ethics and skills in human rights, gender and equity analysis.

105. UNEG Standards define evaluation competencies as “…the qualifications, skills, experience and attributes required by those employed within the evaluation function to carry out their duties as stipulated and to ensure the credibility of the process. Competences are required for all those engaged in designing, conducting and managing evaluation activities, managing evaluators, conducting training and capacity development and designing and implementing evaluation methodologies and systems.”

106. UNEG suggests that some skills are particularly useful for persons conducting evaluations as “evaluators”, while others are needed for persons who manage evaluations as “evaluation managers”. UNEG Standards groups both under the term “evaluators.” UNICEF’s Evaluation Policy, on the other hand, separates “staff engaged in evaluation” from “external evaluators employed by UNICEF.”

107. While not assessed directly by UNICEF, the importance of the quality and credibility of the external evaluators was underlined in the 2012 Survey Report. In it, CO Representatives found a number of factors to be critical in the production of quality evaluations, including quality of the consultant engaged in the evaluation, Country Office leadership, and internal oversight of the evaluation. This is highlighted in Exhibit 3.3 below, in which 68 per cent of respondents rate the quality of the consultant as “highly important”, compared to “country office leadership” at 49 per cent, and internal oversight of evaluation at 48 per cent.

108. The 2012 Case Study highlighted the fact that in a number of ROs, rosters have been developed to ensure a high and consistent quality of evaluations in the region.
Exhibit 3.3 Percentage of CO Representatives who rated “highly important” each of the factors which affect the production of quality evaluations (N=55)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of the consultant(s) engaged in the evaluation</td>
<td>68%</td>
</tr>
<tr>
<td>Country office leadership</td>
<td>49%</td>
</tr>
<tr>
<td>Internal office oversight of evaluation</td>
<td>48%</td>
</tr>
<tr>
<td>Availability of an M&amp;E specialist/focal point in Country Office</td>
<td>42%</td>
</tr>
<tr>
<td>Independence of the Evaluation</td>
<td>41%</td>
</tr>
<tr>
<td>Regional office oversight</td>
<td>20%</td>
</tr>
<tr>
<td>Existence of an evaluation advisory group</td>
<td>13%</td>
</tr>
</tbody>
</table>

Finding 19: M&E staff at UNICEF are confident of their skills in managing evaluations and performance monitoring, and to some extent in national M&E capacity development. They are less confident in evaluating emergency, peace-building and humanitarian situations.

109. UNICEF M&E staff engaged in evaluation includes M&E specialists as well as M&E focal points in some countries or sub-offices where there is no M&E specialist. The skill of the UNICEF staff who manage evaluations is a second factor which falls under Guideline 11. The 2011 M&E Survey Report showed that 64 per cent of M&E staff have less than five years of M&E experience with UNICEF. This reflects a high level of staff turnover within the function in general; although comparative data for other types of posts within the organisation was not available, this breakdown in staff experience suggests that the M&E function does not retain staff at a level necessary for having an experienced team of M&E staff in the organisation.
110. Across UNICEF regions, there is some variation in terms of the level of M&E experience of M&E staff. With regards to CO size, the larger COs tend to have more experienced staff: roughly 40 per cent of M&E staff in large and medium sized offices had more than five years of M&E experience versus only 25 per cent with more than five years of experience in small offices. When disaggregated by level of income in the country, low income countries tend to have a higher percentage of M&E staff with more than five years of experience than focal points in either lower middle income, or upper middle income countries. This finding has an important implication for the UNICEF M&E function. LMI and UMI countries tend to have stronger capacities for M&E in government ministries and partner institutions than LI countries. Yet, when viewing the data, the M&E experience of UNICEF focal points in these LMI and UMI countries is the lower. As was shown in other findings above, there is a risk that UNICEF loses credibility in the eyes of M&E partners in these countries.
M&E staff at UNICEF are confident of their skills in managing evaluations and performance monitoring, and to some extent in national M&E capacity development. They are less confident in evaluating emergency, peace-building and humanitarian situations. According to the 2011 M&E Survey Report, M&E staff scored themselves on average 7 out of 10 in developing the IMEP, results-based management, developing terms of reference and performance monitoring. They scored themselves respectively 4.22/10 and 3.28/10 on humanitarian and peace-building evaluations.
Finding 20: Familiarity with ethical issues is low among M&E staff at UNICEF.

112. There is only one reference to ethics within the Evaluation Policy. UNDP, on the other hand, dedicates more detailed attention to ethics and human rights in its Evaluation Policy under Section III (d): Ethics. “Evaluation should not reflect personal or sectoral interests. Evaluators must have professional integrity and respect the rights of institutions and individuals to provide information in confidence and to verify statements attributed to them. Evaluations must be sensitive to the beliefs and customs of local social and cultural environments and must be conducted legally and with due regard to the welfare of those involved in the evaluation, as well as those affected by its findings. In line with the Universal Declaration of Human Rights, evaluators must be sensitive to and address issues of discrimination and gender inequality.”

113. The current knowledge and practice of ethics in the M&E area was explored in the 2011 M&E staff survey. The majority of respondents (66 per cent) had not received any training on ethical considerations in the design of evaluations, surveys or research. Therefore, M&E staff familiarity with ethics and evaluations is low.

114. Twenty-four per cent of all respondents agreed that evaluations are reviewed for adherence to good ethical standards, as shown in. Over 60 per cent of respondents have not observed ethical problems in M&E work in their CO. Only 10 per cent of respondents have come across what they considered to be ethical problems.

Finding 21: Systematic training for M&E staff has not yet been fully institutionalised within UNICEF.

115. One way that technical knowledge is passed on to M&E staff is through training and professional development. The Evaluation Policy states that staff engaged in evaluation should have access to professional development opportunities, with accountability for ensuring this falling to both the Evaluation Office and the regional offices. It states that:

---

23 When reviews were carried out, they were most often done by CO staff.
“The evaluation function at the regional level focuses on oversight and strengthening the evaluation capacities of UNICEF offices and their government counterparts through the...coordination of evaluation capacity-building activities with the Evaluation Office at headquarters.”

116. The survey captured the extent to which respondents had received recent M&E training (face-to-face training, self-learning or distance learning) within UNICEF or through outside partners or institutions.

117. 2011 M&E Survey Report data showed that 63 per cent of M&E staff received UNICEF or UN M&E training in the past two years. However, 21 per cent have not received UNICEF/UN training on M&E in more than two years, and 16 per cent had not received any UNICEF/UN supported M&E training.

Exhibit 3.7 Latest participation in UNICEF/UN supported M&E training, in per cent of total responses (N=82)

118. Of those who did not receive any M&E training, 38 per cent had been working in the field of M&E for less than one year. What’s more, only 23 per cent of those who did not get UNICEF/UN M&E training had received external M&E training. This means that roughly 10 per cent of respondents had received neither UN/UNICEF nor external M&E training.

119. Beyond the technical areas highlighted above, M&E staff also cited a number of other priority areas for learning/training within the organisation. Out of 67 comments, the most prominent were:

- 13 (19 per cent) cited the need for basic or standard M&E training, approaches and analysis techniques
- 8 respondents (12 per cent) commented on the need for training on statistical analysis, data analysis, sampling strategies or management of databases
- 7 respondents (10 per cent) wanted to know more about the evaluation of policies or advocacy campaigns
- 6 respondents (9 per cent) wished to have training on equity analysis or equity focused planning techniques
• Other areas for training (7 per cent or less) included: monitoring, knowledge management, complex evaluations and budgeting for an evaluation

120. Some of the priority areas highlighted by M&E staff were reiterated in the 2012 Case Study Report, including an interest in skills development for evaluations of policy and equity approaches.

Finding 22: UNICEF has demonstrated good oversight and leadership in Human Rights and Gender but a fairly low/very modest level of reporting on equity, human rights, gender equality and women’s empowerment.

121. The importance of human rights, gender equality and equity is noteworthy, given UNICEF’s considerable efforts to support and define a human rights-based approach and equity approach in the organisation. UNICEF’s mission statement and its accordance to the normative principles of the Universal Declaration of Human Rights (UDHR), Convention on the Rights of the Child (CRC) and Convention on the Elimination of Discrimination against Women (CEDAW) form the backdrop for a Human Rights-Based Approach to Programming (HRBAP) and the Equity approach. In the work of UNICEF and other UN agencies, the application of these principles means that country programmes of cooperation need to identify issues of exclusion and injustice as central concerns in the dialogue with national partners. As it pertains to evaluations, issues of exclusion and injustice in all sectors and in all geographic areas must be included in the way in which an evaluation is carried out. Data disaggregation in order to highlight those excluded populations is one way in which this is manifested in evaluations. Another is to be inclusive in evaluations, and ensure that marginalised groups and excluded populations are considered in the evaluation process. This is closely linked with the human rights principle of participation.

122. In addition to UNICEF, UN organisations such as UNDP, UN Women and UNHCR all refer to human rights more generally in their policies. UNHCR’s policy refers to a:

“Focus on beneficiary rights, consideration of participation in development of evaluation methods; look towards disaggregated data; focus on protection and human rights issues”

123. How are human rights principles manifested in evaluations? The Evaluation Policy cites that “UNICEF will apply, and advocate for increased efforts to evaluate the different effects of its work on women, men, boys and girls, in accordance with the UNICEF gender policy and using gender analysis and gender-disaggregated data. The chapter on Monitoring and Evaluation in the PPP also makes claims to human rights and gender through references to HRBAP, data disaggregation, and participation of key stakeholders, in line with UNEG Norms and Guidelines.

124. Assessment of human rights principles in evaluation has been provided through the GEROS. In particular, it tracks the evaluation’s focus on beneficiary rights, consideration of participation in development of evaluation methods; look towards disaggregated data; focus on protection and human rights issues. More recently, equity evaluations also fall under this category. In particular, the GEROS quality

24 Article 1 of the Universal Declaration of Human Rights states: “All human beings are born free and equal in dignity and rights.” The related principle of non-discrimination (on the basis of race, colour, gender, language, opinion, origin, disability, birth or any other characteristic) is expressed in Article 2 of the CRC. Issues of exclusion, inequality and injustice are therefore central concerns in the dialogue with national partners, and in the development of programmes of cooperation. The disaggregation of indicators –for instance by age, sex, ethnic group, household composition, geographic area – permits the assessment of discrimination and inequality and is therefore essential for making rights-sensitive programming decisions. While the well-being of all children is important, UNICEF gives priority to the most disadvantaged children and the countries in greatest need (see UNICEF Mission Statement). Guidance on reaching marginalised children and families is available in Chapter 6, Section 20. Positive measures are vital to ensure that public policies benefit all children, and that all children have access to quality basic services and other opportunities, based on, among other things, distributive justice.
criteria include four elements which encompass Human Rights, Gender and Equity. These are reviewed and rated and reported on through the GEROS system:

- Did the evaluation design and style consider incorporation of the UN and UNICEF’s commitment to a human rights-based approach to programming, to gender equality, and to equity? This could be done in a variety of ways including: use of a rights-based framework, use of CRC, CCC, CEDAW and other rights related benchmarks, analysis of right holders and duty bearers and focus on aspects of equity, social exclusion and gender. Style includes: using human-rights language; gender-sensitive and child-sensitive writing; disaggregating data by gender, age and disability groups; disaggregating data by socially excluded groups

- Does the evaluation assess the extent to which the implementation of the evaluated object was monitored through human rights (inc. gender, equity & child rights) frameworks?

