



## Economic and Social Council

Distr.: General  
19 April 2013

Original: English

**For action**

---

### United Nations Children's Fund

Executive Board

**Annual session 2013**

18-21 June 2013

Item 8 of the provisional agenda\*

### **Annual report on the evaluation function and major evaluations**

#### *Summary*

The present annual report has been prepared in accordance with the Evaluation Policy (E/ICEF/2008/4) and relevant Executive Board decisions. The report provides information on global developments in evaluation and the current state of the evaluation function in UNICEF, including progress in strengthening the decentralized evaluation function.

A draft decision is included in section VIII.

---

\* E/ICEF/2013/10.



## **I. Introduction**

1. Evaluation in UNICEF serves the organization's goals of protecting and promoting children's rights, addressing children's needs and helping children to achieve their full potential in the course of regular development as well as during times of crisis. It promotes organizational learning, accountability and transparency, with a view to strengthening performance and delivering better outcomes for children. It supports decision-making and advocacy by providing reliable evidence on the implementation and results of policies, programmes and initiatives.

2. This report describes the performance of the UNICEF evaluation function in 2012. In recent years, the annual report on evaluation also included commentary on major evaluations completed in the course of the year. The Executive Board, wishing to see deeper thematic analysis of recent evaluations, has requested a separate report on this aspect. Accordingly, while the present report outlines the status of the evaluation function in 2012, a companion report (E/ICEF/2013/15) presents a thematic synthesis of recent evaluations of UNICEF humanitarian action.

3. The evaluation function in UNICEF operates under the Evaluation Policy approved by the Executive Board in 2008. Recent years have seen transformation within UNICEF and across the wider development landscape. Further changes are in prospect, both within UNICEF, where a new medium-term strategic plan is under preparation, and beyond, as work begins on formulating sustainable development goals to guide future development efforts as well as transformation of the humanitarian agenda to ensure more robust response. Within this shifting environment, the Executive Board has requested UNICEF to update the Evaluation Policy. The revised Evaluation Policy (E/ICEF/2013/14) has been submitted to the Executive Board for consideration.

4. This present report covers key elements of the evaluation function at UNICEF. It outlines efforts made by UNICEF to support evaluation efforts at the global level, especially inter-agency activities within the United Nations system; focuses on the performance of the evaluation function within UNICEF itself, as measured against a set of key indicators; and presents observations on the human and financial resources allocated to evaluation. It then outlines measures taken to strengthen the evaluation function within UNICEF; and reports on efforts to strengthen evaluation capacity at national and global levels. Following the concluding remarks, it presents a draft decision for consideration by the Executive Board.

## **II. Support to evaluation actions at global level**

5. UNICEF undertakes significant evaluation activity at the global level, much of it in support of evaluation within the wider United Nations system. These activities fall mainly into three areas: (a) working groups and meetings that develop, disseminate and harmonize technical knowledge and good practices; (b) inter-agency and joint evaluations and related initiatives; and (c) overall leadership and governance processes. Much of this work is conducted through or in coordination with the United Nations Evaluation Group (UNEG), as well as through the Inter-Agency Standing Committee (IASC). This section outlines the significant efforts made at the global level and the implications for future action.

### **Harmonization of technical knowledge and good practice**

6. From its inception, UNEG has organized inter-agency technical working groups to develop technical and policy guidance and, cumulatively, a large body of work has been completed. The past year saw progress in several working groups in which UNICEF takes a leading role:

(a) The task force on joint evaluation collaborated with the United Nations Development Operations Coordination Office (DOCO) and launched guidance on the preparation of terms of reference for United Nations Development Assistance Framework (UNDAF) evaluations and of management responses to UNDAF evaluations. The task force is developing a resource pack on planning and managing joint evaluations — an area of increasing workload and complexity. The goal is to improve the quality and use of UNDAF evaluations, which have been widely recognized to lag agency-specific evaluations in quality and uptake.

(b) The task force on national evaluation capacity development compiled a brochure, “Practical tips on how to strengthen national evaluation systems”, to guide both technical and non-technical staff. In addition, UNICEF and other United Nations agencies, together with UNEG, are supporting a major partnership initiative (“EvalPartners”) which aims to strengthen civil society’s evaluation capacities (see Section 6 for further details).

(c) The task force on human rights and gender equality finalized guidance on integrating human rights and gender equality dimensions in evaluations, filling a gap in United Nations evaluative practice that UNEG members had identified.

(d) The task force on impact evaluation is preparing technical guidance for impact evaluation, addressing particularly normative work and institutional support within the United Nations. UNEG convened the annual meeting of the Network of Networks on Impact Evaluation, which provided an opportunity to discuss latest developments with evaluation partners from the Development Assistance Committee of the Organization for Economic Cooperation and Development (OECD-DAC) and international financial institutions.

7. Despite the success of most individual task forces, some are inactive while several require wider sharing of effort across members. The utilization and effectiveness of UNEG products has not been systematically reviewed. For these reasons, UNEG has commissioned an external assessment of its work, to provide an informed basis for reflection and action in 2013. UNICEF supports this moment of reflection, with the expectation that it will refocus and refresh the work of the group.

### **Inter-agency and joint evaluations and related activities**

8. UNICEF was engaged in twice the usual number of joint evaluations and related activities in 2012, equally with regard to development activities and humanitarian action. Joint evaluations included the following:

(a) *Evaluation of the United Nations Girls’ Education Initiative (UNGEI)*. UNICEF and the UNGEI cooperating agencies completed this important evaluation, which aimed to review UNGEI’s governance structure and administrative components and identify the key elements that support strong partnerships for girls’ education. Country case studies were conducted in Egypt, Nepal, Nigeria and

Uganda, and a regional case study reviewed the East Asia-Pacific partnership. A joint management response was issued by the Global Advisory Committee on behalf of all UNGEI partners. The evaluation and management response, taken together, informed the new strategic direction for UNGEI, agreed at a global meeting hosted by Uganda. UNICEF added its own management response, addressing the recommendations related to its role as the lead agency and head of the UNGEI secretariat.

