Economic and Social Council

United Nations Children's Fund
Executive Board
Annual session 2013
18-21 June 2013
Item 10 of the provisional agenda*

UNICEF Evaluation Office

Thematic synthesis report on evaluation of humanitarian action

Summary

     This synthesis report provides an overview of the state of evaluation of humanitarian action (EHA) within UNICEF from 2008 to 2012. It is intended to inform discussions on humanitarian-related evaluation and, more generally, on humanitarian action. The report summarizes evaluation coverage and quality over this period, major themes emerging from recent exercises, and illustrates how UNICEF has used evaluation to improve its approach to humanitarian action.

I. Introduction

1.   Humanitarian crises cost lives and prevent the realization of children's and women's rights. Conflicts, natural disasters and complex emergencies kill or injure children and separate them from their families, schools and communities. Emergencies destroy fragile institutions and hit struggling communities, eroding development gains and jeopardizing opportunities for economic and social development. By preying on pre-existing vulnerabilities and creating new ones, emergencies lead to less equitable life chances for children.

2.   Accordingly, UNICEF has been heeding the call to humanitarian action since its inception in 1946. Later, this mandate was articulated in the Convention on the Rights of the Child and its Optional Protocols, which identified a set of universal rights as inalienable and non-severable during emergencies, and conferred specific duties in these situations. More recently, the Core Commitments for Children in Humanitarian Action (CCCs) provide UNICEF and its partners a framework for humanitarian action, and commit it to help realize these rights in emergencies.

3.   Over time, the role of UNICEF in humanitarian action has grown in numeric terms, from 230 responses in 2007 to 292 in 2011.[1] This includes the organization's response to Level-1 emergencies (UNICEF country offices managing on their own), Level-2 emergencies (country offices requiring additional support by the corresponding regional office) and Level-3 emergencies (large-scale emergencies requiring an organization-wide effort).

4.   This increased activity reflects the growing toll emergencies have taken on children. In the late 1990s, disasters alone affected an estimated 66.5 million children a year; this figure is expected to rise to 175 million annually over the next decades because of the combined effects of climate change, population growth, urbanization and economic uncertainty.[2] UNICEF has expended considerable resources to boost capacity to meet growing demand for assistance, raising roughly $1 billion of emergency funding in 2011, compared to $200 million in the late 1990s.

5.   The role of UNICEF in emergency response has expanded qualitatively as well. Since 2005, a process of humanitarian reform has sought to improve inter-agency response through greater predictability, accountability, responsibility and partnership among humanitarian actors. A main pillar of these reforms, the Cluster Approach, aims to strengthen humanitarian response by assigning sectoral coordination responsibilities to a designated Cluster Lead Agency (CLA). As lead agency for three clusters (nutrition, water, sanitation and hygiene education (WASH) and, with Save the Children, education) and two areas of responsibility, (child protection and, with the United Nations Population Fund, gender-based violence), UNICEF bears more cluster lead responsibilities than any other agency.

6.   The reform process entered a second phase in 2011 with the Transformative Agenda. UNICEF has been actively involved in shaping this second wave of reform, which has sought to strengthen response in a number of areas: (a) leadership

_____

[1] Country office annual reports, annex A.
[2] See Angela Penrose and Mie Takaki, "Children's rights in emergencies and disasters", *Lancet 367* (2006): 698-99; and *Feeling the Heat: Child Survival in a Changing Climate*, Save the Children (2009).

(inter-agency Level-3 activation and rapid response mechanisms, development of emergency coordinator rosters and empowered leadership); (b) coordination (better managed, strategically used and more inclusive clusters); and (c) accountability (agreed humanitarian programme cycles and shared strategic frameworks in emergencies, against which to measure performance, and enhanced accountability to affected populations).

UNICEF evaluation of humanitarian action

7.   Given its significant humanitarian profile, UNICEF relies on a range of knowledge sources to help it prepare for emergencies and respond to them in ever-better ways. Evaluation forms one critical pillar, with the UNICEF Evaluation Policy[3] and the CCCs both underlining the organization's pledge to ensure adequate monitoring and evaluation of its humanitarian work. Evaluation represents the systematic and impartial means of meeting the full range of the learning and accountability needs of UNICEF. This includes accountability for results achieved for children and women in emergencies, fiduciary responsibility for investments in humanitarian action and overall learning and improvement.

8.   Evaluation has received heightened stature through policy directives both within UNICEF and in the inter-agency realm. It now forms an explicit element of UNICEF emergency response planning, having been incorporated into the Simplified Standard Operating Procedures (SSOPs) for Level-3 emergencies. At the inter-agency level, it has been a core element of humanitarian reform, which early on urged greater inter-agency collaboration on evaluation to meet the growing need for humanitarian accountability and learning. This trend has carried over into the Transformative Agenda, which has amplified emphasis on joint accountability for results, beginning with shared strategic plans. Although specific guidance for implementing the Transformative Agenda is under development at the time of writing, evaluation remains an explicit, inextricable element in the humanitarian programme cycle: evaluations at the end of a cycle determine whether the joint strategic plans have achieved the goals targeted.

Report overview

9.   This report provides a snapshot of how evaluation has promoted humanitarian accountability and learning over the past five years. Following a brief background to situate evaluation within the broader context of humanitarian action, it presents a series of analyses, beginning with a descriptive overview of EHA coverage and quality, and synthesizes key themes emerging from major evaluations — what they are telling us about UNICEF strengths and challenges in humanitarian action. It then describes the level of EHA utility and use from the perspective of internal clients, and highlights such use in three examples. After reviewing the findings in light of relevant policy currents, the report offers recommendations for action by UNICEF management.