- Do the methodology, analytical framework, findings, conclusions, recommendations & lessons provide appropriate information on human rights (inc. women & child rights)?

- Do the methodology, analytical framework, findings, conclusions, recommendations & lessons provide appropriate information on gender equality and women’s empowerment?

125. Given the criteria noted above, the GEROS provided a rating for evaluations on their human rights and gender equity components. The GEROS found that while “UNICEF has been at the forefront of international efforts to recognise human rights, gender equality, women and children’s empowerment, and socio-economic equity within the evaluation function. This is not being translated into strong performance in these areas.” The 2011 GEROS reported:

- Overall, only 9 per cent of reports included equity issues to a satisfactory level, according to UNICEF standards: out of 89 reports, one report was rated outstanding and seven were rated as confident.

- In human rights, 19 per cent of reports were rated satisfactory – no evaluation was found to be outstanding in this regard.

- For gender equality and women’s empowerment issues: 20 per cent of reports were rated as satisfactory.

Finding 23: Growing interest in equity evaluations at the country level shows the relevance of the universality principle at UNICEF.

126. In line with the universality principle, respondents in the 2012 Case Study Report commented on a more recent focus and interest on the equity approach, and developing improved methods of evaluating the effect of policies and strategies on the most vulnerable populations. Using the Monitoring Results for Equity System (MoRES), respondents cited the need to get government and other partners on board with paying greater attention to the bottom 2 per cent, or the most vulnerable populations, as they carry out evaluations.

127. This emerging interest in equity was confirmed through the GEROS, which stated: “…many evaluators are struggling to understand how to integrate the equity-intention into a robust evaluative process. A couple of reviews also note that this is further compounded by a lack of equity-relevant programme-data available to evaluators even when they do develop analytical frameworks.”

128. The Evaluation Office is leading efforts to better equip UNICEF offices and partners in evaluating pro-equity interventions. A manual on designing and managing equity-focused evaluations was developed.
and disseminated with accompanying training; an electronic resource centre was created in partnership with UN-Women; and a webinar series was delivered in partnership with several organisations.  

However, where it refers to human rights principles more generally, the 2012 Case Study Report highlights that the political context and government-led evaluation may affect governments’ perception of the value of evaluations. This is particularly the case in authoritarian regimes where evaluations are perceived in a negative light, or where governments wish to shy away from highlighting some marginalised groups. The role of UNICEF in these sensitive MI countries, as pointed out in the 2012 Case Study Report, involves building a national government appetite for independent evaluation, showing where it can add value, and communicating the message that evaluations can be beneficial for government ministries in their ability to validate important programmes and initiatives. Therefore there may be cases where UNICEF must uphold its obligations to report on human rights and equity principles in its evaluations in cases where governments may not wish, for example, to disaggregate data and reveal important weaknesses and gaps.

**Guideline 12: Evaluation findings, recommendations and lessons are made public and disseminated to all stakeholders concerned.**

**Finding 24: UNICEF’s dissemination of evaluation finding shows a good level of commitment to information sharing, but there remains room for improvement.**

Good dissemination of findings, recommendations and lessons learned is part of evaluation use. The UNEG Norms and Standards & Ethical Guidelines rightly emphasise that evaluations are valuable to the extent to which they serve the information and decision-making needs of intended users.

Making the results of evaluations public is also part of accountability and transparency. Most development organisations have inculcated public access and transparency of evaluations into their Evaluation Policies.

In the 2012 Case Study Report, external communication and dissemination of evaluation findings among national partners was generally reported as strong among case study countries, regardless of the size of the CO or country typology. Most respondents cite the participation of partners as part of advisory groups and engaging them in planning as key factors which facilitate dissemination and sharing of evaluation findings within the country.

The 2012 Survey Report asked the extent to which evaluation findings and recommendations were systematically disseminated externally. Given three options: “Never”, “Sometimes” and “Always”, CO Representatives overwhelmingly selected “Sometimes”, with 75 per cent of the total responses falling into this category. A remaining 18 per cent responded that they “always” disseminated evaluation findings, while the remaining 7 per cent fell into the “Not Sure/Don’t know.” No CO reported never having disseminated evaluation findings.

A similar question was asked of M&E specialists in the 2011 M&E Survey Report, and revealed a similarly mixed level of compliance. While 62 per cent of respondents suggested high compliance (either a 4 or a 5), a further 37 per cent of respondents rated compliance at a level 3 or lower. This finding

25 UNICEF Executive Board 2012. Annual report on the evaluation function and major evaluations in UNICEF

26 In that survey, respondents were asked to rank from 1-5 the level of compliance (1= low compliance and 5=high compliance) to good practice on communicating the finding of M&E activities. While 62 percent of respondents suggested high compliance (either a 4 or a 5), a further 37 percent of respondents rated compliance at a level 3 or lower.
confirms that compliance in disseminating findings is good, but that COs have some improvements to make before meeting a standard of full compliance.

**Guideline 13: Evaluation must be duly considered, with management responses and action plans made public.**

**Finding 25:** Management responses are increasingly tracked and followed-up at UNICEF. There are, however, certain instances where accountability for management responses remains burdensome or unclear.

135. The management response as a requirement is one way to better ensure the use of evaluation results. Evaluation Policies across the development spectrum include this important phase of evaluations to better track and ensure use of evaluation results: OCHA specifies that the management response is required three months after the completion of an evaluation. For UN Women, a management response is required after 6 weeks.

136. In addition to UNEG guidelines, the QCPR report correctly points out that individual entities have invested in strengthening their monitoring of management responses and recommendations. However, the report also points out that the 

> “...absorptive capacity of the system is increasingly being tested as evaluation outputs increase across the system. Various studies concur that the system may not have the capacity to engage effectively with the recommendations of so many evaluations.”

The overburdening of programme staff with evaluation findings and management responses remains an issue to consider across all agencies.

137. UNICEF’s Evaluation Policy requires that a management response be completed for each evaluation. The Annual Report on the Evaluation Function suggests that advocacy and oversight has resulted in improved compliance in the preparation of management responses and in uploading completed responses to the tracking system. Upload rates increased from 12 per cent in 2009 to 47 per cent in 2010 and 65 per cent in 2011 (for evaluation reports completed the previous year).

138. To further improve compliance, appropriate performance indicators have been included in the evaluation dashboard in VISION; the relevant guidance is being revised to take account of recent experience with the system; and an e-learning module is under preparation.

139. The Independent review of UNDP’s Evaluation Policy also pointed out that “... its Management’s responses, and use, of independent evaluation findings and recommendations is improving.” A review of the Evaluation Resource Centre (ERC) database, the equivalent of the Global Evaluation Database at UNICEF suggests that the requirement for management responses to corporate-level evaluations is also increasingly fulfilled although there is still room for improvement. However, it also added that

> “... as far as the partner government representatives are concerned., the management response is not shared with partner governments and there is actually a break in the chain from recommendations to management response to changes in programming. This has obvious implications for country ownership of the evaluation products, and how management responses

---

27 Quadrennial comprehensive policy review of operational activities for development of the United Nations system, p.56, par. 3

are used to influence decisions that partner countries endorse, for which UNDP is eventually accountable to the Executive Board.”

140. The 2012 Case Study Report respondents suggested that they understood that evaluation findings and recommendations must ultimately become the object of a management response. This is stated in the UNICEF Evaluation Policy and is now part of the Management Information Dashboard, which notes the global level of completion of management actions across countries and regions. This is mostly seen as a positive accountability measure within the organisation, with the caveat of many interviewees that management responses do not fully capture the use of completed evaluations.

141. While the intent of management responses is largely understood by Representatives, follow-up on management responses is not always consistent or clear. For example, accountability for the implementation of certain evaluation recommendations may lie with government partners. In the cases of joint programming, they may lie with other UN or external partners. In both cases, UNICEF Representatives are accountable for elements of a management response that they may have difficulty controlling. Some case study interviewees cite the engagement of UNICEF senior management and government partners in the evaluation process (from planning and onwards) as a strong factor affecting the internal use of evaluation findings and the implementation of management responses.

142. Another factor which is unclear at the CO level is the relevance of all evaluation recommendations, for example, when management response recommendations do not appear to be sustainable or appropriate to the programming at hand. In the same vein, some interviewees stated that the formulation of management responses should not be driven by an EO requirement or included as a component in the performance reviews of senior managers. They feel that this may create a disincentive for some UNICEF COs to carry out important evaluations.

Guideline 14: At minimum, the executive summary of the evaluation must be translated into relevant local language(s) and made available to stakeholders.

Finding 26: Communication of UNICEF’s evaluations in local languages and to a range of stakeholders is not well tracked or known.

143. In a decentralised programming environment, reaching the local population with the results of any evaluation is an important exercise. UNICEF’s Evaluation Policy is not unique in its aim to reach local populations, including stakeholders, although defining who the stakeholders are that should be reached is not made clear in the Evaluation Policy.

144. To facilitate wider use and dissemination of evaluation findings at UNDP, for example, the executive summary of all independent evaluations are translated into the three working languages of UNDP (English, French and Spanish). The Evaluation Office is responsible for a communication and outreach programme to expand the effective use of independent evaluation findings and the development of a learning community. Like UNICEF, UNDP country offices are encouraged to translate a summary of evaluations into local languages and use other means in order to inform stakeholders of findings and for learning. UNDP’s Policy states that “Evaluation findings and recommendations should be presented in a manner that will be readily understood by target audiences.”

145. The extent to which this guideline has been applied is unclear and not reported. As Finding 25 points out, the Independent review of UNDP’s Evaluation Policy found a disconnect between recommendations and the management response with government stakeholders. Anecdotally, it appears common that non-government stakeholders, some of whom may have participated during an evaluation, were never made aware of evaluation findings or recommendations.
Guideline 15: Evaluation findings of relevance to the Executive Board are to be brought to its attention.

Finding 27: UNICEF has well-established systems to ensure sharing of evaluation findings at the level of the Executive Board.

146. The UNICEF Executive Board has embedded the requirement to report back periodically on UNICEF’s field level application of its programmes and strategies across the organisation, as well as how policy and pilot interventions are being implemented. UNICEF has well-established policies in place to encourage transparent reporting of its results with government partners and donors at the country level. In all cases and at all levels, there is a strong importance placed on demonstrating continued relevance, effectiveness, and efficiency of programmes and interventions.

147. At the senior management level, there is a broad appreciation for the role played by the EO to institute and monitor evaluation quality as a clear indicator that can be tracked. The GEROS results are shared widely at Headquarters and Executive Board levels. The EO has also been lauded for taking steps in improving overall evaluation quality for emergency evaluations, such as developing e-learning tools, standard operating procedures and improved guidance for emergency evaluations in the PPPM.