(b) *Joint evaluation of joint gender programmes in the United Nations system.* This joint initiative, led by the United Nations Entity for Gender Equality and the Empowerment of Women (UN-Women), involved UNICEF, United Nations Development Programme (UNDP), United Nations Population Fund (UNFPA), the Millennium Development Goals Achievement Fund, and the Governments of Norway and Spain. The evaluation is examining the value of joint gender programmes in improving United Nations system coherence, to determine how far this has enhanced the achievement of results on gender equality and women's empowerment. Launched in 2012, the evaluation will conclude in mid-2013.

(c) *Joint evaluation of the UNFPA-UNICEF joint programme on female genital mutilation/cutting (FGM/C).* Managed jointly by UNICEF and UNFPA, this evaluation is engaging with major national counterparts to assess how far and under what circumstances the joint programme has accelerated the abandonment of FGM/C over the period 2008-2012. Due to conclude in July 2103, the evaluation focuses on 13 countries where the joint programme is implemented. In-depth case studies have been undertaken in four countries (Burkina Faso, Kenya, Senegal and Sudan).

(d) *Inter-agency study on the monitoring and reporting mechanism on grave violations against children in situations of armed conflict.* Led by UNICEF, in collaboration with United Nations Department of Peacekeeping Operations and the Office of the Special Representative of the Secretary-General on Children and Armed Conflict, the study will fill a knowledge gap on good practices related to the monitoring and reporting mechanism.

9. Inter-agency joint evaluations are usually more complex to manage and take longer than those carried out by a single agency, but UNICEF clients generally judge the results to be worth the investment. The quality of these carefully prepared and well-resourced evaluations has been consistently good. In contrast, the quality of inter-agency joint evaluations of responses to humanitarian emergencies has been mixed, despite considerable efforts by participating agencies.<sup>1</sup> UNICEF has been a key contributor to inter-agency efforts to revise various joint evaluation methodologies in humanitarian situations, including real-time inter-agency evaluations, with an eye toward improving quality and utility. Engagement in joint evaluations — “evaluating as one” — is an important mechanism for undertaking comprehensive evaluation of joint programmes, especially in the absence of effective arrangements for United Nations system-wide evaluation. However, with limited evaluation capacity across the United Nations, joint evaluation priorities need to be judiciously balanced with requirements for internal evaluation within the respective agencies.

---

<sup>1</sup> See E/ICEF/2013/15.

### Overall leadership and governance processes

10. Evaluation leadership in inter-agency matters across the United Nations is fractured and governance arrangements are weak; these were among the conclusions reached by the review of arrangements for system-wide evaluation commissioned by the Deputy Secretary-General in 2012.<sup>2</sup> UNEG, as a professional association, has no formal responsibilities within the United Nations system to direct or manage evaluation activities, although it has provided a useful forum for coordination, consultation and sharing of views among evaluation specialists within the United Nations.

11. *Independent Evaluation of Delivering as One.* UNEG provided advisory support to this important study on *Delivering as One*, which informed Member States on progress with the “One United Nations” agenda. UNEG also provided a channel for advice on evaluation-related issues considered in the course of the quadrennial comprehensive policy review of operational activities for development within the United Nations system, including the above-mentioned review of arrangements for United Nations system-wide evaluation.

12. UNEG has served as a useful sounding board and source of technical advice; it could provide a substantial platform for advice and advocacy on evaluation issues of concern to the United Nations. Pressing issues include, for example, the need for increased efficiency, effectiveness and value for money, where evaluation can provide relevant input, and the question of building monitoring and evaluation capacity in the Resident Coordinator’s office and more widely across United Nations country teams to support adequate assessment of United Nations performance and results at the country level. Through UNEG, UNICEF will work with other members to provide a clearer and more powerful professional voice on issues concerning evaluation.

## III. The evaluation function in UNICEF: performance and results

### Introduction and overview

13. UNICEF first reported to the Executive Board on the performance of the evaluation function in 2006, and has since done so annually. As information systems have been refined, more elements have been added.

14. A suite of six clusters of key performance indicators provides a general overview of performance. Enough data is now available to offer some analysis of trends.<sup>3</sup> Based on trend data, four major conclusions are warranted:

(a) The number of evaluations being conducted and the topical coverage have fallen below minimally adequate levels in recent years; these numbers need to increase;

(b) Impact and summative evaluations are declining relative to evaluation effort overall. While this may be a temporary trend linked to new programme

<sup>2</sup> Angela Bester and Charles Lusthaus, *Comprehensive review of the existing institutional framework for the system-wide evaluation of operational activities for development of the United Nations system. Final Report*, March 2012.

<sup>3</sup> Trends can in most cases be identified over the past three years. For some issues, reliable trend data goes back a decade.

initiatives, it would be of concern if it signals that UNICEF has lost sight of outcomes and impact;

(c) Evaluation quality has shown only limited improvement since the mid-2000s;

(d) UNICEF has successfully instituted a culture in which management responses to evaluations are routinely provided.

**Indicator 1: Coverage of evaluations managed, and submission rate to the Global Evaluation Database**

15. Until a more effective oversight system was introduced four years ago, UNICEF offices claimed to be conducting as many as 250 evaluations each year. Reports were submitted for little more than half of this number, and many of these arrived late. Management responses were seldom prepared. Consequently, in setting up the Global Evaluation Reports Oversight System (GEROS), the Evaluation Office emphasized better training on evaluation principles, timely submission of evaluation reports and management responses to central databases, and publication of all evaluation reports on the UNICEF external website. GEROS provided for timely assessment and rating of reports, feedback on the quality of reports to offices, and compilation of performance information into a management information dashboard.