10.   Taken together, these analyses portray UNICEF as an organization generating a number of significant evaluative exercises of its own humanitarian work and, together with partners, at inter-agency level. These evaluations have largely been of

_____

[3] E/ICEF/2008/4.

sufficient quality to warrant action, with UNICEF management using these evaluations for positive change.

11.   However, key coverage and quality gaps remain. Few evaluations have taken place at decentralized level, and when they have, quality has been uneven, suggesting low capacity at this level. At headquarters level, capacity has not kept pace with growing demands, especially for evaluations focused on results. Policy developments related to the new strategic plan for 2014-2017 and the Transformative Agenda offer opportunities to address these gaps.

II.   Coverage and quality of humanitarian-related evaluations

12.   Evaluation of humanitarian action is similar to non-humanitarian evaluation within UNICEF in that it subscribes the norms and standards of the United Nations Evaluation Group. However, EHA presents a number of specific challenges: (a) limited access to affected populations; (b) fluid and often unclear programmatic objectives; (c) accordingly, results frameworks that are constantly in flux; and (d) a lack of baseline data, against which to gauge performance. At the same time, EHA often has to be undertaken quickly and with a light approach but without compromising thoroughness, rigour or credibility. Taken together, these challenges require competencies beyond those of other evaluations.

13.   A 2005 review highlighted UNICEF management's commitment to humanitarian-related evaluation but noted uneven levels of coverage and quality across the organization.[4] Since then, several developments have prompted the need to revisit these challenges. These include the implementation of the humanitarian reform, establishment of the UNICEF Evaluation Policy and revision of the CCCs, all of which underscore the need for an appropriate and credible humanitarian evaluation effort. Accordingly, in early 2013, the Evaluation Office undertook a systematic exploration of UNICEF-led or co-led EHAs from 2008 to 2012. This examination included a descriptive analysis of all relevant reports categorized as evaluations in the UNICEF Evaluation Reports Database, coupled with the quality ratings of these reports in the Global Evaluation Reports Oversight System (GEROS).[5] The findings of this analysis are summarized below.

Evaluation activity has increased, but is not commensurate with the full range of UNICEF humanitarian accountability and learning needs

14.   Humanitarian-related evaluations have become an increasingly common feature of the evaluation profile of UNICEF, in line with the organization's growing humanitarian role. As figure 1 conveys, evaluations embodying a sole or significant humanitarian focus have become more commonplace. These evaluations have achieved broad thematic coverage and examined the full spectrum of emergency operations, from emergency preparedness to response and transition.

---

[4] Review of UNICEF evaluation of humanitarian action, UNICEF, 2005.
[5] This report analysed all exercises labelled as "evaluations", even though upon closer inspection many of these were deemed reviews, assessments, lesson-learning exercises or other studies.

Figure 1
Number and percentage of EHAs and other evaluations, 2008-2012[6]
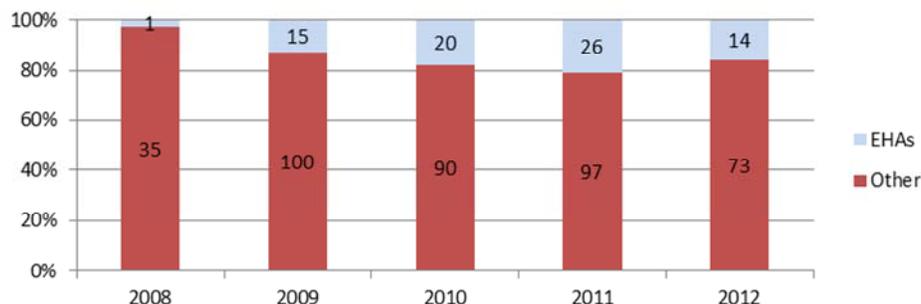
15.   Despite the significant human, financial, programmatic and reputational risks at stake in the humanitarian work of UNICEF, EHA remains a small fraction of the organization's evaluation landscape. Of the 471 evaluations produced by UNICEF during 2008-2012, only 76 (16 per cent), were humanitarian in scope.[7] This low level of activity is even more pronounced in emergency-affected countries. Although UNICEF responded to 1,025 humanitarian situations between 2008 and 2011, in only 34 cases (3 per cent) did it carry out evaluations of its response. Of these, a large number (29 in total) have taken place in Level-2 and Level-3 settings since 2010, leaving Level-1 emergencies — the most common emergency scenario UNICEF faces — largely unevaluated.

Humanitarian-related evaluation has largely been a headquarters endeavour

16.   When EHAs have occurred, these have largely taken place at headquarters level, specifically through the Evaluation Office. Of 76 such evaluations conducted between 2008 and 2012, two-thirds (50) were managed at headquarters level. Moreover, these were virtually always led or co-led by the Evaluation Office, even though the great majority of these were commissioned by other UNICEF divisions or offices. By contrast, only 24 EHAs occurred at the country-office level, a figure consistent with the low level of coverage of Level-1 emergencies.

17.   Regional offices are conducting even fewer such evaluations. Despite the increasingly regional nature of emergencies — evidenced in the Horn of Africa, the Sahel, Syria and Mali — only two EHAs were managed at the regional-office level over the entire 2008-2012 period, and one of these was co-managed by the Evaluation Office and the Office of Emergency Programmes (EMOPS).

The focus on results achieved for children and women in humanitarian action has been low

18.   The CCCs represent the normative standards against which UNICEF seeks to hold itself to account for programmatic and operational results. They outline specific commitments to bolster evaluation and organizational learning that UNICEF undertakes across the emergency continuum. However, as figure 2

_____

[6] Figures for 2012 will likely increase, as they only include reports submitted by the official deadline of 12 January 2013. By contrast, 2008 figures could be the result of low overall reporting, as the frequency of report uploads prior to 2009 was uneven.

[7] Ibid.

illustrates, very few exercises have focused on results achieved for children in humanitarian action. Although this percentage increased in 2012, only a minority of cases have such a focus.