148. Responses during the 2012 Case Study Report suggest that the relevance and use of evaluations by the Executive Board is strong and well appreciated. The Annual Report of the Executive Board features information on the number of evaluations completed, their general rating and the number of staff involved in Evaluations. The Global Evaluation Report Oversight System (GEROS) contains quality rating on evaluations (and associated evaluation findings) submitted to the UNICEF Evaluation office.

3.2 Findings on MTSP Priorities and Other Areas not Covered by the Current UNICEF Evaluation Policy

Finding 28: Efforts by the EO to improve humanitarian evaluations have resulted in a greater appreciation of their use in recent years. But there are some challenges in carrying out such evaluations in emergency situations within the decentralised structure.

149. UNICEF’s MTSP placed specific focus on improving the evaluation of humanitarian situations. The UNICEF Evaluation Policy refers to the evaluation of humanitarian action, particularly in the context of Real-time Evaluation (RTE). Discussions during the 2012 Case Study Report suggest that RTEs are but one of several types of evaluation carried out in emergency settings. Evaluability is another consideration for evaluations of humanitarian emergencies, as well as strategic evaluations such as those which respond to specific organisational capacities in emergencies. The report highlighted the efforts that have been made by the EO to develop tools for these situations. The EO has also been lauded for taking steps in improving overall evaluation quality for emergency evaluations, such as developing e-learning tools, standard operating procedures and improved guidance for emergency evaluations in the PPPM. These measures go some way to respond to the gaps outlined in the 2011 M&E Survey, which placed particular attention on the gaps in knowledge of M&E staff on humanitarian evaluations.

150. Demand for emergency evaluations can be sub-divided depending on the levels of the emergency, with Level 1 and Level 2 evaluations carried out at the CO and RO levels with some support from the EO level as needed. Level 3 emergency response evaluation has a higher demand, particularly from HQ, due to the need to ensure accountability through evaluation. This demand is particularly notable at HQ, which takes a much more active role in evaluating Level 3 responses as compared to Levels 1 or 2, due to a lack of capacity in many COs to evaluate humanitarian emergencies. This has, however, placed a heavy burden on the EO to manage humanitarian evaluations in a more centralised way. This is contrary to the Evaluation Policy’s decentralised structure.
151. In addition to the lack of knowledge on humanitarian evaluation methodologies, the timeliness of humanitarian evaluations, and RTEs in particular, has been a factor that has dissuaded COs from engaging further in such evaluations. Done during a response, they place pressure on teams on the ground to engage in such evaluations while responding to an on-going crisis.

**Finding 29: Evaluations across MTSP focus areas, strategies and operational effectiveness are well balanced in the organisation.**

152. The MTSP highlights the importance of ensuring a balance in evaluation across the UNICEF focus areas and in all areas of effectiveness. UNICEF’s Evaluation Policy is more explicit in its demands for evaluation coverage at the CO level. In order to strengthen the evaluation function, it gives responsibility at all levels (CO, RO and Division Directors) to *Institute measures to ensure that evaluations are strategically selected and that all major programme components are evaluated during the programme cycle, allocating adequate resources.* At the CO and RO levels, this has caused some confusion and concern, especially in COs which are becoming more strategic and upstream in their selection of evaluations to carry out.

153. According to the Annual Report of the Evaluation Function, evaluations conducted in 2010 and reviewed under the Global Evaluation Reports Oversight System (GEROS) were spread across the medium-term strategic plan focus areas as follows: 25 per cent examined child survival and development issues; 15 per cent examined education and gender; 10 per cent examined HIV/AIDS; 16 per cent examined child protection; 1 per cent examined policy advocacy and partnerships, and 33 per cent examined cross-cutting themes and organisational performance. This shows coverage across all areas of strategic focus.

**Finding 30: Increased management attention to evaluation and strengthened organisational capacity in evaluation are mutually reinforcing. Progress is due in part to GEROS and the VISION dashboard.**

154. Two additional areas of focus for evaluation, as highlighted in the MTSP, are organisational capacity in evaluation and a heightened management attention to the evaluation function. Both are mutually reinforcing: the increased attention brought about by rating tools such as GEROS and the vision dashboard have led to the need, at RO and CO level, to develop the organisational capacity to manage and improve the function as a response. As noted in other findings, RO systems have been developed for quality assurance at the CO level. Heightened coordination between Evaluation Specialists at the CO, RO and EO levels, through network meetings and webinars have helped to share lessons and bring greater professionalism to the function.

155. At the senior management level, there is a broad appreciation for the role played by the EO to institute and monitor evaluation quality as a clear indicator that can be tracked. The GEROS results are shared widely at Headquarters and Executive Board level, and many at the HQ level feel that evaluation quality and use will be enhanced due to these improvements.

**Finding 31: M&E staff accountability is not well defined in the current Evaluation Policy, affecting evaluation independence and quality.**

156. UNEG has developed a Normative Framework for the Evaluation Self-assessment in order to allow member organisations to review their own Evaluation Policies and functions. It outlines the key Norms and Standards alongside assessment criteria in order to highlight potential gaps. Many of the areas outlined are covered in the above findings. In particular, Finding 17 points to the reporting of the M&E staff at the decentralised level, and the fact that 30 per cent of M&E staff report to someone other than the Representative or Deputy Representative. This reporting structure is currently not well defined within the Evaluation Policy, reflecting weak structural independence in the organisation.

157. The 2012 Case Study Report and 2011 M&E Survey Report underline the important role of M&E staff in carrying out the responsibilities and duties outlined in the Evaluation Function. While overall
accountability within the CO lies with the Representative, the accountabilities of staff responsible for evaluation are not clearly specified. At the design and implementation stage of evaluations, for example, the Evaluation Policy is silent on who is accountable for the development of independent TORs, the selection of consultants, or the quality review of draft reports. Instead, these are specified in the 2009 Policy Directive at 3.2.5 under the technical support provided by M&E specialists. This lack of clarity may underline the mixed compliance results on a number of independence and quality criteria in Finding 17.

158. GEROS and the VISION dashboard have created the incentive for greater overall accountability for the evaluation report at both the CO and RO levels. It has been noted in the 2012 Case Study Report that ROs have felt an increased accountability for results in their region due to the VISION dashboard. Yet, at times, they note that poor quality evaluations in their region come from COs that have chosen not to share TORs at the RO level. The Evaluation Policy is silent on accountability for sharing TORs, methodologies and draft reports with the ROs.

Finding 32: The steps to ensure transparency of evaluations at all stages are poorly defined in the Evaluation Policy.

159. UNICEF guidance on transparency as outlined in the Evaluation Policy was highlighted in Finding 24 as lacking in practice. UNEG assessment criteria as defined in its Self-Assessment go much further than UNICEF’s Evaluation Policy in this regard. For example, UNEG Standards propose that “A disclosure policy should ensure the transparent dissemination of evaluation results, including making reports broadly available to the Governing Bodies and the public, except in those cases where the reasonable protection and confidentiality of some stakeholders is required.” UNICEF’s Evaluation Policy is silent on its disclosure of evaluation products at all stages. This includes the sharing of an evaluation plan with key stakeholders, sharing of final TORs with stakeholders and how evaluations should be made publicly available at local levels. UNEG Ethical Guidelines for Evaluations suggest that at country level, evaluation findings shall be presented and discussed at the appropriate national or local level, to enable stakeholders to respond to them, and ideally before the evaluation report is complete. As noted in Finding 24, the Evaluation Policy is clear on the sharing and dissemination of Evaluation results and findings, but in practice, has found some gaps in consistent implementation.

160. UNICEF’s Evaluation Policy is also unclear on the issues surrounding sensitive issues in evaluation reports. Findings 7, 13, and 23 highlight the practical challenges when UNICEF COs must engage in evaluations of sensitive topics or where governments are less willing to engage in evaluation. In these cases, nothing is stated in the Evaluation Policy on how UNICEF at all levels can report on dissenting views, objections or requests for confidentiality. This issue is of particular importance on evaluations with disaggregated data covering human rights or equity issues.
4. Conclusion and Recommendations

4.1 Conclusions

161. The purpose of this review is to assess the demand for evaluation and the utilisation of evaluation results and, making use of existing documentation, to make an overall assessment of the current status of the evaluation function in UNICEF. The consultancy used the 15 guiding principles outlined in the Evaluation Policy to present findings on UNICEF’s decentralised evaluation function and the governance and management of evaluation activities within UNICEF, including the role of management at all levels in addressing evaluation needs and demands, in ensuring effective utilisation of evaluation results, and in taking account of cross-cutting issues. The following conclusions and recommendations are grouped to comment on both the Evaluation Policy as well as its implementation while providing clear evidence of what worked and what did not work with regard to the demand for evaluations, quality of evaluations, and use of evaluations. However, these sub-headings should not distract from the inter-related and contributory nature of each of these factors with one another.

The Evaluation Policy - Overall

162. Upon review of the Evaluation Policy and in particular its guiding principles under the context of demand for evaluation and the quality and use of evaluations, the policy remains relevant and closely aligned to the UNEG Norms and Standards. The decentralised evaluation function championed by the Evaluation Policy aligns well with the operational nature of UNICEF’s decentralised programming environment, given UNICEF’s role to act as an authoritative voice on children’s issues at the country level and the necessity to reflect the differences and particularities of each country and region. The decentralised nature of the evaluation policy is also relevant in the current policy context of the UN. As noted, it is well aligned to the most recent QCPR which highlights a continued focus on decentralised evaluation and national capacity development. The challenge lies in fully implementing and in some cases clarifying the policy, rather than radically altering it.

163. The Evaluation Policy does have some drawbacks, however. First of all, the policy fails to delink the funding for evaluations from the size of programming expenditures. Under the current policy, the dollar amount of funding for evaluations is proportional to the size of the programme. This will have increasing importance in the coming years as COs in MI countries with fewer resources but greater technical evaluation needs try to remain relevant in the evaluation environment. This will also have relevance in COs that require human and financial resources for both evaluations and national capacity development in evaluation. The Evaluation Policy isn’t clear on whether or not country-led evaluations, rather than UNICEF-led evaluations, are adequate in the evaluation of “all major programme components” in a cycle. This obligation within the Policy has been interpreted in different ways at the CO level.

164. The Evaluation Policy lacks a clear set of accountabilities for M&E or evaluation personnel. This may have been the intent of the Policy so as to leave flexibility at the decentralised CO level to set out the resources and accountabilities required. Data suggests that in most cases, COs have indeed committed to staffing for M&E resources. Therefore, the absence of these staff within the policy obfuscates their critical role within the CO for evaluation as well as research, surveys and for tasks such as performance monitoring or the development of national capacities in evaluation. In fact, the study has shown that staff allocation of time does not necessarily favour evaluation over other tasks, including studies and research, or other monitoring tasks within the organisation. The incentives and skills required for M&E staff to champion national M&E capacity development are similarly undefined.