16. In key respects, these internal capacity development and oversight measures have been successful. The submission rate of evaluation reports to the Global Evaluation Database reached 98 per cent in 2012 due to close cooperation between the Evaluation Office, regional offices and country teams. Further, all reports and quality assessments are freely available to the public through the UNICEF website and can be easily searched by various criteria, such as theme, country, region or date.<sup>4</sup>

17. However, the number of evaluations being submitted has dropped from 140 in 2010 to 99 in 2011. It is regrettable that 38 country offices, including several managing large programmes, appear to have conducted no evaluations whatsoever over the period 2009-2011. Others conducted just one or two, despite the requirement in the 2008 Evaluation Policy that all programme components should be evaluated at least once in the programme cycle. Moreover, while its humanitarian response has grown in terms of both numbers and resources, UNICEF has conducted evaluations of only 34 of the 1,025 emergencies to which it responded from 2008 to 2011. Importantly, the vast majority of these evaluations have been of Level 2 and Level 3 emergencies, leaving Level 1 emergencies — by far the most common humanitarian situation UNICEF faces — largely unexamined. It is possible that a number of evaluations were carried out but not submitted to the central database. However, if the GEROS numbers are accurate, it appears that many country offices, and the organization as a whole, are missing the opportunity to generate critical knowledge for management, learning and accountability. A new performance indicator on evaluation coverage is needed, along with means to verify the accuracy of the information collected against this indicator and to ascertain the reasons for poor evaluation coverage, where this persists.

---

<sup>4</sup> <http://www.unicef.org/evaldatabase/index.html>.

**Indicator 2: Topical distribution**

18. The 2011 results show evaluation coverage across all areas of strategic focus. The figures for the topical distribution of evaluations are set out in table 1 below.

19. Over time, the numbers of evaluations of UNICEF regular programmes in each sector or focus area can vary significantly. The most notable trend has been the doubling of the proportion of multi-sector evaluations, from 12 per cent in 2009 to 25 per cent in 2011. This is likely to be a by-product of several changes: the move toward holistic programming for equity goals; and the move “upstream” where policy-led initiatives (e.g., cash transfers) affect results across several sectors. Consequently, this seems to be a generally positive change. Similarly, the decline of HIV/AIDS evaluations to 4 per cent of the total in 2012, from previously 10 per cent, is considerable. The Evaluation Office will continue to monitor these trends.

Table 1  
**Topical distribution of evaluation reports, 2011**

<i>Topic</i>	<i>Percentage*</i>
Multi-sectoral	25
Education and gender	25
Child survival and development issues	13
Child protection	10
Policy advocacy and partnerships	4
HIV-AIDS	4
Cross-cutting themes and organizational performance	20
Not classified	3

*Source:* GEROS report 2012.

\* Total exceeds 100 per cent as some reports cover several topics.

**Indicator 3: Types of evaluations conducted**

20. Evaluations assess results at different programmatic levels; this provides an indication of institutional focus. A strong results orientation, for example, would be demonstrated by a preference for undertaking evaluations focused on impacts and outcomes, rather than on outputs. The data from 2009-2011 shows contrasting trends (see table 2). On the positive side, output-level evaluations have decreased from 33 per cent in 2009 to 27 per cent in 2011. However, impact-level evaluations have decreased even more, from 43 per cent in 2009 to 27 per cent in 2011. Meanwhile, outcome-level evaluations have increased from 24 per cent in 2009 to 46 per cent in 2011.

Table 2  
Types of evaluation conducted: trends 2009-2011

<i>Indicator</i>	<i>2009 (%)</i>	<i>2011 (%)</i>	<i>Trend direction</i>
Output level evaluations	33	27	Towards fewer: as desired
Impact level evaluations	43	27	Towards fewer: undesirable
Formative level evaluations	45	59	Towards more: cause?
Summative level evaluations	55	41	Towards fewer: undesirable

21. These trends may reflect the evolution of corporate strategies. The increased focus on strengthening national capacities may explain the growth in outcome evaluations, and the growth in accurate models and more frequent survey data may allow impact measures to be served by monitoring approaches rather than by evaluations.

22. Another trend concerns the timing of evaluation in the programme cycle. Formative evaluations conducted early in the cycle — assessing whether the programming is unfolding as planned in order to provide timely evidence to guide adjustments — have increased by nearly a third, from 45 per cent in 2009 to 59 per cent in 2011. Summative evaluations, conducted at the end of the programme cycle, investigate cumulative results. The proportion of these evaluations has fallen from 55 per cent in 2009 to 41 per cent in 2011. The equity refocus may be driving these shifts as programme managers seek assurance that the most disadvantaged are being reached with restructured program strategies. However, it is too soon to be certain, and the combination of lessened coverage and fewer summative evaluations risks widening evidence gaps concerning eventual outcomes and impacts.

#### **Indicator 4: Quality of UNICEF evaluations**

23. The quality of evaluations is obviously important. High quality is needed to ensure that evaluation results are reliable and can be used with confidence, and that the investment in evaluation is justified.

24. It should be noted that the proportion of good or excellent evaluation reports has increased slightly in recent years. A sizeable proportion of evaluations reports fall into the category of “almost satisfactory” or “almost confident to act”. Given the stringency of the rating process, these are typically evaluation reports where a few technicalities or gaps in reporting have pushed the assessment below the level required for a rating of “satisfactory”. However, the evaluation in such cases has generally been adequate, with shortcomings in reporting rather than in substance.

25. Nevertheless, a significant proportion of evaluations fall into the category rated as “poor”.<sup>5</sup> Here, it is helpful to place the results of recent GEROS assessments within the perspective provided by earlier meta-evaluations of evaluation quality. Focusing on the most problematic evaluations shows two trends.

26. First, the 2011 figure of 23 per cent of UNICEF evaluations rated as “poor” shows a welcome improvement from the baseline figure of 33 per cent during 2000-

<sup>5</sup> The figures were 15 per cent in 2009 and 30 per cent in 2010; the average for 2009-2011 is 23 per cent.