Figure 2
**Number and percentage of results-focused evaluations, 2008-2012**
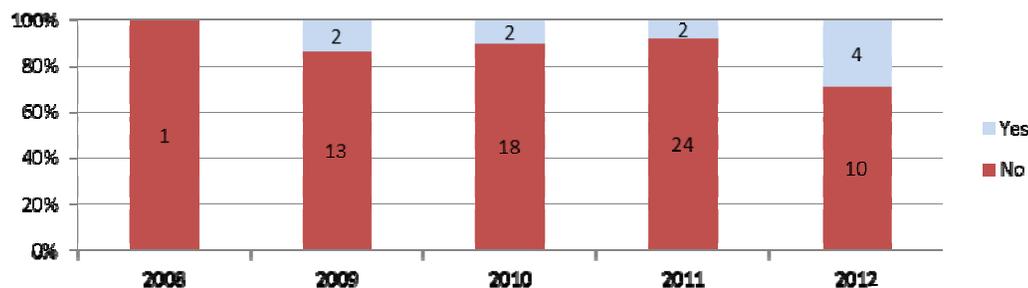
19. Instead, there has been a tendency to commission lower-level lessons-learned exercises and internal operational reviews, even though a majority of these were submitted as "evaluations" to the Evaluation Reports Database. Even at headquarters level, where most of the evaluations focusing on results have occurred, these have almost entirely been internal reviews of operational contributions to targeted results; internal lesson-learned exercises largely grounded in perceptions of results rather than objective measurement of such results; evaluability exercises focusing on future results measurement; or inter-agency real-time evaluations, which, despite the term "evaluation", have largely emphasized coordination issues rather than results delivered. Although most of these have been rigorous and effective, none has provided UNICEF a clear and credible sense of whether it is achieving results for children and women in the way that rigorous outcome-focused evaluations can. One exercise, a 2008 evaluation of the UNICEF response to the Indian Ocean tsunami, did aim to assess results against the CCCs but encountered significant measurement challenges.

When UNICEF has committed to evaluations of humanitarian action, they have largely been of sufficient quality to warrant management action

20. All UNICEF evaluations undergo an external peer review, the GEROS, which rates their quality on a scale from 1 to 4, where 1 represents the lowest rating and 4 the highest.[8] Despite the aforementioned coverage challenges, when UNICEF has committed to undertake EHAs, they have generally been of a high calibre. As figure 3 attests, more and more evaluations are receiving a rating of "Confident to Act" or "Very Confident to Act". The mean and median GEROS scores for these four years were 2.6 and 3.0, respectively, or generally "Confident to Act".

_____

[8] The scale used in this mechanism has changed over time, from "Poor" and "Excellent" (2008-2009) to "Not Confident to Act" and "Very Confident to Act" (2010-2011), and to "Unsatisfactory" and "Outstanding" (2012). The Confidence scale is used in this report for ease of comparison, and since most of the evaluations covered were rated on this scale.

Figure 3
Breakdown of EHA GEROS ratings, by number and percentage, 2008-2012
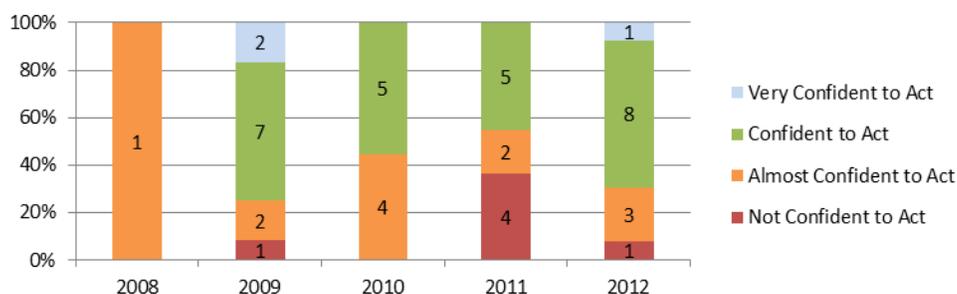

21. Within this overarching pattern, two important distinctions are noteworthy. First, the headquarters-led exercises has been of higher quality than those at decentralized levels — with a mean score of 2.9, or "Confident to Act", compared to 2.1, or "Almost Confident to Act", at decentralized level. Second, high quality has been harder to achieve in joint and inter-agency evaluations. Although no headquarters-led exercise received a "Not Confident to Act", all of those receiving an "Almost Confident to Act" rating were joint or inter-agency evaluations, which present unique challenges over which UNICEF has limited control.

Joint and inter-agency exercises have become a core feature of the UNICEF evaluation profile

22. Quality challenges notwithstanding, joint and inter-agency exercises have nonetheless become an integral aspect of the UNICEF evaluation profile. Since 2005, the process of humanitarian reform has led to a growing call for agencies to synchronize their efforts, including in the area of evaluation. Some of these collaborations are prescribed by policy directive, such as humanitarian financing evaluations required by donors or inter-agency real-time evaluations automatically triggered by the Inter-Agency Standing Committee. Joint and inter-agency evaluations are intended to minimize duplication of efforts and maximize collaborative learning around themes of common interest to the entire humanitarian community. They are also meant to reduce transaction costs borne by individual organizations — and, more importantly, by staff on the ground, who must set aside time amidst their many competing priorities to engage in the evaluation.

23. Figure 4 speaks to this trend. The percentage of such joint or inter-agency exercises rose dramatically between 2009 and 2010; they now constitute a majority of UNICEF humanitarian-related evaluations.