165. While the Evaluation Policy highlights the importance of country-led evaluations and joint evaluations, it does not explicitly couch them within the broader Aid Effectiveness agenda. A more clear elucidation of how the aid effectiveness agenda is implicated within the evaluation function would provide...
a stronger impetus to UNICEF COs to leverage partnerships at the HQ, RO and CO levels, and to advocate for greater country leadership in evaluation.

166. Finally, new and emerging issues now populate the aid and programmatic agenda and these are not reflected in the Policy. In particular, the Equity Approach within UNICEF, which has received a good amount of attention in recent years by the EO, is not present in the current policy.

Implementation of the Evaluation Policy

167. Overall the synthesis study suggests that UNICEF’s current evaluation practice holds up well to the guiding principles espoused by the Evaluation Policy. UNICEF staff have noted a continued demand for evaluations both within and outside the organisation for the purposes of evidence-based policy advocacy, decision making and accountability. While overall, financial resources, leadership and donor demand are strong determinants that prioritise evaluations, national priorities for evaluation are also playing an increasing role in determining demand for evaluations, particularly in MI countries. In addition, it is noted that joint evaluations with other UN agencies or with government partners and multi-country evaluations are not well defined or strongly integrated into UNICEF’s current portfolio of evaluations, a problem that is shared in the entire UN system.

168. In the context in COs where evaluations were not carried out, the CO had not done so for reasons which are often country-specific and which tend to focus on a lack of external demand or priority from the government. This underlines the important role that UNICEF has in building national capacity in evaluation. In some cases, while Senior Management recognised the need for some level of review of performance, this was not done through a formal evaluation but rather through informal monitoring channels, studies or other research. In humanitarian contexts, the demand for evaluation is determined in part by funding thresholds. Carrying out humanitarian evaluations currently reflects an imbalance between demand for such evaluations and the technical know-how to produce and oversee them, and has placed a considerable burden on the EO for management and oversight.

169. With regards to evaluation quality, the GEROS system and management dashboard have fomented the correct incentives to produce improved quality evaluations at the CO level according to the Evaluation Policy and UNEG guidelines. The EO and ROs have invested significant resources in tracking and supporting evaluation quality with a view to improving UNICEF’s overall evaluation credibility. This effort has not gone unnoticed by Senior Management at all levels of the organisation and has helped to build and sustain a culture of evaluation in the organisation.

170. This comes, however, with some concerns about the extent to which evaluations should be the only product assessed for quality. Studies, reports, and surveys are other types of assessments which are not included in the GEROS mechanism. Other concerns revolve around the need to balance the quality role at the CO level with the myriad other tasks to which evaluation specialists attend, including monitoring, situation assessment, and capacity building. This points the way to a greater reliance on partnership in evaluation in order to leverage knowledge and take advantage of skills in other organisations.

171. Not surprisingly, COs with highly rated evaluations are characterised by strong planning, well developed Terms of References, good national evaluation capacity and an adherence to a quality assurance process that often includes RO involvement. Constraints in terms of having the right mix of staff skills and sufficient financial resources to carry out an evaluation agenda will continue to persist; these constraints suggest that a one-size fits all approach within UNICEF is becoming more and more difficult to administer. For example, certain MI countries require more advanced expertise in capacity development in order to support government with policy evaluations, while other countries require advanced expertise in emergency evaluations. Some COs face challenging political contexts for evaluations, while others face resource constraints and will rely on partnerships and a strong regional presence to carry out evaluations. Quality is also affected by the skills of M&E staff: M&E staff at UNICEF are confident of their skills in managing evaluations and performance monitoring, and to some extent in national M&E capacity development. They
are less confident in evaluating emergency, peace-building and humanitarian situations and their familiarity with ethics for evaluations is low.

172. Evaluations have no value to UNICEF or its partners if they are not adequately and widely used for their intended purposes. Concerning the use of evaluation results, evaluations appear to be most frequently used for decision-making purposes and evidence-based policy advocacy. Their greatest use is at the country level, where UNICEF uses evaluation results to refocus or advocate for its programming interventions with partners in favour of children’s rights. Evaluation results are often used to formulate management responses but their tracking and implementation has only recently become more systematic at the country level.

173. External communication and dissemination of evaluation results is seen as generally good at the country level, especially when government partners or other stakeholders are involved in the planning and design of evaluations, or where they participate in an advisory group capacity. However, language barriers and full engagement with governments in some countries has hindered full transparency of evaluations, including issues such as stakeholder presentations and response to evaluation findings, and whether or not the local language was used for evaluation summaries. Communications and dissemination at headquarters is also seen as relatively strong in responding to donor needs, and use of evaluations at headquarters results in stronger technical inputs to programme strategies and the UNICEF MTSP. The processes and dissemination strategies for the internal sharing of evaluation results for knowledge generation, however, are less fully developed within the organisation. As such, knowledge generation and management emerging from evaluations can be improved.

4.2 Recommendations

174. The following recommendations are intended to contribute to the overall review of the Evaluation Policy and strengthen its implementation.

Recommendations on the Evaluation Policy

Recommendation 1: UNICEF Executive Board, through the Evaluation Policy, should commit to spending one per cent of programme resources specifically on evaluation and should seek a way to ensure minimum funding levels for evaluations and capacity development for country-led evaluations.

175. In 2006, the panel reviewing the UNICEF evaluation function pointed out that the “ability to budget for evaluation is a key element of independence.” The organisation has been a leader in setting out a budget for assessments and should now go further by ensuring that one per cent of programme resources be spent specifically on evaluation. The EO, in coordination with the Regional Offices, Programme Division and the Division on Policy and Practice should take a leadership role in advocating for solutions to the financial constraints facing some COs in carrying out evaluations. This may include financial mechanisms to ensure that programmes set aside resources for evaluations, or a formal discussion during the Programme Budget Review on evaluations to ensure that evaluation and M&E remain part of the core business in all UNICEF COs – keeping in mind that funding for evaluation based on the size of the programme to be evaluated is clearly detrimental to meaningful evaluations, especially in countries where UNICEF is engaged in influencing changes through advocacy and policy influence rather than direct service delivery.

Recommendation 2: The UNICEF Evaluation Policy should more clearly incorporate the role of the M&E staff at the CO level.

176. While accountability for the decentralised evaluation function rests with the Country Representative, this review suggests that in some cases, delineating some accountability at the level of the M&E specialist could be beneficial. This would allow for a greater appreciation of the role of the M&E specialist in, for
example, capacity development for country-led evaluations. It may also help to ensure a stronger role for the M&E officer and greater accountability for some areas of evaluation governance.

**Recommendation 3:** The UNICEF Evaluation Policy should be updated to include greater clarity on the aid effectiveness agenda, the requirement to evaluate every programme at least once per cycle, and the Equity approach.

177. A number of updates of the Evaluation Policy are timely. The first is to include greater clarity linking the aid effectiveness agenda with the Evaluation Policy. This should help provide impetus for greater coordination on evaluation between UNICEF and other partners, and provide the EO with the incentive to advocate with donors and Representatives to better understand the important role of capacity development for country-led evaluations.

178. In a similar vein, more clarification is required on the requirement to evaluate every programme at least once per cycle, particularly in instances in MI countries where governments take the lead on evaluations (with or without UNICEF support) or in instances where a policy evaluation covers multiple programmatic areas.

179. Finally, UNICEF’s focus on the equity approach in recent years should be reflected in the updated Evaluation Policy.

**Recommendations on the Implementation of the Evaluation Policy**

**Recommendation 4:** The EO should propose guidance and oversight on issues such as sharing of evaluation work programmes, TORs and draft reports with stakeholders; obtaining their views; and managing dissenting opinions on evaluations.

180. The goal of having a decentralised evaluation function is to engage most closely where UNICEF is working most directly – at the country level. The relevance of the decentralised evaluation function has been reiterated during the recent QCPR in 2012. While the UNICEF Evaluation Policy is clear with regard to engagement with country-led programming, national capacity development, and engaging with stakeholders throughout the evaluation process, several findings in this review have shown that improvements need to be made in practice. This points strongly to the transparency criterion outlined in the UNEG guidelines. The EO should propose guidance and oversight on issues such as sharing of evaluation work programmes, TORs and draft reports with stakeholders; obtaining their views; and managing dissenting opinions on evaluations. The EO should also work to clarify in the Evaluation Policy or in other guidance exactly how evaluations should be made public at the CO level, including the language in which reports or summaries of reports should be made available and their dissemination.

**Recommendation 5:** UNICEF ROs and EO need to provide better oversight and incentives to UNICEF’s role in capacity development of national evaluation systems, with a focus on leveraging partnerships with other actors in order to be more efficient in this regard.

181. UNICEF has instituted a number of mechanisms and guidance on capacity development of national evaluation systems. For this reason, it is often seen as a leader in this area within the UN. One adverse effect of the VISION dashboard has been perceived decline in relative attention afforded to this area of engagement due to few clear incentives. M&E staff and RO staff feel that without a more clear focus on national capacity development, this important function will be seen as less important than evaluations. The EO should work to encourage a stronger oversight or incentive mechanism to monitor capacity development steps in countries. For example, country level monitoring of the number of relevant country-led evaluations carried out with UNICEF support should be encouraged and this indicator should be part of evaluation staff performance assessments. Similarly the activities linked to building national evaluation
associations. This would help to balance the attention afforded to this area as compared to evaluation quality indicators. It would also help to facilitate the transitional role that UNICEF is playing in some countries with regard to evaluations.

**Recommendation 6:** The EO should clarify the extent to which joint evaluations are encouraged and work with UNEG and other agencies in learning more about the value of joint evaluations in the current programming environment.

182. As highlighted in the 2012 Survey Report, more clarity and oversight is required in the implementation of the Evaluation Policy on the extent to which joint evaluations are encouraged. This recommendation goes beyond a single UN organisation and reflects a wider need of the UN to take stock of the culture gaps in engaging in more joint evaluations. There is some evidence that joint evaluations can be welcome and can lead to better quality reports. On the other hand, it is possible that the dearth of joint evaluations reflects the limited number of joint programmes being carried out. The QCPR highlighted the poor implementation of the M&E of UNDAFs, and the Evaluation Policy, while highlighting joint evaluations, does not specify this in relation to Paris Declaration objectives. It is suggested therefore that the EO work to clarify this in the Evaluation Policy strategy and that UNICEF carry out additional research jointly with UNEG on why joint evaluations are not more pervasive across the UN as a whole.

**Recommendation 7:** The EO should develop KPIs for the design of evaluations to help ensure better compliance with accepted good practices.

183. Several findings in this review highlight the uneven application of good practices in independence and transparency of evaluations at all stages of the evaluation process. In some cases, additional oversight by RO and EO could alleviate some of the challenges in meeting high standards; however, the burden in doing so could also outweigh the benefit. Instead, the EO should develop KPIs for the design of evaluations, and the accountabilities of the M&E staff at the CO level should be enhanced and clarified in the Evaluation Policy, as noted in recommendation 2 above. This places the burden of ensuring transparency and independence, along with other criteria, firmly in the hands of the M&E staff, with oversight and accountability directly to the Representative or Deputy Representative. This will simultaneously reduce the follow-up required at the RO level to track and follow evaluations in all COs and instead provide more timely and needed evaluation support.