2002.<sup>6</sup> However, while the baseline figure was reduced following a major capacity strengthening push during 2004-2006, there has been little overall change since then. It seems to be difficult to reduce poor performance from the 20-25 per cent zone under current arrangements. Sections 5 and 6, therefore, outline new or reinforced initiatives under way, and indicate additional measures for consideration in coming years.

27. Second, within the overall ratings, some elements of evaluation practice are showing very positive movement. Against an ambitious standard, assessment of the incorporation of human rights, gender and equity in evaluations shows an increase from 18 per cent in 2010 (the first year that these issues were rated) to 33 per cent in 2011. Similarly, satisfactory stakeholder participation increased from 40 per cent in 2010 to 52 per cent in 2011.

#### **Indicator 5: Use of evaluation, including management responses**

28. The 2008 Evaluation Policy requires that a management response should be completed for each evaluation. Advocacy, training and oversight have resulted in remarkable improvements in compliance: in the preparation of management responses; and in uploading completed responses to the tracking system. Submission of management responses to the Global Tracking System has reached 93 per cent, from the 2009 baseline of 10 per cent. Achieving such a high level of responses is a major achievement over the past three years.<sup>7</sup>

29. Utilization depends not only preparing the response, however; implementation of management response commitments is key. Compliance has increased from 62 per cent in 2011 (the first year of the management response tracking system) to 82 per cent in 2012. As these figures are based on self-reporting by the offices, independent means are needed to verify that the responses were implemented and to the level of completion claimed.

#### **Indicator 6: Corporate-level evaluations**

30. Corporate-level evaluations, managed by the Evaluation Office, are listed in the Integrated Monitoring and Evaluation Framework (IMEF). The 2010-2011 IMEF presented commitments to conduct 15 corporate-level evaluations, 13 of which (86 per cent) were completed, and 2 (13 per cent) cancelled (due to data and funding shortfalls). The 2012-2013 global IMEF, now under implementation, was drawn up on the basis of wide consultations to ensure that the topics selected for evaluation were relevant and met demand. Of the 14 evaluation topics approved, 7 are underway and 5 more are scheduled to begin in 2013. To allow adequate oversight from the Evaluation Office, it is proposed to reschedule the remaining two to begin in 2014 (see annex).

31. The global evaluation agenda has improved in relevance. However, it does not offer complete coverage of major programmes and strategies. The Strategic Plan for 2014-2017 provides an opportunity for developing a comprehensive approach to evidence generation, uniting evaluation, national statistical monitoring and research,

<sup>6</sup> The present rating term for this category is “not confident to act”; the former term was “poor”.

<sup>7</sup> There is good evidence that management responses simply were not being done in the past: the 2000-2002 meta-evaluation showed that 75 per cent of offices had no recorded response to their evaluations.

underpinned by the Monitoring Results for Equity System (MoRES), including the humanitarian performance monitoring component. Within this framework, and tightly integrated with the new Strategic Plan, it is proposed that a global evaluation plan be presented to the Executive Board in 2014.

#### **Additional performance indicators**

32. The performance indicators now in use have prompted action at all levels of UNICEF and demonstrate that information to help to manage the evaluation function can usefully be generated at modest cost. Still, there are some important gaps in evaluation coverage. Data on expenditure specifically on evaluation is needed for any calculation of efficiency. Beyond the tracking of formal management responses, information on utilization and the value added by evaluation is needed to ensure that the function yields maximum benefit. Information on complementary sources of evidence is also required to develop an overall evidence generation strategy. Finally, performance measures are needed to manage and reinforce initiatives to strengthen internal, partner and national capacities. For these reasons, UNICEF has concluded that it is necessary in coming years to pilot a supplementary set of simple but informative indicators.

### **IV. The evaluation function in UNICEF: human and financial resources**

33. In recent years, UNICEF has updated the Executive Board every second year on the specifics of human and financial resources allocated to the evaluation function. This allows reporting of true trends, as distinct from unstable single-year fluctuations. Following last year's detailed biannual update,<sup>8</sup> the present report offers a broader perspective on human resources and financing.

#### **The human resource base**

34. UNICEF invests heavily in evaluation personnel, compared to almost all other development agencies.<sup>9</sup> Approximately 75 per cent of country offices have a professional post with "Evaluation" in the job title, which is an outstanding level of presence. The percentage of national staff is relatively low (32 per cent), which is good for career growth opportunities. The percentage of female staff is relatively high (44 per cent).

---

<sup>8</sup> E/ICEF/2012/13.

<sup>9</sup> While true of development agencies, UNICEF notably lags behind United Nations agencies with a humanitarian mandate, with the UNICEF ratio of EHA-dedicated staff member to evaluation output placed at 1:12, versus 1:4.7, for WHO and 1:3.7 for WFP, respectively.

**Table 3**  
**Human resources for evaluation**

<i>Indicator</i>	<i>2012 (%)</i>
Percentage of offices with a post including evaluation in the job title	75
Percentage of female staff	44
Percentage of monitoring and evaluation officer's time devoted to evaluation	14
Percentage of posts bundling evaluation duties with sectoral duties	32
Incumbent's longevity in monitoring and evaluation posts in UNICEF	< 2 years: 20% 2-5 years: 49% >5 years: 31%
Percentage of evaluation posts located in regional or country offices	89

35. However, solutions are needed to address the following issues:

- (a) About 25 per cent of all offices have no such post;
- (b) Another 25 per cent only have such posts at levels 1 or 2, unsuited to managing highly technical, complex or sensitive assignments;
- (c) There is a pronounced concentration of posts in larger offices. Smaller offices, especially in middle-income countries, will require innovative approaches and solutions to support quality evaluation in the wide variety of UNICEF office types, sizes and contexts;
- (d) Monitoring and evaluation officers are allocated multiple roles and duties, reducing the average time available specifically for evaluation to just 14 per cent of all working time;
- (e) Monitoring and evaluation posts are often bundled with other roles, including programme duties that might be evaluated, thereby creating potential conflicts of interest;
- (f) Monitoring and evaluation staff members serve, on average, just 4.5 years on monitoring and evaluation duties, due to mixed incentives, including a weakly articulated career path and the high level of success enjoyed by monitoring and evaluation staff competing for managerial posts. A cadre of longer-serving professionals is needed to meet the increasingly challenging technical demands of evidence generation.