Figure 4
Percentage of joint and inter-agency exercises as a share of all UNICEF EHAs, 2008-2012
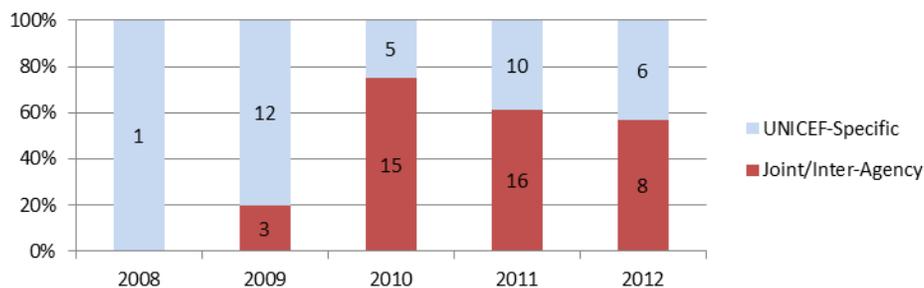
24.   The Office for the Coordination of Humanitarian Affairs (OCHA) has led most of these evaluations, with UNICEF and other agencies providing technical, financial and human resource support in a co-management or advisory capacity. In keeping with overarching trends described above, the Evaluation Office has represented UNICEF in most of these evaluations. More recently, the Evaluation Office collaborated with the World Food Programme (WFP) and the Government of the Netherlands to help co-manage an evaluation of WFP's leadership of the global logistics cluster.

25.   Together with partners, UNICEF has been working to improve the quality of joint and inter-agency evaluations. These efforts have begun to bear fruit. Since the quality-enhancing collaboration began in 2011, mean ratings have risen from 1.9 ("Almost Confident to Act") to 2.8 ("Confident to Act").

III.   Synthesis of themes emerging from major humanitarian evaluations

26.   In early 2013, the Evaluation Office commissioned an analysis of all 23 evaluations embodying a substantial humanitarian focus and receiving a GEROS rating of "Confident to Act" or "Very Confident to Act".[9] The analysis was not intended to provide a comprehensive overview of all evaluations, or all issues, or all of UNICEF operations. Rather, it constituted a narrow exploration of the most credible reports, aimed at identifying recurrent themes. The analysis of the 23 reports revealed five underlying themes, summarized below.

UNICEF appears to be contributing to humanitarian effectiveness in significant ways, saving thousands of lives every year

27.   Despite the aforementioned need for more bona fide results-focused evaluations, numerous reports do present separate strands of evidence that UNICEF is indeed successful in reaching children and saving lives. Overall, 13 reports acknowledge the contribution of UNICEF to humanitarian action, albeit with differing degrees of hard evidence.

_____

[9] The analysis included one report rated "Almost Confident to Act" for heightened geographic parity.

28. Several programmatic areas are of particular note. Five evaluations identify successful WASH operations. In response to the Haiti earthquake in 2010, UNICEF provided safe water to 1.2 million people. In the Sahel in 2012, it helped treat 2 million children suffering from malnutrition, including more than 700,000 severely malnourished. Although it is impossible to know how many lives these resilience efforts have saved, reports do point to the significant contribution of UNICEF in this area. In Central Asia and South Caucasus, UNICEF improved the ability of schools to undertake disaster preparedness and risk reduction measures, thereby enhancing their resilience to potential disasters. Furthermore, the organization's cash transfer project for protection of blanket feeding in Niger improved the resilience of the population, without any direct effect on market prices.

29. Several factors in the internal operational systems of UNICEF appear to be at play in explaining its successes to date, notwithstanding the need for further improvement. These include donor efforts to help UNICEF to build its humanitarian capacity;[10] UNICEF supply, logistics and finance capacity;[11] and its capacity in getting the right staff on the ground at the right time.[12] The specific steps UNICEF has taken to increase predictability, speed and accountability in its response since the Haiti earthquake are cited as critical to its responses in the Horn of Africa and the Sahel. These include the establishment of the SSOPs for Level-2 and Level-3 emergencies, the Monitoring of Results for Equity System (MoRES) in humanitarian action, also known as the Humanitarian Performance Monitoring system, and the fast-track recruitment process.

Sometimes deemed risk-averse in emergencies, UNICEF is becoming more risk-aware

30. Emergencies inherently bear high risks. Twelve reports found the skills and organizational culture required for understanding, valuing, and managing risk to be lacking. The 2010 Haiti review portrayed UNICEF as being risk-averse rather than risk-aware, and its staff as diligently adhering to internal procedure — even when risks could have been managed to achieve greater results. In its 2010 response to the Sa'ada conflict in Yemen, UNICEF was one of the first agencies on the ground. However, adherence to procedures better suited to development contexts, coupled with slow procurement, caused delays that reduced the organization's ability to respond effectively and hence compromised its credibility. In Somalia during the Horn of Africa crisis in 2011, UNICEF responded slowly to the rapid escalation in risks to the affected population and, rather than acting pre-emptively, only stepped up the response when a United Nations system-wide response was declared.

31. Nevertheless, as UNICEF increasingly engages in risk management, it has shown that it can indeed innovate for better results. The SSOPs developed in response to Haiti have been deployed and further refined for the Horn of Africa and the Sahel. A 2012 review of the fast-track recruitment process developed for emergency deployments showed this mechanism to be effective in getting the right people with the right skills on the ground at the right time — despite early concerns about the quality of personnel being recruited. The WASH programme in Haiti was

_____

[10] These include the DFID-UNICEF programmes of cooperation and the Netherlands-funded evaluation of education in emergencies.

[11] These were highlighted as crucial in the Sahel and Haiti responses.

[12] This was cited as a factor in the UNICEF fast-track recruitment process.

found to be more effective than other sectors because it engaged with the community and managed local capacity rather than avoiding the risks thought to be associated with this approach.

32. Among the best depictions of emergency responses are those where UNICEF was already involved beforehand in disaster risk reduction — by definition a form of risk management. Examples include the well-recognized outpatient therapeutic feeding programme in Ethiopia and the region-wide disaster preparedness programme in Central and Eastern Europe and the Commonwealth of Independent States (CEE/CIS).