**Recommendation 8:** The EO and ROs should agree on a way to identify COs that have not carried out any evaluations for more than two years and assist them in developing an evaluation strategy.

184. Finding 10 in the review noted that some COs had not carried out any evaluations over several years. It is suggested that the EO and ROs agree on a way of tracking and responding in cases where COs have not conducted any evaluations over the last three years. The EO and RO should then take steps to ensure that those COs are supported with clear strategies to gather evaluative data for short, medium and long-term planning and decision making.
Appendix I List of Findings

Finding 1: UNICEF’s decentralised evaluation function has its grounding in the country-driven programming of the organisation.

Finding 2: UNICEF evaluations are seen as relevant across all levels of the organisation, although knowledge generation in general is seen as weak.

Finding 3: UNICEF M&E staff at COs, ROs and the EO have integrated national capacity development into their day-to-day work in response to this noted priority.

Finding 4: While UNICEF efforts in national capacity development have improved the culture for evaluation at the country level, some challenges have been noted in how UNICEF prioritises national capacity development (e.g., in terms of providing adequate evaluation expertise, balancing capacity development with monitoring evaluation quality, and the RO role).

Finding 5: UNICEF engages in evaluations in partnership with national governments, UN agencies, and other partners to some extent, but this guideline is not well-defined and the results of these efforts are not consistent.

Finding 6: Evaluations are most often used for evidence-based policy advocacy and decision-making purposes.

Finding 7: While UNICEF’s evaluation system is decentralised, there are challenges in ensuring sufficient human and financial resources for this structure to meet expectations.

Finding 8: Although UNICEF is meeting its minimum expenditure target of 3 per cent for evaluation, studies and research, its expenditures on evaluation alone are significantly below one per cent.

Finding 9: The demand for and quality of evaluations at the CO level are affected by the availability of financial resources, the leadership of the Representative, and donor requests. The Evaluation Policy also ensures evaluations are carried out for all programmes at least once per cycle.

Finding 10: COs that have not carried out evaluations in the past few years have cited lack of resources and lack of country-level demand for evaluation as reasons.

Finding 11: UNICEF COs with highly rated evaluations are characterised by strong planning and Terms of Reference, good national evaluation capacity, and an adherence to a quality assurance process that often includes RO involvement.

Finding 12: Ensuring adequate human and financial resources for monitoring and evaluation at the country level is necessary to strengthen the decentralised evaluation function; however, this has different implications across regions and COs.

Finding 13: Although the Evaluation Policy does not set out clear definitions of results, the GEROS review accounts for output, outcome and impact evaluations.

Finding 14: The overall demand for evaluation to serve accountability purposes has been increasing in recent years, due in part to an increased focus on results by donors.
Finding 15: UNICEF’s recent focus on evaluation quality through GEROS is an incentive for COs to improve evaluation quality.

Finding 16: UNICEF’s GEROS system is the primary tool for assessing the quality and credibility of evaluation reports. Results are mixed, with some improvements between 2009 and 2010.

Finding 17: The overall picture regarding compliance with good practices for evaluations carried out at the decentralised level is mixed.

Finding 18: The quality of the evaluator is recognised as one of the key indicators of quality evaluations.

Finding 19: M&E staff at UNICEF are confident of their skills in managing evaluations and performance monitoring, and to some extent in national M&E capacity development. They are less confident in evaluating emergency, peace-building and humanitarian situations.

Finding 20: Familiarity with ethical issues is low among M&E staff at UNICEF.

Finding 21: Systematic training for M&E staff has not yet been fully institutionalised within UNICEF.

Finding 22: UNICEF has demonstrated good oversight and leadership in Human Rights and Gender but a fairly low/very modest level of reporting on equity, human rights, gender equality and women’s empowerment.

Finding 23: Growing interest in equity evaluations at the country level shows the relevance of the universality principle at UNICEF.

Finding 24: UNICEF’s dissemination of evaluation finding shows a good level of commitment to information sharing, but there remains room for improvement.

Finding 25: Management responses are increasingly tracked and followed-up at UNICEF. There are, however, certain instances where accountability for management responses remains burdensome or unclear.

Finding 26: Communication of UNICEF’s evaluations in local languages and to a range of stakeholders is not well tracked or known.

Finding 27: UNICEF has well-established systems to ensure sharing of evaluation findings at the level of the Executive Board.

Finding 28: Efforts by the EO to improve humanitarian evaluations have resulted in a greater appreciation of their use in recent years. But there are some challenges in carrying out such evaluations in emergency situations within the decentralised structure.

Finding 29: Evaluations across MTSP focus areas, strategies and operational effectiveness are well balanced in the organisation.

Finding 30: Increased management attention to evaluation and strengthened organisational capacity in evaluation are mutually reinforcing. Progress is due in part to GEROS and the VISION dashboard.

Finding 31: M&E staff accountability is not well defined in the current Evaluation Policy, affecting evaluation independence and quality.
Finding 32: The steps to ensure transparency of evaluations at all stages are poorly defined in the Evaluation Policy.
Appendix II  Terms of Reference

ToR for external consultancy
Status of the evaluation function, including current use of evaluation results, and current and emerging needs and demands
Evaluation Office
August 2012

Introduction

UNICEF is currently undertaking a review of its Evaluation Policy and of the evaluation function. This comes at a time of major changes within UNICEF and the organisation’s external operating environment, and the review is expected to lead to preparation of an updated evaluation policy. Further details of the policy review appear in Annex 1.

A number of studies of relevance to the review have already been undertaken, including a wide-ranging survey of M&E staff in the field (see Annex 2). To complement these, UNICEF now requires a systematic assessment of the demand for evaluation and the utilisation of evaluation results within UNICEF. Demand and utilisation are key aspects of the evaluation function, and analysis of these will contribute towards a comprehensive assessment of the current status of the evaluation function in UNICEF.

Purpose and objectives of this consultancy

The purpose of this consultancy is to assess the demand for evaluation and the utilisation of evaluation results and, making use of existing documentation, to make an overall assessment of the current status of the evaluation function in UNICEF. The consultancy will review arrangements across UNICEF for the governance and management of evaluation activities within UNICEF, including the role of management in addressing evaluation needs and demands, in ensuring effective utilisation of evaluation results and in taking account of cross cutting issues. The study will also assess emerging and changing evaluation needs and demands within UNICEF and beyond, taking account of a rapidly changing internal and external operating environment. Consideration will also be given to recent trends and developments in evaluation approaches and methods, including Equity-focused evaluations. Findings and conclusions will be summarised in a report for wide consultation across UNICEF, as well as for technical peer review and validation.

The assessment will draw on the UNEG peer review framework, which is itself based on UNEG norms and standards with a particular focus on independence, credibility and utility.

The assessment will focus on demand for evaluation and utilization of evaluation results. It will be important to distinguish the various needs and demands of different groups within UNICEF and of stakeholders beyond the organisation, to identify the strength of demand in various contexts (including its absence), and to identify the various factors driving or constraining demand. Assessment should be made of how needs and demands are considered in the planning, management and implementation of evaluation activities.

Similarly, the assessment will analyse the various uses made of evaluation results in different settings, will consider utilization gaps or shortcomings, and will identify factors driving or constraining use. It will be important to consider the quality and reliability of evaluation results and how these factors affect utilization.
Evaluation quality and perceptions of quality are important dimensions for consideration. Analysis should be made of any linkages between the demand for evaluation, the quality of evaluation processes and results, and the eventual utilisation of evaluation results. The demand for evaluation may be affected by various perceptions among management, staff and stakeholders of how evaluation is conducted and what it can provide. Similarly, varying perceptions of evaluation processes and results will affect the way in which evaluation results are used.

The assessment will be made within the broad context of the evaluation function, which includes the following:

- governance, roles and responsibilities;
- credibility, transparency and independence;
- planning, coordination and harmonization, including UN coherence;
- allocation and management of human and financial resources, including internal capacity building;
- evaluation partnerships and joint activities;
- evaluation capacity building, including support for national evaluation capacity development;
- quality of evaluation products and processes;
- use of evaluations, including not only the formal Management Response mechanism but also informal learning and follow up;
- communication, knowledge management and learning.

On the basis of the review, and drawing on existing documentation, the consultant will prepare an overall assessment of the current status of the evaluation function in UNICEF. This will include identification of any gaps, anomalies and issues in the current evaluation policy and function, as well as noting good practices and opportunities for improvement.

**Methods/process, and deliverables**

The consultant will collect information through a variety of sources, including existing material (see Annex 2). In particular, the consultant will:

- Conduct preliminary analysis of evaluations for which quality assessment has been completed in the past 3 years, in order to select a sample including evaluations both of good quality and of poor quality, with a view to analysing factors driving quality, and assessing linkages between evaluation quality, demand for evaluation, and eventual use of evaluation results. The analysis should also identify COs which have not produced any evaluations in recent years.

- On the basis of the preliminary analysis, design and conduct 10-12 case studies of evaluation activities in selected COs to understand key factors enabling or constraining good performance in terms of demand and planning for evaluation, production of good quality reports and use of evaluation results. The set of COs will be illustrative rather than representative but should ensure coverage of COs producing good, bad and no evaluations respectively in the last 3 years. Case studies should as far as possible be selected to cover all UNICEF regions; to include big, medium and small size offices; cover a range of development contexts (LDC, MIC); and should include at least 2 COs addressing significant humanitarian issues. The case studies will include desk review, a short questionnaire and phone interviews. No travel is foreseen. A report analysing the findings of the case studies will be produced.
Design and conduct a short survey (about 12-15 questions) to be administered to all COs, ROs and HQ divisions, focusing on the demand and use of evaluation by senior managers in HQ, ROs and CO, including emerging evaluation needs and demands, and addressing issues of evaluation quality. This will be complemented by a 22 key informant interviews (12 from the case studies, plus 10 from selected ROs, Board Members and HQ Divisions). Regional M&E Chiefs should be consulted. A report analysing the findings of the survey will be produced.

Based on the above analysis and reports, and drawing on existing documentation, the consultant will produce a draft synthesis report on the overall status of the evaluation function in UNICEF, to be presented to key stakeholders. The consultants will then finalise the report based on comments/feedback received.