36. The location of these posts strongly reflects UNICEF as a decentralized organization. UNICEF assigns 89 per cent of its monitoring and evaluation staff to field offices, where they can engage closely with programme staff. This decentralized model differs sharply from the pattern in most other agencies, which rely largely on their central evaluation offices. However, it also creates a demand for support and quality assurance by regional colleagues that can be hard to meet, given the many competing responsibilities.

37. The UNICEF evaluation cadre is heavily engaged in various inter-agency activities. This contribution carries opportunity costs, including unmet needs related to the UNICEF programme. The strong monitoring and evaluation field presence of UNICEF is not mirrored by many other agencies, which make very limited

investments in monitoring and evaluation staffing at country level. As the United Nations system evolves, there is a need for a coordinated approach to meeting monitoring and evaluation staffing needs in all programme contexts, including humanitarian situations.

### **Strengthening internal evaluation staff capacity**

38. Properly skilled staff members are a key driver for good-quality evaluations, together with strong leadership and adequate financial provision. Significant investment continues to be made to improve internal evaluation capacity at all levels. These efforts include knowledge management activities, such as web-enabled communities of practice and web-based seminars (“webinars”), a help desk function, e-bulletins and newsletters, as well as conventional network meetings, study visits and training seminars.

39. Significant adjustments and improvements were undertaken in 2012, including the following:

(a) New self-instruction through online training courses. These include a broad new course targeting both programme staff and monitoring and evaluation specialists, providing instruction on managing good quality evaluations and effective use of evaluations;

(b) New tools and training on equity-focused evaluation. These efforts range from broad general instruction to detailed methodological guidance in this area of concentrated corporate effort. The manual on designing and managing equity-focused evaluations was translated into Arabic, French, Russian and Spanish, and has been disseminated. The manual was followed by training sessions and a highly successful new e-learning programme available inside and outside UNICEF, which has attracted nearly 5,000 people from 168 countries.

40. These new efforts are significant, but it is uncertain whether they will be sufficient to take evaluation in UNICEF to a new level and yield improvements in evaluation coverage and quality. It is important to supplement the updated evaluation policy with a comprehensive strategy that would focus on developing staff capacity and increasing engagement of external human resources.

### **Financial resources**

41. According to expenditure data from 2010, around 3 per cent of the UNICEF programme budget is spent on evaluation, analysis, research, studies, data acquisition, databases, surveys and statistics. This is in line with the target of 3-5 per cent set in the 2008 Evaluation Policy for total expenditure on these various channels of evidence generation. However, given the shift toward more upstream programming, support to national policies and advocacy, and to global advocacy in support of child rights, the argument can be made that investment in evidence generation needs to increase substantially.

42. With specific regard to evaluation, the 2010 data showed that field offices spent 0.33 per cent of their programme budget in evaluation activities (excluding staff costs). Globally, expenditure rises to 0.4 per cent if the Evaluation Office resources are added. Considering the coverage and quality issues noted elsewhere in this report, it is clear that the function is under-resourced. It is below the notional target of 1 per cent referenced in the recent review of United Nations system-wide

evaluation,<sup>10</sup> and far below the expenditure guideline of 3 per cent set out in the evaluation policies of some comparator organizations, including UN-Women and the United States Agency for International Development (USAID).<sup>11</sup>

43. Moreover, the funding of the Evaluation Office in past years has come from a mix of core resources and other resources, either directly contributed to the office or assigned to the office by programme units contributing to a specific evaluation. Other resources have accounted for 38 per cent of the budget over the past three years. In 2012, for the first time, an allocation from global thematic funds was provided in recognition that delinking funding from the sections whose work was to be evaluated would facilitate independence. However, thematic funds are available only intermittently and not for all evaluation topical areas. Therefore, increased funding from core resources is required to support the global evaluation agenda. This would safeguard the independence of global evaluations and, as core resources would be more predictable, speed their delivery.

Table 4  
**Financial resources for evaluation**

<i>Indicator</i>	<i>2010 (%)</i>
Percentage of field office programme budget spent on evaluation	0.33
Percentage of total UNICEF programme budget spent on evaluation (field plus NYHQ)	0.4
Percentage of total UNICEF programme budget spent on evaluation, monitoring, research, study, surveys	2.95
Percentage of Evaluation Office budget coming from non-core resources: actual spending	38 (2010-2012)

44. Resource allocation for evaluation has become harder to measure, given recent changes in UNICEF management information systems. Data coding on evaluation expenditure is less precise than before, which makes monitoring and reporting harder. There will be an opportunity to correct this when updating corporate systems.

45. Increased investment in evaluation is needed, but this is only a part of the picture. Even well financed evaluations can be poorly done. Resources need to be managed wisely throughout the evaluation process, and prerequisites for good evaluation need to be in place, including well-designed programmes and a sound information base provided by adequate monitoring data.

<sup>10</sup> Angela Bester and Charles Lusthaus, *ibid*, p.39: “While there is no official standard for the budget ratio, a widely-used guide is to spend 1 percent of the total budget on evaluation.”

<sup>11</sup> *Evaluation policy of the United Nations Entity for Gender Equality and the Empowerment of Women*, UN-Women, 2012; *USAID Evaluation Policy*, USAID, 2011.

## V. Strengthening the evaluation system

### Reviewing the evaluation function and policy

46. In 2012, the Evaluation Office commissioned external consultants to collect and analyse evidence on the performance of the decentralized evaluation function. The synthesis report on the UNICEF evaluation function and policy is available on the Evaluation Office website.<sup>12</sup> Data from the report is cited in the following sections, combined with evidence from other sources.