Humanitarian partnerships are challenging, but UNICEF has worked hard to become a better partner

33. Many UNICEF programmes are implemented by non-governmental organizations through partnership cooperation agreements. The quality of these agreements and the organization's relationship with its implementing partners are the subject of findings and recommendations in eight reports. Many implementing partners lack the capacity to meet UNICEF reporting requirements or the financial means to wait for the disbursement of UNICEF funding; the inability to find appropriate implementing partners with adequate capacities emerges in 11 reports. In addition, large-scale emergencies create long supply chains managed by implementing partners, without sufficient coordination capacity to make these partnerships work. Several evaluations find UNICEF investment in the capacity needed to communicate and coordinate effectively with implementing partners and others to be low at decentralized levels. The placement of junior staff as cluster coordinators in at least two Level-3 emergencies (Haiti and the Horn of Africa) resulted in these important coordination roles becoming secondary to the priorities of more senior programme staff.

34. The Kenya country office used partnership cooperation agreements in 2011-2012 to obtain supplies faster and to avoid competing in the market with its partners. Other country offices in the Horn of Africa found ways to fast-track partnership cooperation agreements, to release supplies based on a note for the record, and to establish impromptu meetings on partnership cooperation agreements and on contracts. Most significantly, the SSOPs for Level-3 emergencies include fast-tracking procedures for partnership cooperation agreements, thereby addressing a bottleneck several EHAs had previously highlighted.

35. A handful of evaluations have shown that investing time and resources to cultivate partnerships in disaster-prone countries can pay dividends when emergencies occur. UNICEF Niger, for example, has cultivated important partnerships during non-emergency periods. A 2012 real-time evaluation of the Sahel emergency found that the best-functioning partnerships are long-term relationships developed long before emergencies strike.

36. Partnership has similarly been challenging at the global level, but has also witnessed improvement. A 2010 review of the UNICEF co-leadership in the global education cluster with Save the Children found the pairing of a United Nations agency and a non-governmental organization to be a complex undertaking at first. However, the two organizations have worked closely to address these challenges, not least by developing a joint management response to the review. In 2012, to enhance its cluster lead agency performance, UNICEF undertook a number of

structural changes, including moving implementation of the clusters from Programme Division to EMOPS and their co-location in Geneva. An evaluation currently under way on the organization's role as a cluster lead agency aims to help further strengthen UNICEF management of those cluster lead functions.

UNICEF has learned to work with rather than around local capacity but must do more to foster participation in needs assessments and programme design — including by affected populations

37. Evaluations have found that supplementing rather than supplanting existing national capacity generally leads to more effective humanitarian operations. This was the case in the 2010 WASH interventions in Haiti and the Pakistan floods and in some countries during the 2011 Horn of Africa response. In other cases, as noted, for example, in a 2008 evaluation of the Indian Ocean tsunami, UNICEF supplemented local capacity. The more common involvement of national institutions in Pakistan, Ethiopia and the Sahel suggests that UNICEF appears to have become more sensitized to the need to work *with*, rather than *around*, national actors.

38. Carrying out needs assessments remains challenging. Eights reports indicate that needs assessments did not take place, were incomplete or were merely general situation analyses. Other reports found that UNICEF inadequately communicates its intentions, current situation or reasoning to implementing partners, local authorities or affected populations. Where needs assessments have been done, local participation has been very low. Six reports indicate that needs assessments were reported as not involving the participation of affected groups.

39. This low engagement with affected populations has reduced the organization ability to design well-tailored, equity-focused programmes. Eight reports provide evidence of the UNICEF equity focus not having been achieved in humanitarian operations. In the 2011 Horn of Africa response, for example, a focus on systems resulted in UNICEF programmes failing to reach girls and women, people living with disabilities and children in difficult-to-access schools.

UNICEF is not effectively managing the flood of information generated in and on emergencies

40. Nine reports indicate that monitoring and evaluation was absent or poorly executed in specific emergencies. Where monitoring and evaluation frameworks have been developed, the experiences from the Education in Emergencies and Post-Crisis Transition programme and the response to the Horn of Africa and the Sahel emergencies suggest that more could be done to refine "good-enough" sets of locally adapted indicators.

41. These evaluations propose that monitoring and evaluation efforts work best when they are contextualized, when they are well coordinated with partners, and when they are integrated into preparedness efforts rather than surprising country offices and regional offices when an emergency strikes.

42. The aforementioned challenges associated with the capacities of implementing partners extend to monitoring and evaluation as well, with recent evaluations in Northern Kenya and the Sahel revealing that many local implementing partners and suppliers do not have sufficient logistical capacity to deliver over the "final mile" and lack monitoring and evaluation capacity for accurate reporting.

43.   Conversely, where information has been available, it has rarely been used effectively because there is simply so much of it, particularly when information and communication technologies had not been properly planned for ahead of time. An evaluation from Niger provides an example of relevant data being generated but not used. In some cases, as in Somalia, there is a large physical distance between management teams and information sources, making contextualization and verification of incoming data extremely difficult.

44.   Experience from the Horn of Africa suggests that for some UNICEF humanitarian teams there is a tendency to consider only information that is directly relevant to their geographical and sectoral areas of responsibility. For example, the refugee influx to Ethiopia and Northern Kenya took UNICEF by surprise, even though a stronger regional information management perspective would likely have revealed its likely occurrence. The tendency to manage information in silos means that coordinated responses to connected issues, such as WASH and nutrition or WASH and gender-based violence, are sometimes missed.