**Timeline and deliverables**

- Inception report 2 weeks after signature of contract
- Case studies draft report after 3 weeks of identification of selected countries, and final one after two weeks from reception of comments
- Draft Demand and utilization survey after 3 weeks of acceptance of inception report, and final one after two weeks from reception of comments
- Draft final synthesis report on the status of the evaluation function and PPT with key findings/reccs 1 week after finalization of the above two reports
- Final report on the status of the evaluation function and PPT with key findings/reccs two weeks after reception of comments

<table>
<thead>
<tr>
<th>Deliverables</th>
<th>Wk1</th>
<th>Wk2</th>
<th>Wk3</th>
<th>Wk4</th>
<th>Wk5</th>
<th>Wk6</th>
<th>Wk7</th>
<th>Wk8</th>
<th>Wk9</th>
<th>Wk10</th>
<th>Wk11</th>
<th>Wk12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inception report after signature of contract</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case studies draft report after identification of selected countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reception of comments of Case studies draft report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case studies final report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Draft Demand &amp; utilization survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reception of comments of Demand &amp; utilization survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demand &amp; utilization survey final report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Draft synthesis report on the status of the evaluation function &amp; PPT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reception of comments of synthesis report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Qualifications

- Excellent and proven knowledge and experience in designing and implementing complex organisational reviews
- Proven knowledge and experience with decentralised evaluation systems in big organisations, preferably UNICEF or UN agency ones
- Excellent analytical and writing skills in English required.
Appendix III Framework for Analysis

Framework for Analysis

The following section sets out the framework for the analysis and justification for the survey and the key informant interviews. The overarching questions highlighted below have been formulated in an appropriate way in the survey questionnaire (see Appendix IV) and interview protocols presented in Appendices V to VIII.

Exhibit 4.1 Framework for Analysis

<table>
<thead>
<tr>
<th>Key Criteria</th>
<th>Key Questions</th>
<th>Sub Questions</th>
<th>Sources</th>
<th>Reference Documents and Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand for evaluation(^{30})</td>
<td>What are the key factors determining demand for evaluation?</td>
<td>What types of evaluations are demanded, in terms of relative importance?</td>
<td>Survey, Case studies, interviews</td>
<td>UNICEF Planning documents such as the IMEP and IMEF define the agreed-to evaluations to be carried out.</td>
</tr>
<tr>
<td>UNICEF demand</td>
<td></td>
<td>Which managers are demanding evaluations and why?</td>
<td>Survey, Interviews</td>
<td>Documents such as the PPP and MTSP define the demand for evaluations at the institutional level.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Which focus areas are demanding evaluations and why?</td>
<td>Survey</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>What other factors are determinant for evaluations?</td>
<td>Interviews</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>What are other ways in which Senior Managers are collecting evaluative data</td>
<td>Case studies</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>for accountability or decision making?</td>
<td>Case studies</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>How are UNICEF demands for evaluations prioritised?</td>
<td>Case studies, survey</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>What are emerging needs in evaluation (e.g. joint evaluations, etc.)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{29}\) The UNEG Framework for Professional Peer Reviews of the Evaluation Function of UN Organisations (2011) has been reviewed and informs this framework for analysis. While the review is not using a peer review model per se, much like the UNEG Framework, this review provides an assessment of key facets of the Evaluation Policy, including independence, credibility and utility. However, this review has a particular focus on the utility of evaluation results.

\(^{30}\) While the focus is on investigation existing demand for evaluation, weak demand or absence of demand will be investigated. Informants will be prompted to explain weak/inexistent demand when necessary. This will inform the case study design and will allow the evaluation team to document the linkages between demand and use.
<table>
<thead>
<tr>
<th>Key Criteria</th>
<th>Key Questions</th>
<th>Sub Questions</th>
<th>Sources</th>
<th>Reference Documents and Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of Evaluations&lt;sup&gt;31&lt;/sup&gt;</td>
<td>How do key UNICEF senior staff assess UNICEF's planning process for evaluations?</td>
<td>Do UNICEF planning tools (e.g. IMEP, IMEF) lead to better evaluation quality?</td>
<td>Case studies</td>
<td>UNICEF Planning documents such as the IMEP and IMEF define the agreed-to evaluations to be carried out.</td>
</tr>
<tr>
<td>Planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Governance, roles and responsibilities</td>
<td>How does governance affect quality of evaluations?</td>
<td>Are the key oversight functions leading to better evaluation quality?</td>
<td>Case studies, Survey,</td>
<td>Section IV of the UNICEF Evaluation Policy: Accountability for Evaluation in UNICEF sets out the key roles and responsibilities for the Evaluation function at CO, RO and HQ levels.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Does the evaluation function sufficiently involve key users throughout the evaluation process, leading to better evaluation quality?</td>
<td>Case studies</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Do such things as advisory committees improve the quality of evaluations?</td>
<td>Survey, Interviews</td>
<td></td>
</tr>
<tr>
<td>Management of human and financial resources for evaluation</td>
<td>Are sufficient resources being allocated to carry out quality evaluations?</td>
<td>Do Senior Managers feel that they have the financial resources necessary to carry out quality evaluations?</td>
<td>Survey</td>
<td>Section IV of the UNICEF Evaluation Policy: Accountability for Evaluation in UNICEF sets out the allocation and management of resources for evaluation. For example, UNICEF Country Representatives are responsible for ensuring that adequate UNICEF resources are dedicated to evaluation,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Do Senior Managers feel that they have the technical resources required to carry out quality evaluations?</td>
<td>Survey</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Do Senior Managers have the human resources necessary to carry out an oversight function of the evaluation?</td>
<td>Survey, Key informant</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>interviews</td>
<td></td>
</tr>
</tbody>
</table>

<sup>31</sup> Data collections tools differentiate quality of evaluation report from quality of evaluation management process. This is reflected by specific questions for each category.
<table>
<thead>
<tr>
<th>Key Criteria</th>
<th>Key Questions</th>
<th>Sub Questions</th>
<th>Sources</th>
<th>Reference Documents and Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support for national evaluation capacity development</td>
<td>To what extent does the evaluation function prioritise the development of national capacities for country-led evaluation systems?</td>
<td>What are the enabling factors and/or bottleneck for UNICEF to support national evaluation capacity development?</td>
<td>Survey, Case studies, Key informant interviews</td>
<td>In its Evaluation Policy, UNICEF states that it will seek to work with UNEG and other partners to establish and implement a common strategy for strengthening evaluation capacity in programme countries. The UNICEF MTSP (2006-2013) highlighted six main areas for strengthening evaluation during the MTSP period. The first of these is national capacity building and strengthening national leadership in country-level evaluation so that evaluation is “owned” by the country. In 2010, UNICEF developed guidance materials on national capacity development for country-led evaluation systems.</td>
</tr>
</tbody>
</table>

32 These include (a) National capacity-building and strengthened national leadership in country-level evaluations; (b) Strengthen evaluation within the United Nations system and with other partners; (c) Evaluation in humanitarian crises; (d) Evaluations related to MTSP focus areas, strategies and operational effectiveness; (e) Strengthen organisational capacity in evaluation; and (f) Heightened management attention to the evaluation function.
<table>
<thead>
<tr>
<th>Key Criteria</th>
<th>Key Questions</th>
<th>Sub Questions</th>
<th>Sources</th>
<th>Reference Documents and Definitions</th>
</tr>
</thead>
</table>
| Coordination and harmonization, including UN coherence; Evaluation partnerships and joint activities; | What are the coordination methods which are seen to favour quality evaluations?  
What types of partnerships and joint activities are perceived to favour quality evaluations? | Do joint evaluations lead to better evaluation quality?  
What are the key lessons learned from carrying out joint evaluation activities? | Survey, Key informant interviews, Case studies | Section VI. Evaluation with partners in the UNICEF Evaluation Policy: The increasing scope of global partnerships highlights the fact that successful interventions require joint ventures. Because of this, new approaches in evaluation need to be developed. UNICEF will need to seize the growing opportunities to participate in multi-partite evaluations, thereby drawing attention to results and impacts for children on the evaluation agendas of other organisations. There is evidence that joint evaluation tends to increase the independence and quality of evaluation, a further reason to promote evaluation with partners. |
| Quality of evaluation products and processes | What are the factors which determine the quality of UNICEF evaluations? | What are seen as the key conditions for good quality evaluations?  
What are seen as the key constraints of good quality evaluations?  
What processes have been seen to lead to better quality evaluations? | Survey, Key informant interviews, Case studies | The UNICEF Evaluation Policy and the UNICEF-Adapted UNEG Evaluation Reports Standards outline the key elements of the products and processes for UNICEF evaluation. UNICEF GEROS defines criteria for good quality evaluations within UNICEF. |
<table>
<thead>
<tr>
<th>Key Criteria</th>
<th>Key Questions</th>
<th>Sub Questions</th>
<th>Sources</th>
<th>Reference Documents and Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Use of Evaluations</strong>&lt;br&gt;General use of evaluations</td>
<td>What are the primary uses of UNICEF evaluations?</td>
<td>What are the primary uses of evaluations (Reporting results, decision-making, implementation, evidence-based policy advocacy, knowledge management? Are evaluation findings being used more or less by certain types of countries (LDC, LMI, and UMI)? Are evaluation findings being used more or less by certain types of offices? Are joint evaluations as useful as UNICEF-specific evaluations? Do government stakeholders use UNICEF’s evaluations? Do Senior Managers view evaluations as drivers of programme change?</td>
<td>Survey, Key informant interviews, Case studies, Case studies</td>
<td>Section B. <em>Guiding principles for the evaluation function</em> in the UNICEF Evaluation Policy sets out key criteria for how evaluations should be carried out. These are based on the UNEG Code of Conduct for Evaluation in the UN system</td>
</tr>
<tr>
<td>Use of evaluations, including formal Management Response mechanism and informal learning and follow up;</td>
<td>What UNICEF mechanisms and structures drive the use of evaluations?</td>
<td>Which managers are using evaluations and why? Which focus areas are using evaluations and why? What other factors are determinant for internal use of evaluations? Is the management response format adequate to monitor UNICEF’s use of evaluations?</td>
<td>Survey, Key informant interviews, Case studies</td>
<td>Several policies in UNICEF help to define evaluation use, notably the UNICEF Evaluation Policy Section A. <em>Purpose and use of evaluation</em> and the UNICEF-Adapted UNEG Evaluation Reports Standards. At the CO level, Representatives are responsible for following up on evaluation recommendations and reporting on the status of the follow-up.</td>
</tr>
<tr>
<td>Key Criteria</td>
<td>Key Questions</td>
<td>Sub Questions</td>
<td>Sources</td>
<td>Reference Documents and Definitions</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>Credibility(^{33}), transparency and independence(^{34});</td>
<td>How do senior managers perceive the credibility, transparency and independence of UNICEF evaluations?</td>
<td>Do senior managers view independence, transparency and credibility as important factors in evaluation quality and use?</td>
<td>Survey, Key informant interviews, Case studies</td>
<td>Section B. Guiding principles for the evaluation function in the UNICEF Evaluation Policy sets out key criteria for how evaluations should be carried out. These are based on the UNEG Code of Conduct for Evaluation in the UN system. Section A, covers the Purpose and use of evaluation and the UNICEF-Adapted UNEG Evaluation Reports Standards.</td>
</tr>
<tr>
<td>Communication, knowledge management and learning.</td>
<td>Does UNICEF communication/dissemination of its evaluations, affect use?</td>
<td>What is the relative importance of communication and dissemination compared to other factors affecting use of evaluations?</td>
<td>Survey, Key informant interviews, Case studies</td>
<td>The UNICEF Evaluation Policy and the UNICEF-Adapted UNEG Evaluation Reports Standards outline the key elements of the use of evaluations.</td>
</tr>
</tbody>
</table>

\(^{33}\) As per UNEG standards, credibility is defined through five core elements: structure logic and clarity; full description of objective evaluation; purpose objective and scope are explained; appropriate and sound methodology; findings, conclusions and recommendations are based on sound analysis

\(^{34}\) As per the GEROS Review Template, independence is defined as Implementation and control of the evaluation activities.
Appendix IV Survey Questionnaire

UNICEF Country Representatives on the Demand, Quality and Use of Evaluation

If you require further information concerning this survey or if you experience any technical difficulties, please contact Mr. Emmanuel Trépanier at etrepanier@universalia.com.