47. The synthesis report found that the Evaluation Policy remains relevant and closely aligned with the UNEG norms and standards. It found that while the policy is an important foundational guide for offices to understand corporate goals and their accountabilities, some parts of the policy need updating. These updates are addressed in the revised Evaluation Policy and not detailed here. The review also found evidence of a number of constraints to implementation of the policy. This section spells out some of the issues requiring attention.

### Governance and leadership

48. The synthesis report found that a decentralized evaluation function aligns well with the operational nature of the decentralized programming environment of UNICEF. This in turn places a premium on leadership by senior management at regional and country levels. To support managers, Geros results and other data on key performance indicators are routinely fed back via dashboards and scorecards at all levels. The synthesis report found that this effort is appreciated; it has helped to build awareness and sustain action. Regional elaborations of these scorecards have been put in place, as in Middle East and North Africa (MENA).

49. The observed need to strengthen the evaluation function is leading regions to provide effective oversight, though largely contained to evaluation of regular programmes. In Central and Eastern Europe and the Commonwealth of Independent States (CEE/CIS), MENA and the Latin American and Caribbean regions, the regional directors wrote in 2012 official memos to all the representatives asking for increased attention to the evaluation function. CEE/CIS and Latin America and the Caribbean have identified monitoring and evaluation as a core function, to be prioritized within regions featuring middle-income countries with their particular needs and capacities. These regional offices have begun to explore what impact this could have on strengthening human and financial resources for the evaluation function. Their strategic thinking leans towards adopting different monitoring and evaluation structures from previous models (e.g., multi-country sharing of monitoring and evaluation officers): requiring revision of job descriptions; placing a premium on partnerships; and supporting a swift shift to policy-level evaluation.

50. More generally, renewed attention is being given to empowering, and holding accountable, a broader range of staff for fulfilling evaluation roles. The 2008 Evaluation Policy did not state sufficiently clearly the roles and accountabilities of different evaluation stakeholders. Evaluation accountabilities are set out more clearly in the revised Evaluation Policy and the revised internal programme manual as well as in a new online staff training course on managing evaluations.

---

<sup>12</sup> *Synthesis report on the UNICEF evaluation function and policy*, Universalis, 2013.

51. Despite these advances, UNICEF has never set clear performance standards for the evaluation function at regional and country levels. There is now a need to do so while retaining the flexibility to address contrasting policy, programme and operational settings. The introduction of clear performance standards will clarify expectations while giving offices flexibility in deciding how to meet evaluation challenges.

#### **Assessing the demand for evaluation**

52. Coverage gaps and the relatively low numbers of evaluations make it critical to understand how demand arises. An external assessment of the demand and use of evaluation, which surveyed UNICEF representatives, provides some answers.<sup>13</sup> Interestingly, many representatives believe both internal and external demand for evaluations are rising (44 per cent internal, 46 per cent external), while a smaller number perceive declining demand (16 per cent internal, 5 per cent external). The rest perceive no shift. Sources of demand include the professional judgment of staff and national partners (generating 32 per cent of the demand), donor requests (29 per cent) and mandatory requirements (26 per cent).

53. Yet increasing demand is not met by increasing investments or increased numbers of evaluations. While financial constraints may be part of the story, we see evaluation gaps in many relatively resource-rich offices. Financial constraints therefore do not fully explain the shortfalls.

54. A further paradox concerns the stated use of evaluations. UNICEF representatives state that the most important goal of evaluation is to support evidence-based policy advocacy (50 per cent of responses), followed by programmatic decision-making needs (28 per cent) and accountability (17 per cent). These goals align exactly with UNICEF evaluation policy aims. However, if policy advocacy were so important, then one would expect to see more impact evaluations and more summative rather than formative evaluations.

55. Looking to future needs, representatives expressed great interest in equity-focused evaluations (60 per cent), impact evaluations (52 per cent) and policy evaluations (42 per cent). Innovation is likely to be another growing need. While UNICEF is well advanced in strengthening capacities for equity-focused evaluations, the other fields need more support and there will be challenges in developing new methodological approaches to match the diverse programming environments across UNICEF.

#### **Enhancing strategic planning and coverage of evaluations**

56. The intensified pressure to deliver relevant analyses is encouraging innovation and collaboration in strategic planning and evaluation. The regional offices and regional management teams are again the drivers.

57. In MENA, country offices responding to the Syrian crisis have developed, with regional support, a sub-regional monitoring and evaluation system for humanitarian programmes supporting refugees, while WCARO is presently launching a real-time evaluation of the unfolding crisis in Mali. In CEE/CIS, the regional management

<sup>13</sup> *Internal Review of the UNICEF Evaluation Policy and Function. Survey Report.* Universalis, 2012.

team approved a regional knowledge and leadership agenda addressing 10 key results are as concerned with reduction of violations of child rights and closing equity gaps. The focus is on programmes under implementation in a significant number of countries, with sufficient material to be readily documented and evaluated. Aware of the abundant experience to draw upon, CEE/CIS has launched multi-country evaluations on five distinct themes to assess the extent to which child rights violations and equity gaps have been reduced, and identify how UNICEF has contributed to these changes.

58. All regions have a policy of reviewing country-level integrated monitoring and evaluation plans in advance of their implementation, but this effort is being expanded in scope and comprehensiveness. Several regional offices have reviewed the “evaluability” of country programme documents and country programme action plans, and are helping offices adjust their programme designs and monitoring and evaluation strategies to ensure that strong evaluations will be possible when needed in the programme cycle.

### **Promoting and supporting quality evaluations**

59. High-quality evaluation requires not only necessary human and financial resources, but also appropriate procedures and technical approaches. A key evaluation design moment is the development of the terms of reference. The 2000-2002 baseline showed that 75 per cent of evaluations rated as “Poor” started on the wrong foot with weak terms of reference. Recent GEROs data shows that in around 50 per cent of cases, well-prepared terms of reference had a positive overall effect on the quality of final reports. Quality support in preparing good terms of reference can be usefully provided by the regional office or external firms contracted to review and comment on the terms of reference.