Synthesis summary

45.   Despite gaps in evaluation coverage and systematic results measurement, the most credible evaluations of humanitarian action underscore that UNICEF is aiming high — and is achieving some notable successes. Like any global organization, it faces challenges in navigating the inherent complexities of the humanitarian realm. Working across organizational boundaries to achieve results, whether with partners embodying diverse mandates and managerial cultures or with governments and affected populations, has not always been as straightforward as originally envisioned in the humanitarian reform process. Moreover, achieving results in rapidly changing situations has been onerous: information flows have been overwhelming, and staffs have not always been empowered to take action when and where it was needed. However, even in these areas of identified shortcomings, UNICEF appears to be making progress.

IV. Evaluation utility and use

46.   The foregoing analyses point to achievements and gaps in UNICEF humanitarian action as borne out in the evaluations reviewed. What they do not articulate, however, is whether these evaluations are useful to clients — or, more importantly, whether clients are using them to make meaningful improvements. Clients have produced management responses to all recent global-level evaluations of humanitarian action, in compliance with the UNICEF Evaluation Policy, but utilization goes beyond this formal process. In other words, have evaluations led to positive changes in UNICEF humanitarian work?

47.   The present section helps answer this question. Drawing on three case studies, one from each level of the organization, it illustrates how UNICEF has used evaluations of humanitarian action to improve the way it carries out its humanitarian work.

Country case study

Evaluating the unprecedented: The 2010 earthquake in Haiti

Haiti represents an example of how UNICEF has used evaluation meaningfully to effect large-scale policy changes within the organization. The devastating earthquake that struck Haiti on 12 January 2010 set in motion one of the largest and most complex emergency responses on record. Enormous resources flowed in to help humanitarian organizations respond. UNICEF Haiti was transformed, virtually overnight, from one of the organization's smallest offices to one of its largest. In addition to ensuring its programme accountabilities to the millions of children affected by this tragedy, UNICEF needed to guarantee proper coordination and make sure that resources were being managed, effectively and efficiently.

Two main streams of evaluative activity ensued in order to meet these knowledge needs. A series of inter-agency real-time evaluations was triggered; one (rated "Confident to Act") seeking to examine inter-agency coordination around the early response effort and another (rated "Almost Confident to Act") on the later recovery effort. Led by OCHA, with assistance from UNICEF and others, the two inter-agency real-time evaluations were used to retool operations on the ground. UNICEF paid more attention to programming in an urban context by better engaging the Haitian Government and civil society groups, and played a key role in transitioning the work carried out by its clusters to the hands of the Government.

In parallel, the UNICEF Office of the Executive Director commissioned an independent review of the organization's own response to the Haiti earthquake. Rated "Confident to Act" in GEROS, the review has resulted in a number of significant changes. The Haiti country office has incorporated the review's findings and recommendations into its programme and management plans, including for human resources, monitoring and reporting. At the global level, the review led to the development of the SSOPs for Level-3 emergencies as well as the Corporate Emergency Activation Procedures that trigger the SSOPs. Together, these procedures have led to a more predictable means of declaring such emergencies, determining who will be responsible for what aspects of it, and managing the response across UNICEF operations.

Regional case study

Evaluating in real time: The 2012 nutrition crisis in the Sahel

The Sahel experience demonstrates how evaluation can facilitate timely learning that feeds into concrete, and sometimes life-saving, operational decisions — and with an explicit view to enhancing equity. It also represents an example of strong collaboration across the organization to achieve high use of high-quality evaluation. Poor rainfall, combined with a complex array of other factors, provided the tipping point that led to crop failure and a serious loss of livestock in the Sahel belt of Western Africa in 2011. Food prices spiked and a nutrition crisis ensued. Projections estimated some 12 million people would likely be affected and that the lives of 1.1 million children were at risk of severe acute malnutrition.

Faced with a deteriorating situation, UNICEF scaled up for a regional response to the treatment of severe acute malnutrition in nine countries. The response was challenging not only because of its scale, but also because many of the most affected areas were difficult to access because of insecurity. In addition, implementing partners with adequate capacities were scarce, and national health systems lacked capacity to reach affected populations.

A real-time evaluation of the UNICEF response was launched in four of the affected countries, three months after the declaration of the crisis, to provide rapid and early learning to improve the organization's response and help mitigate future crises. Managed by the West and Central Africa Regional Office, together with the Evaluation Office and EMOPS, the real-time evaluation enabled UNICEF to establish a more coherent framework to guide its integrated, multi-pronged response. This included concrete recommendations on how to reach more children by extending the number of centres proving malnutrition treatment, using mobile clinics, promoting community-based case management of severe acute malnutrition, increasing capacity in remote areas, and better integrating service provision. The report received a GEROS rating of "Very Confident to Act", and its results have been widely used in the region. All involved country offices developed detailed management responses immediately after the exercise, and the regional office is currently implementing the recommendations.

Global case study

**Thinking before programming: DFID-UNICEF programme of cooperation, 2012-2015**

UNICEF received in 2012 the latest in a series of significant contributions by the United Kingdom Department for International Development (DFID) to boost its capacity for humanitarian action. UNICEF was required to achieve results in four specific capacity-related areas — and to demonstrate whether these results had been achieved. In essence, UNICEF had to think about evaluation even before the programme got under way.

The Evaluation Office commissioned an "evaluability" assessment — an exercise conducted at the early stages of a programme to help determine whether it will be ready for a meaningful evaluation later on. The evaluability assessment aimed to help management think through a number of key questions early on and thus help it achieve success. Was the programme logic clear? Was it clearly communicated to stakeholders, and understood and accepted by them? Were planned programme activities the most strategically sound interventions for achieving the targeted results, and were the resources sufficient? Were the management arrangements optimal? Would the data being collected demonstrate whether results were being achieved?