1. BACKGROUND

1.1 How long have you been a Country Representative in your current country office?
- Less than 6 months
- More than 6 months
- Please select if you are comfortable responding to this survey based on your current country office
- Please select if you were a Country Representative in another country office prior to this one and would prefer to respond based on your previous country office
- Please select if you believe that you have insufficient knowledge of the current country office context to respond to a survey on evaluation demand, quality and use

1.2 What country are you currently working in?
If you have worked in your current country office for less than six months, you may select the country office of your previous assignment.
- Afghanistan
- Albania
- Algeria
  ... 171 additional choices hidden ...
2. DEMAND FOR EVALUATIONS

The following questions are designed to help us get an understanding of the nature of requests or requirements for evaluations and the kinds of evaluations being requested in your programme country. For the purpose of this discussion, please consider the evaluations carried out within the past two years. Note that UNICEF uses evaluation for four broad purposes, namely: Reporting results (accountability); Decision-making (improving country/regional programming); Evidence-based policy advocacy (use evidence on what works for children to inform/influence public policies and programmes); Knowledge management (informing good practice).

2.1 Country offices carry out different numbers of evaluations each year. What is the approximate number of evaluations carried out in your office last year (2011)?

An evaluation is an exercise that attempts to determine as systematically and objectively as possible the worth or significance of an intervention, strategy or policy. This excludes surveys, research and studies although they may draw on similar methodologies.

- 1-3
- 3-5
- More than 5
- Don't Know/ Not Sure
- None

If you selected “None” please provide any details on the factors that led your office to not carry out evaluations last year.


2.2 How has the overall internal (i.e. within UNICEF) demand for evaluations changed over the last 2-3 years?

Please select the option that best suits your programme country.

- Demand for evaluations has increased
- Demand for evaluations has stayed the same
- Demand for evaluations has decreased
- Don’t know/ Not sure

Please provide any reasons for your answer above.
2.3 How has the overall external (e.g., from governments, donors, etc.) demand for evaluations changed over the last 2-3 years?

Please select the option that best suits your programme country.

- Demand for evaluations has increased
- Demand for evaluations has stayed the same
- Demand for evaluations has decreased
- Don’t know/ Not sure

Please provide any reasons for your answer above.

2.4 Please rate the following factors driving the demand for evaluations for your country office?

<table>
<thead>
<tr>
<th>Factor</th>
<th>1-Low Importance</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5-High Importance</th>
<th>Not Sure/ Don't Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of evaluations carried out in your country office in the previous years (tradition of evaluation)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation capacity (HR) of the country office</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leadership of Regional Office</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leadership of Evaluation Office</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Donor requests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Government requests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial resources in your country office for evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political context</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandatory requirements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leadership of Representative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## 2.5 In which direction do the following factors drive the demand for evaluations for your country office?

Please select the option from the dropdown menu that best suits your office.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Increases demand for evaluations</th>
<th>No effect on demand for evaluations</th>
<th>Decreases demand for evaluations</th>
<th>Don’t know/ Not sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of evaluations carried out in your country office in the previous years (tradition of evaluation)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Evaluation capacity (HR) of the country office</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Leadership of representative</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Leadership of Regional Office</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Donor requests</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Topic</td>
<td>Impact</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>---------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Government requests</td>
<td>Increases demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No effect on demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decreases demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Don’t know/ Not sure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial resources in your country office for evaluation</td>
<td>Increases demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No effect on demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decreases demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Don’t know/ Not sure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political context</td>
<td>Increases demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No effect on demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decreases demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Don’t know/ Not sure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandatory requirements</td>
<td>Increases demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No effect on demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decreases demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Don’t know/ Not sure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leadership of Evaluation Office</td>
<td>Increases demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No effect on demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decreases demand for evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Don’t know/ Not sure</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please describe any additional information on demand for evaluations in your country office:
2.6 Please rank the relative importance that you place on the following purposes for carrying out evaluations in your office?

(Rank 1=least important, 4=most important)

<table>
<thead>
<tr>
<th>Purpose</th>
<th>1-Least Important</th>
<th>2</th>
<th>3</th>
<th>4-Most Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting on results (accountability)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Decision-making (improving country/regional programming)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Evidence-based policy advocacy (use evidence on what works for children to inform/influence public policies and programmes)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Knowledge management (informing good practices)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

2.7 What do you see as the emerging needs for evaluation at UNICEF?

Please indicate your level of interest and demand in the following types of evaluations:

<table>
<thead>
<tr>
<th>Evaluation Type</th>
<th>1-Low Interest</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5-High Interest</th>
<th>Not Sure/ Don't Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity-focused evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Policy evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Impact evaluations</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Joint UN evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Joint evaluations with other partner</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Country led evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

2.8 Please highlight any other types of evaluations that you think will have greater importance in UNICEF’s future.


2.9 Please describe any factors in the political context which influence the demand for evaluations in your country.


3. QUALITY OF EVALUATIONS

The following questions are designed to help us get a better understanding of what factors affect the quality of evaluations at UNICEF.

3.1 Please rate the factors affecting the production of quality evaluations in your country office?

<table>
<thead>
<tr>
<th>Factor</th>
<th>1-Low Importance</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5-High Importance</th>
<th>Not Sure/Don't Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Internal office oversight of evaluation</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>B. Regional office oversight</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>C. Quality of the consultant(s) engaged in the evaluation</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>D. Country office leadership</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>E. Existence of an evaluation advisory group</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>F. Availability of an M&amp;E specialist/focal point in Country Office</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>G. Independence of the Evaluation</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>H. Other</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Other (Please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.2 Given that poor quality of evaluations remains an issue in decentralised evaluations in UNICEF, what measures would you suggest in order to improve quality of UNICEF evaluations?

3.3 What key measures would you suggest in order to improve the quality of government-led evaluations?
4. USE OF EVALUATION

The following questions are designed to help us get a better understanding of the utilisation of UNICEF’s evaluation results, including findings and recommendations.

4.1 Please indicate the overall level of use of evaluations in your country office for each of the following purposes:

<table>
<thead>
<tr>
<th>Purpose</th>
<th>1-Low Use</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5-High Use</th>
<th>Not Sure/Don’t Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting results (accountability)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decision-making (improving country programming)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evidence-based policy advocacy (use evidence on what works for children to inform/influence public policies and programmes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge management (informing good practices)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you selected “Low use” for any of the purposes above, please explain:

[Blank space]


4.2 In your opinion, what are the factors that affect the internal use of UNICEF evaluations results in your country office?

<table>
<thead>
<tr>
<th>Factor</th>
<th>1-Low Importance</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5-High Importance</th>
<th>Not Sure/ Don't Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who initiates/demands the evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Purpose of evaluation (accountability; decision-making; evidence-based policy-advocacy; knowledge management)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Senior management leadership</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Quality of evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Independence of evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Dissemination strategy</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Donor involvement</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Joint evaluations</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Establishment of reference or advisory group</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Timeliness of evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Development of a management response</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Other</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

4.3 In your opinion, what are the factors that affect the external use of UNICEF evaluations results by stakeholders (government, donors, partners)?

<table>
<thead>
<tr>
<th>Factor</th>
<th>1-Low Importance</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5-High Importance</th>
<th>Not Sure/ Don't Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Independence of evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Dissemination strategy</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Joint evaluations</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Establishment of reference or advisory group</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Timeliness of evaluation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Political context</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Other</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
Other (please specify)

4.4 Are UNICEF evaluation findings systematically disseminated externally?
○ Always
○ Sometimes
○ Never
○ Don't Know/ Not Sure

4.5 What is your perception on the extent to which the following groups use UNICEF evaluation findings and recommendations?

<table>
<thead>
<tr>
<th>Group</th>
<th>1-Low Use</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5-High Use</th>
<th>Not Sure/ Don't Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNICEF Staff</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other UN Agencies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Government</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Donors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civil Society</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.6 Please describe any factors in the political context, which influence the use of evaluations in your country?

5. CONCLUSION

5.1 Please describe any recommendations that you have in order to improve UNICEF’s overall evaluation function in the future?

The UNICEF Evaluation Office and Universalia wish to thank you for your participation in this survey.
Appendix V Case Studies Interview Protocol

For selected UNICEF Country Representatives, Deputy Representatives, Programme Officers and Evaluation Specialists

Introduction

The purpose of the UNICEF Evaluation Policy is to ensure that the organisation "has timely, strategically focused and objective information on the performance of its policies, programmes and initiatives to produce better results for children and women. The policy aims to foster a common institutional understanding of the evaluation function in UNICEF, and further strengthen evidence-based decision-making and advocacy, transparency, coherence and effectiveness."\(^{36}\)

UNICEF is currently undertaking a review of its Evaluation Policy and the Evaluation function. This review is expected to lead to the preparation of an updated Evaluation Policy. The objectives of this review are to assess the demand for evaluation and the utilization of evaluation results. You have been contacted to complement data collected through document review and surveys. For the purpose of this discussion, please consider that UNICEF uses evaluation for four broad purposes, namely:

- Reporting results (accountability);
- Decision-making (improving country/regional programming);
- Evidence-based policy advocacy (use evidence on what works for children to inform/influence public policies and programmes);
- Knowledge management (good practices and lessons learned).

Please reflect on the last 3-4 evaluations carried out within your UNICEF country office, or the last evaluation that you personally worked on.

Demand for evaluations

1) How did the demand arise for these evaluations (requirement for accountability, decision making, advocacy, knowledge generation)?

2) Were you clear on their intended purposes? Do you feel that staff/other stakeholders were clear on the intended purpose of the evaluations? If not, why not?

3) Of the four purposes mentioned, what are the purposes which are in greatest demand within your country office?

4) What are the key elements which drive evaluation requests in your office (programmes, leadership, history of good quality, government, other)?

5) What elements from the UNICEF organisational culture contribute to generating (or not generating) a demand for evaluations?

6) Through what processes are requests for evaluations prioritised within your country office (IMEP, joint UN planning, government, donor requests, other)?

---

\(^{35}\) While these are preferred respondents, in some cases others may be interviewed. It may not be feasible to ask all questions to all the respondents. Responses will remain confidential according to UNEG guidelines.

7) Are there any unique requests for UNICEF evaluations in your particular country? Are there certain internal actors or external stakeholders demanding evaluations? (if Humanitarian Country) Do you have specific requests for humanitarian evaluations?

8) In general, what are the key factors inhibiting the demand for evaluations?

Quality of evaluations

1) Can you briefly outline the process through which the evaluation(s) were conducted?