60. Analysis conducted by CEE/CIS has indicated that another system feature — a formal review of the draft final report — systematically improved the quality of evaluation report, either by presentational suggestions or by prompting additional data collection and re-analysis. Reports whose drafts were not reviewed ended with the worst ratings. The added value of an independent review at key moments is now considered proven.

61. A clear target for future effort is, therefore, to extend the coverage of quality support and quality assurance, whether internally or with the assistance of external specialists. The obligatory review of the terms of reference and the draft reports — these can be commissioned at low cost — should become a performance standard that is monitored and used by senior managers.

62. Some country offices are already establishing innovative arrangements and partnerships in this area. For example, the Burundi country programme has developed a partnership with Carleton University in Canada around quality assurance, while the Zimbabwe programme is establishing an external centre of excellence to review their evaluations and support other research and knowledge management roles.

### **Enhancing evaluation usage and management responses**

63. Thanks to effective monitoring and follow-up, progress has been recorded in preparing management responses and submitting them to the global tracking system.

Beyond the formal management response mechanism, UNICEF has been developing good dissemination and utilization practices. System-wide data on demand and use are now available, for the first time, from the commissioned surveys, mentioned above.

64. Representatives reported a high level of use of evaluations. Evaluation findings are incorporated into evidence-based policy advocacy (indicated by 53 per cent of respondents as a “high use”) and programmatic decision-making (42 per cent). However, the surveys showed room for improvement: 75 per cent of representatives said evaluation findings were “sometimes” systematically disseminated externally whereas only 18 per cent responded “always” (and 7 per cent responded “not sure/don’t know”). UNICEF offices need to improve dissemination to realize the full benefit of investments in evaluation. There are a number of recent examples of good practice: production of fact sheets (CEE/CIS, the Uganda and Zimbabwe programmes); dissemination of findings through the United Nations country team (Iran); presentation of findings to parliament through the national evaluation association (Morocco); and the systematic documentation and sharing of innovations on the dissemination and use of evaluations (East Asia and Pacific Regional Office).

65. Beyond dissemination, the effectiveness with which evaluation evidence is employed has not been specifically examined. Two global evaluations will address this question in 2013: a holistic look at the policy advocacy performance of UNICEF and a sectoral evaluation of the work of UNICEF in upstream education advocacy. These evaluations will look not only at success stories, but also at instances where advocacy efforts did not succeed.

## **VI. National evaluation capacity development**

### **Global partnership to enhance evaluation capacities**

66. Strong national evaluation capacities can play a key role in promoting development effectiveness. We already see, for example, an effective use of evaluation in guiding decision-making around several major social programmes in Latin America. The potential impact of strong national evaluation systems has been leading to investment by UNICEF and other donors for almost 20 years, with support and encouragement from evaluation professionals and supporters in programme countries. This effort has so far concentrated mainly on national and regional evaluation associations. However, more interconnected and better resourced global efforts are required. The past year has seen major strides in drawing together a coherent global partnership, providing a springboard to boost evaluation capacity development at all levels:

(a) Establishment of EvalPartners, an initiative to strengthen civil society evaluation capacity (complementing efforts to strengthen public sector capacities). The 31 charter members include all regional evaluation associations, UNEG, UNDP, UN-Women, the African Development Bank, the Governments of Spain and the United States of America, the Rockefeller Foundation and other major stakeholders;<sup>14</sup>

<sup>14</sup> See <http://www.mymande.org/evalpartners>.

(b) Establishment of an equity-based mission to serve the disadvantaged in society;

(c) Agreement on a conceptual framework focused on strengthening evaluation associations at national and regional levels (complementing initiatives serving to strengthen individual skills and knowledge);

(d) Launch of the EvalPartners website (tied in with social media) to facilitate communication;

(e) Mapping of regional and national evaluation associations or “voluntary organizations for professional evaluation”, identifying 150 associations in 110 countries, as well as case studies on their effective performance;<sup>15</sup>

(f) Convening of the International Forum on Civil Society Evaluation Capacities.<sup>16</sup> The forum, which took place in December 2012, facilitated the sharing of good practices and lessons learned by evaluation associations and other key stakeholders, and identified priorities for EvalPartners in 2013. All partners committed to the Chiang Mai Declaration pledging to work with civil society partners in support of evaluation;<sup>17</sup>

(g) Preparation of an advocacy strategy promoting the enabling environment for evaluation, for implementation in 2013-2015: targeting, inter alia, the inclusion of robust evaluation commitments in the processes associated with the anticipated sustainable development goals.

67. UNICEF has been a major driver behind EvalPartners and an active participant, with generous support from several donor agencies and foundations. This support will continue as EvalPartners matures. Realizing the ambitious goals of the partnership will require clear target-setting, careful performance monitoring and, in due course, independent assessment.

### **Global learning**

68. EvalPartners benefits from the global learning platform “MyM&E”, a user-friendly knowledge management system intended support individual and national evaluation capacity strengthening. The numerous ways MyM&E facilitates capacity strengthening can be viewed on the website.<sup>18</sup> An implicit proof of its value to users is the 500 per cent growth in visitors to the website since 2011, to almost 250,000 visitors in 2012 (originating from over 168 countries) and the 800 per cent increase in downloads, to almost one million pages in 2012.

69. MyM&E constantly updates its products. For example, 2012 saw the launch of a new e-learning programme on development evaluation, taught by 33 world-level keynote speakers.<sup>19</sup> As of December 2012, some 8,000 people from over 168 countries had registered. User satisfaction is very high, with 72 per cent of participants expressing satisfaction and 88 per cent stating their intent to attend future e-learning courses. Efficiency is very high once the development costs have been absorbed, with a cost per participant of about \$2.

---

<sup>15</sup> See <http://www.mymande.org/evalpartners/international-mapping-of-evaluation>.

<sup>16</sup> See <http://mymande.org/evalpartners/forum>.