The assessment helped management strengthen the linkages between programme activities and targeted results, and put in place a communications strategy for greater cohesion between programme objectives and how it was being implemented at decentralized level. This also led to a more strategic process of allocating programme funds. Rated "Confident to Act", the exercise demonstrates how the evaluation and management functions can work together to address issues before they become problems.

48. These examples demonstrate that UNICEF can and does use major exercises to improve its humanitarian work in important and far-reaching ways. However, these examples do not show how widespread usage is. In order to understand the prevalence of using EHAs, the Evaluation Office, in early 2013, contracted an assessment of the perceptions of its internal clients regarding the usefulness and quality of the 76 evaluations of humanitarian action: in which ways they have used these exercises, and the factors influencing their use. This included an electronic survey and interviews with nine clients at the global and regional policy levels. While not a scientific study intended to draw statistical conclusions, the analysis nonetheless sheds valuable light on the experience of UNICEF clients. As figure 5 suggests, clients responding to the survey feel that the quality and usefulness of EHAs have improved overall in the period under review.

Figure 5
Client perceptions of quality and usefulness of evaluations, 2008-2012
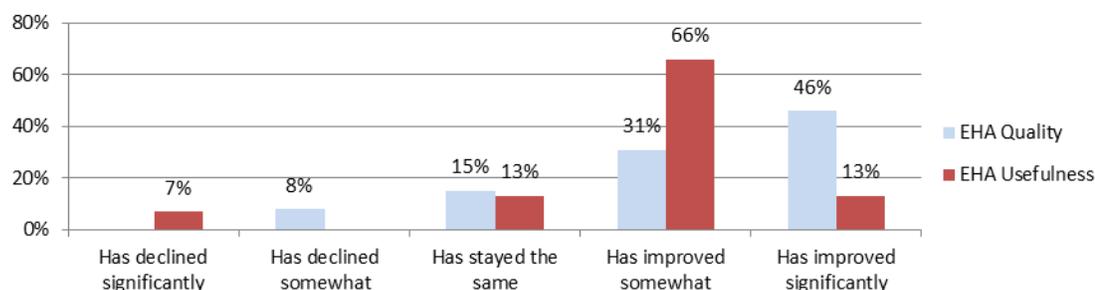
49. This overall sentiment extends to individual evaluations as well. In the survey, respondents were asked to rate, on a scale of 0-100, the usefulness of evaluations with which they were familiar, with 0 corresponding to no usefulness and 100 to extremely high usefulness.[13] The average rating was 65 over the five-year period (70 at headquarters, 63 at decentralized level). The three case studies above were among the evaluations of highest usefulness. Others included the 2011 evaluation of disaster risk reduction in CEE/CIS; a 2010 evaluation of the UNICEF education programme in Timor-Leste; an inter-agency real-time evaluation of the 2011-2012 Horn of Africa response; and a review of the global education cluster co-leadership role of UNICEF with Save the Children. Among the evaluations considered to be of lowest usefulness were three inter-agency real-time evaluations, three country-level evaluations, one global programme evaluation and one internal lessons-learned exercise.

50. Significantly, relatively high-quality evaluations were not always useful, and sometimes lower-quality evaluations were used. A 2010 inter-agency real-time evaluation in Pakistan, for example, rated "Confident to Act", did not receive a high usefulness rating. Conversely, evaluations such as a 2010 real-time evaluation in Yemen and the inter-agency evaluation of the Central Emergency Response Fund, both of which received an "Almost Confident to Act", were used.

51. Several clients, when interviewed, maintained that while quality is vital in determining whether an evaluation will be used, several other factors are equally important. Such factors include the level of attention senior management pays to the evaluation, how actively the evaluation manager and evaluation team involve stakeholders throughout the process while remaining sensitive to their many competing priorities, and how adeptly the evaluation manager balances independence with the need to work with clients to address apprehensions surrounding the evaluation. Direct engagement between the evaluation manager and the client has been particularly vital in lesser-quality evaluations, in which close collaboration has ensured proper use of relevant material without overusing poorly evidenced analysis.

_____

[13] Some evaluations received ratings by several clients and others by only one client.

V. Key issues and future directions

52. As the foregoing analyses convey, UNICEF is producing an expanding number of high-quality evaluations in the humanitarian field, both alone and together with partners. Importantly, the organization is also using these evaluations in a variety of ways to effect positive change. The high levels of overall quality and use appear to justify the investment of time, human capital and financial resources UNICEF has made. Overall, three broad issues emerge from the analysis that warrant management attention and action.

53. The organization's level of evaluation effort must be better tailored to its broad learning and accountability needs. That the bulk of evaluations continue to be undertaken at global level is unsustainable if learning is to occur where knowledge is needed most. While not all Level-1 emergencies must be evaluated, the virtual non-existence of evaluation of these emergencies is glaring. So, too, is the lack of evaluations at the regional level, particularly given the increasing frequency of regional emergencies. The lack of rigorous measurement and assessment of results represents another major gap. Two evaluations slated for completion by mid-2013 — the evaluation of the UNICEF emergency preparedness systems and the evaluation of the role of UNICEF as cluster lead agency — aim to address this gap to a degree. They are not enough, however. A more broad-based, strategic and results-oriented approach needs to be in place to determine what will be assessed when, and with what amount of evaluation effort.

54. A related issue has to do with institutional capacity for high-quality evaluations at decentralized levels. The Evaluation Office, in consultation with the Division of Human Resources, EMOPS and other divisions, has spearheaded a number of initiatives to boost capacity for EHA.[14] These efforts must continue.

55. Capacity gaps are not limited to the decentralized level, however. At headquarters, where much of the activity described has taken place, the Evaluation Office has assumed an expansive number of responsibilities. Besides managing global-level humanitarian evaluations, these include leading or co-leading complex evaluations at decentralized levels (to fill capacity gaps), engaging in major joint and inter-agency evaluations, and contributing to the many evaluation-relevant developments within humanitarian reform. The Evaluation Office is currently reviewing its capacity to meet these expectations and responsibilities.