2) What do you consider to be the key characteristics of good quality evaluations?

3) What do you consider to be the key internal processes which ensure the quality of the evaluations being conducted (prompt for oversight, governance, leadership, quality of consultant, advisory group, others?)

4) Do you have any suggestions on how the process of UNICEF evaluations could be strengthened in order to improve the quality of evaluations?

5) How does working with a) partner organisations and b) government affect (positively or negatively) the quality of evaluations?

Use of evaluations

1) How have the evaluation findings and recommendations been communicated and disseminated inside and outside of the agency?

2) In what way does the environment surrounding UNICEF enable or hinder the use evaluations?

3) For what specific purpose(s) have evaluation results been used (both internally and externally)?

4) In your experience, what are the key factors that lead to an evaluation being used by senior management?

5) If there was a management response to these evaluations, was it instrumental to increase the use of evaluation findings and implementation of evaluation recommendation?

Conclusions

1) What can you or your office do better in terms of the way in which evaluations are used?

2) Do you have any other comments related to the demand, use or quality of the UNICEF evaluation products?
Appendix VI Case Studies Interview Protocol (Countries with No Evaluations Conducted)

For selected UNICEF Country Representatives, Deputy Representatives, Programme Officers and Evaluation Specialists where formal evaluations have not been shared.

Introduction

The purpose of the UNICEF Evaluation Policy is to ensure that the organisation "has timely, strategically focused and objective information on the performance of its policies, programmes and initiatives to produce better results for children and women. The policy aims to foster a common institutional understanding of the evaluation function in UNICEF, and further strengthen evidence-based decision-making and advocacy, transparency, coherence and effectiveness." UNICEF is currently undertaking a review of its Evaluation Policy and the Evaluation function. This review is expected to lead to the preparation of an updated Evaluation Policy. The objectives of this review are to assess the demand for evaluation and the utilization of evaluation results. You have been contacted to complement data collected through document review and surveys. For the purpose of this discussion, please consider that UNICEF uses evaluation for four broad purposes, namely:

- Reporting results (accountability);
- Decision-making (improving country/regional programming);
- Evidence-based policy advocacy (use evidence on what works for children to inform/influence public policies and programmes);
- Knowledge management (good practices and lessons learned).

Based on data contained on GEROS and Annual Reports, we understand that your UNICEF country office has not conducted any formal evaluations in the last three years. The questions below have been designed to better understand the specific evaluation context in your country office. If your country office has completed evaluations, please let us know and a different protocol will be used for this discussion.

Demand for evaluations

1) Despite not carrying out evaluations, has there been a strong demand for evaluations in your country office (as a requirement for accountability, decision making, advocacy, knowledge generation)?

2) If so, what has prevented your country office from conducting these evaluations (resources, staffing, expertise,)?

3) If not, what do you think are the main reasons explaining the absence of a demand for evaluations?

4) Are there any factors related to the involvement of external stakeholders (partner organisations, governments, etc.) that have prevented your country office from conducting evaluations?

37 While these are preferred respondents, in some cases others may be interviewed. It may not be feasible to ask all questions to all the respondents. Responses will remain confidential according to UNEG guidelines.

5) Are there factors related to the quality of evaluations in the past that has prevented your country office from conducting evaluations?

6) If not through evaluations, please describe the primary means through which your office reports on results, demonstrates accountability to donors, and makes decisions, for evidence-based policy advocacy and/or for knowledge management purposes? (e.g., surveys, reviews, studies, nationally led systems or studies.)

7) What are the factors involved which has led your country office to assess programs in the manner which you describe above? Do you see that other UN agencies have followed suit in this manner?

Quality of evaluations

1) Can you briefly outline the process through which these other measurement tools (e.g., assessments, reviews, studies, etc.) have been carried out to ensure quality?

Use of evaluations

1) How have the findings and recommendations of other types of assessments conducted by your country office been communicated and disseminated inside and outside of the agency?

2) For what purpose(s) have the results of these other types of assessments been used internally and externally?

3) If they have not been used, are there any factors related to the potential use of evaluations (e.g., uptake by senior management, political environment, needs expressed by partners, etc.) that have prevented your country office from using evaluations?

Conclusions

1) What can you or your office do better to ensure that evaluations are conducted and used in the future?

2) Do you have any other comments related to the demand, use or quality of the UNICEF evaluation products?
Appendix VII Key Informants Interview Protocol

For Regional Directors, Regional M&E staff, and HQ Directors

Introduction

The purpose of the UNICEF Evaluation Policy is to ensure that the organisation "has timely, strategically focused and objective information on the performance of its policies, programmes and initiatives to produce better results for children and women. The policy aims to foster a common institutional understanding of the evaluation function in UNICEF, and further strengthen evidence-based decision-making and advocacy, transparency, coherence and effectiveness." UNICEF is currently undertaking a review of its Evaluation Policy and the Evaluation function. This review is expected to lead to the preparation of an updated Evaluation Policy. The objectives of this review are to assess the demand for evaluation and the utilization of evaluation results. You have been contacted to complement data collected through document review and surveys. For the purpose of this discussion, please consider that UNICEF uses evaluation for four broad purposes, namely:

- Reporting results (accountability);
- Decision-making (improving country/regional programming);
- Evidence-based policy advocacy (use evidence on what works for children to inform/influence public policies and programmes);
- Knowledge management (good practices and lessons learned).

Please reflect on the last 2 evaluations carried out by your office.

Demand for evaluations

1) How did the demand arise for these evaluations (requirement for accountability, decision making, advocacy, knowledge generation, other)? Were you clear on their intended purposes? Do you feel that staff/other stakeholders were clear on the intended purpose of the evaluations? If not, why not?

2) Of the four purposes of uses of evaluations mentioned, what are the purposes which are in greatest demand?

3) What are the key elements which drive evaluation requests in your office? (programmes, focus areas, leadership, history of good quality evaluations, partners)

4) What elements from the UNICEF organisational culture contribute to generating a demand for evaluations?

5) Through what process are requests for evaluations prioritised within your office (IMEF, annual planning, joint UN planning, government, donor requests, other)? Any comments on prioritization?

39 Responses will remain confidential according to UNEG guidelines.

6) Are there any unique interests for UNICEF evaluations in your particular region/programme area? In other words, are there certain internal actors or external stakeholders demanding evaluations? What are those interests?

7) What do you see as the emerging needs for evaluation at UNICEF (equity focussed evaluation, joint UN evaluation, joint evaluations with other partner, government led evaluation)?

8) In general, what are the key factors inhibiting the demand for evaluations?

**Quality of evaluations**

1) Can you briefly outline the process through which evaluations were conducted?

2) What do you consider to be the key factors to ensure good quality evaluations?

3) Regional office only: what are the key elements which drive evaluation quality in your region? (programmes/focus areas/leadership, history of good quality evaluations, partners)?

4) a) Regional office only: Are there any particularities in your region that affect the way evaluations are carried out? (i.e. partner organisations, governments, quality of consultants)

   b) Headquarters only: Are there any particularities in your division that affects the way evaluations are carried out (i.e. Challenge in data collection, partner organisations, quality of consultants)?

5) What do you consider to be the key internal processes which ensure the quality of the evaluations being conducted (prompt for oversight, governance, leadership, quality of consultant, advisory group)?

6) How does working with a) partner organisations and b) government affect (positively or negatively) the quality of evaluations?

**Use of evaluations**

1) For what specific purpose(s) have evaluation results been used (both internally and externally)?

2) In your experience, what are the key factors that lead to an evaluation being used by senior management (leadership, quality, previous history of quality)?

3) In what way does the political environment surrounding UNICEF enables or hinders the use evaluations?

4) a) Headquarters only: Are there any particularities in the way in which evaluation results are used in your division?

   b) Regional Office only: Are there any particularities by programme/focus area in the way in which evaluation results are used in your region?

5) Regional Office only: what are the key factors that lead to an evaluation being used by senior management in the country offices in your region (leadership, quality, previous history of quality)?

6) What can you or your office do better in terms of the way in which evaluations are used?

**Conclusion**

1) Do you have any other comments related to the demand, use or quality of the UNICEF evaluation products?
Appendix VIII Key Informants Interview Protocol for UNICEF Board Members

Introduction

The purpose of the UNICEF Evaluation Policy is to ensure that the organisation "has timely, strategically focused and objective information on the performance of its policies, programmes and initiatives to produce better results for children and women. The policy aims to foster a common institutional understanding of the evaluation function in UNICEF, and further strengthen evidence-based decision-making and advocacy, transparency, coherence and effectiveness."

UNICEF is currently undertaking a review of its Evaluation Policy and the Evaluation function. This review is expected to lead to the preparation of an updated Evaluation Policy to be presented at the Board meeting in June 2013. The objectives of this review are to assess the demand for evaluation and the utilization of evaluation results. You have been contacted to complement data collected from UNICEF senior managers and through document review and surveys. For the purpose of this discussion, please consider that UNICEF uses evaluation for four broad purposes, namely:

- Reporting results (accountability);
- Decision-making (improving country/regional programming);
- Evidence-based policy advocacy (use evidence on what works for children to inform/influence public policies and programmes);
- Knowledge management (good practices and lessons learned).

Demand for evaluations

1) Do you feel that UNICEF as an organisation carries out a sufficient number of evaluations to sufficiently respond to the purposes outlined – such as reporting on results, decision-making, evidenced-based policy advocacy and knowledge management?

2) Is there a certain purpose of evaluations for which you think UNICEF should focus greater attention and/or resources?

3) What are the factors underlying demand for evaluations that UNICEF needs to consider?

4) In the future, how do you see demand for evaluations evolving for agencies such as UNICEF (i.e. greater need for certain types of evaluations; Executive Board requesting specific corporate evaluations)?

Quality of evaluations

1) UNICEF Evaluation office currently carries out an assessment of the quality of evaluations (GEROS) the highlights of which it shares with the board in its Annual Report on the evaluation function. Do you find it useful, or are there ways it can be improved? How do you currently use this information?

2) What do you consider to be the key factors that UNICEF has been doing well to improve the quality of its evaluations in recent years? What does it need to do better in the coming years, in your opinion?

---

41 Responses will remain confidential according to UNEG guidelines.

3) At the Executive board level, how do you see UNICEF’s evaluation quality vis-à-vis the quality of other agencies or development organisations?

**Use of evaluations**

1) For what specific purpose(s) have evaluation results been used by Board members?

2) Please comment on the Executive Board’s involvement in the discussion of recommendations of the corporate evaluations managed by UNICEF.

3) Do you think UNICEF should make even more efforts to strengthen National Evaluation capacities and systems, so that more and more country-led evaluations could be used?

4) In what way does the Executive Board enable the use of evaluations? Please comment on how UNICEF as an agency has done in terms of using its evaluations and its management responses? How does this compare with other multilateral agencies?

5) In Board meetings, do you sense that Evaluations drive programme change within UNICEF? Why or why not? What else do you see which drives programme change in the organisation?

**Conclusion**

1) Do you have any other comments related to the demand, use or quality of the UNICEF evaluation products?