<sup>17</sup> See [http://www.mymande.org/sites/default/files/Declaration\\_evalpartners\\_English.pdf](http://www.mymande.org/sites/default/files/Declaration_evalpartners_English.pdf).

<sup>18</sup> See [www.mymande.org](http://www.mymande.org).

<sup>19</sup> See <http://mymande.org/elearning>.

### **Strengthening country-led monitoring and evaluation systems**

70. As in past years, UNICEF continues to strengthen national evaluation systems, especially in partnership with other United Nations agencies. The Executive Board has encouraged this area of work; it is also highlighted in the United Nations General Assembly resolution on the quadrennial comprehensive policy review.

71. A 2012 mapping exercise identified some 120 UNICEF country offices engaged in strengthening national evaluation institutions and helping raise national level demand. The East Asia and Pacific Regional Office and the Regional Office for South Asia have been especially active through the United Nations Evaluation Development Group for Asia and the Pacific (UNEDAP). In tandem, UNEG and UNEDAP have agreed to develop a regionally tailored e-learning programme within the EvalPartners initiative.

72. A critical mass of experiences with university-based national centres of excellence is now emerging. Previous reports have noted efforts in the Latin American and the Caribbean and in South Asia. MENA led the new activity in Egypt (where a professional diploma in research and evaluation was developed with the University of Assyout and the University of Helwan); in Lebanon (the American University of Beirut); and Sudan (the Ahfad University for Women in Khartoum). In the Eastern and Southern Africa region, the Burundi programme established a memorandum of understanding with the national university to strengthen national evaluation capacities.

73. Within a gamut of spontaneous South/South collaborations, UNICEF helped Morocco's National Office for Human Development connect with the Mexican National Council for Evaluation of Social Development Policy, to develop systems evaluating public policies. UNICEF also found southern partners to support demand in Cambodia, North Korea and Myanmar. Nationally focused efforts were seen in Ukraine and Zimbabwe, which assessed their nascent national evaluation systems in order to design a national evaluation capacity development strategy. All countries can benefit from the UNEG task force output noted earlier, the document, *Practical tips on how to strengthen national evaluation systems*, which has been translated and disseminated widely.

74. Given the growing range of activities in this area, agreement on tools and methods to assess strengths and the gaps in national evaluation systems is needed, along with routine monitoring and sharing of progress in country-led efforts. This would help to strengthen this important area of evaluation effort.

## **VII. Conclusion: challenges and prospects for evaluation at UNICEF**

75. This report has outlined ways in which the evaluation function at UNICEF has continued to play a significant role in the work of UNICEF and — through joint evaluation activities — in the work of the United Nations system. It describes how the performance of the function has improved, with better management and support mechanisms in place, though greater attention is needed to ensure better evaluation coverage and quality. Globally, the work of UNICEF and its many partners has generated a widely welcomed contribution to evaluation capacity development;

however, in many countries much remains to be done to institutionalize effective national evaluations systems.

76. The evaluation function is well placed to support implementation of the new Strategic Plan within UNICEF, to measure and assess progress under the new strategy, and to inform decisions on new policies and programmes. An updated evaluation policy, an evaluation based in the Strategic Plan, and an implementation strategy to strengthen the evaluation function will underpin the evaluation activities across the organization, as UNICEF continues its work for children in a changing world.

## **VIII. Draft decision**

77. *The Executive Board,*

1. *Takes note* of the annual report on the evaluation function and major evaluations in UNICEF (E/ICEF/2013/13);

2. *Reaffirms* the central role played by the evaluation function in UNICEF and the importance of the principles set out in its Evaluation Policy (E/ICEF/2008/4);

3. *Welcomes* the evidence presented in the report of continued strengthening of the evaluation function at the decentralized level, and *encourages* UNICEF to take further steps;

4. *Takes note* of the key performance indicators tracking the effectiveness of the evaluation system and the administrative data on human and financial resources;

5. *Notes* the positive work to strengthen the skills, capacities, and systems of national partners and to strengthen South-South learning;

6. *Requests* UNICEF to:

(a) Report on steps taken to ensure relevant evaluation results are systematically considered and used in preparing key policies, strategies and programmes.

## Annex

### Global evaluations: Progress in implementing the integrated monitoring and evaluation framework for 2012-2013

(Status as at 31 March 2013)

<i>Title of global evaluation</i>	<i>Status</i>	<i>Comment</i>
1. Multiple indicator cluster surveys	Implementation	Completion expected in 4/13
2. Evaluability and process evaluation of equity refocus	Implementation	Completion expected 9/13
3. Cluster lead agency role in humanitarian action	Implementation	Completion expected 6/13
4. Child protection in emergencies	Implementation	Completion expected 6/13
5. Community approaches for total sanitation	Implementation	Completion expected 7/13
6. Emergency preparedness systems	Implementation	Completion expected 9/13
7. Upstream work in the education sector	Implementation	Request for proposal issued 1/13. Completion expected 9/13
8. Violence against children	Pending	Child Protection section is contributing funds. Expected mobilization 6/13
9. Preventing stunting	Delayed	Complex issue. Will require a scoping exercise in 2013. To start 2014
10. Social protection	Pending	Expected mobilization 5/13
11. Advocacy and policy change	Pending	Expected mobilization 5/13
12. Application of results-based management	Superseded; Pending	Evaluability of Strategic Plan exercise deemed more critical. To begin 5/13
13. Communication for development	Pending	Expected mobilization 7/13
14. National capacity development	Delayed	Complex issue. Will require a scoping exercise in 2013. To start 2014
<i>Mobilization:</i>	Terms of reference finalized or under design; resources being gathered	
<i>Implementation:</i>	Work under way or complete; analysis in process	
<i>Completed:</i>	Final report delivered; dissemination under way	
<i>Pending:</i>	Maintained in schedule; no action has yet been taken	
<i>Superseded:</i>	Original emphasis has shifted to a modified, higher priority	
<i>Delayed:</i>	Proposed to begin in 2014	