56. These expectations are only likely to increase with the new demands the Transformative Agenda will place on evaluation offices or with the growing number of emergencies likely to confront the humanitarian community in future.

57. The capacity gap extends to the client side as well. As noted in internal consultations, clients rightly view their own engagement as critical to evaluation uptake. However, they are also understandably concerned about the time, labour and technical knowledge that such engagement entails — all rare commodities for colleagues engaged in emergencies. A dedicated knowledge management function is necessary to undertake this responsibility, and to help clients learn and act on lessons from evaluations and other knowledge sources. Although EMOPS once

---

[14] Measures include an e-learning module, to be rolled out in 2013; revisions to the UNICEF Programme Policy and Practice Manual; and evaluation management partnerships with regional and country offices and other divisions.

maintained a dedicated knowledge management post, unstable funding prompted the need to do away with the post in 2012. Since then, the division has undertaken stopgap measures to make up for the loss, with other staff members absorbing some knowledge management responsibilities amid their already-challenging emergency responsibilities. However, this is not a sustainable way of ensuring that knowledge generated from various sources is relevant, and the gained knowledge is properly managed and applied to meet the wide range of UNICEF accountability and learning needs. Recent attempts by EMOPS to define the contours of these needs — through a 2012-2013 survey and review of the humanitarian knowledge management function — and to integrate knowledge management within the division's office management plan for 2014-2017, provide opportunities to embed the humanitarian knowledge management function within the organization.

58. The loss of the knowledge management function speaks to a final major gap: the lack of integration of evaluation into a broader knowledge strategy in support of the CCCs. High-quality, results-oriented evaluation rests in part on reliable monitoring data. The organization's under-emphasis of results-oriented evaluation stems in part from a long-standing lack of such data. EMOPS has sought to address this data gap through the Humanitarian Performance Monitoring system, which, after an initially challenging start-up period, has been refined, gradually gaining traction with every new emergency. EMOPS has also established an information management post to equip global cluster coordinators with better information to guide UNICEF and its partners in the field. More broadly, EMOPS has attempted to integrate evaluation into the various policy initiatives it spearheads for strengthening learning and accountability in emergencies — for example, by ensuring that evaluation concerns are incorporated into the SSOPs for Level-3 emergencies — and by including the evaluation perspective in working groups on accountability to affected populations and knowledge management.

59. This momentum must be sustained. The information generated through the Humanitarian Performance Monitoring system is a necessary but insufficient source of results information that meaningful evaluation requires. There is a need for a closer and more collaborative interaction between EMOPS and the Evaluation Office, at the outset of an emergency as well as during non-emergency periods, to ensure that EMOPS is profiting from the Evaluation Office's technical and strategic evaluation expertise. This includes more strategic conversations around the accountability and learning needs likely to be at play in a given emergency or programme. It also entails dedicated technical conversations at the outset of an emergency response or a humanitarian programme to discuss "evaluability", as highlighted in the DFID-UNICEF case study above. All of this requires an Evaluation Office capacity dedicated to provide such strategic guidance beyond management of requested evaluations. It also requires dedicated knowledge management capacity for EMOPS to engage meaningfully in evaluation.

60. Recent policy currents suggest a building momentum for improving the quality of evaluations of humanitarian action and making them strategic and more useful to UNICEF. The elaboration of the next strategic plan, for 2014-2017, has focused attention on monitoring and the role of evaluation in demonstrating results achieved by UNICEF. With humanitarian action likely to be better integrated into the strategic plan than ever before, UNICEF is poised for greater collaboration around humanitarian-related monitoring and evaluation. At the inter-agency level, the Transformative Agenda sets the stage for a more prominent role for monitoring and

evaluation. Seeking to improve coordination in strategic planning and implementation, coupled with improved accountability for performance and to affected populations, the Transformative Agenda promises to increase significantly the demand for more rigorous results-oriented evaluation. Accordingly, UNICEF, together with the sister agencies, has advocated strengthening inter-agency evaluations of large-scale emergency responses that focus on results. Taken together, these policy orientations should enhance the role of evaluation as a tool for learning and for UNICEF meeting its commitments to those affected by humanitarian situations in the years to come.

VI. Recommendations

61. The Evaluation Office offers the following recommendations for management's attention:

(a) **The Division of Human Resources, in consultation with the Evaluation Office and EMOPS, should incorporate key EHA elements into its training of country representatives and deputy representatives, and develop a roster of EHA-qualified monitoring and evaluation staff for surge deployments.** This includes sensitizing managers to the importance of EHA, scaling the level of evaluation effort to the specific accountability and learning needs at hand, and ensuring adequate EHA capacity;

(b) **The Evaluation Office should continue its support for enhancing EHA capacity at decentralized levels.** Capacity-building efforts should be targeted to UNICEF regional offices as well as country offices;

(c) **EMOPS should follow through on its commitment to embed the humanitarian knowledge management function within the organization.** It should consult closely with the Evaluation Office on relevant issues to ensure that evaluation is better integrated into the broad landscape of humanitarian knowledge sources; that fit-for-purpose evaluation is advocated in humanitarian action; that knowledge from evaluations is synthesized and disseminated for maximum uptake; and that adequate capacity is in place to engage fully in evaluations;

(d) **The Evaluation Office should develop, for consideration by senior management, a broad-based, strategic and results-oriented approach to addressing evaluation gaps and providing balanced evaluation coverage to meet UNICEF learning and accountability needs.** This review has identified significant gaps in rigorous assessment of the results of humanitarian action, as well as in evaluation coverage of Level-1 emergencies and regional emergencies. Action is needed to determine what should be evaluated, when, and with what amount of evaluation effort.