Review of the World Food Programme's Humanitarian and Development Effectiveness

2006 - 2011

4 April 2012

Preface

This review provides an independent and evidence-based assessment of the effectiveness of the World Food Programme's (WFP) humanitarian and development programming. The approach to reviewing multilateral organizations was developed under the guidance of the Organisation for Economic Co-operation and Development (OECD)'s Development Assistance Committee Network on Development Evaluation (DAC-EVALNET). It complements the periodic assessments done by the Multilateral Organization Performance Assessment Network of the multilateral organizations' operational effectiveness.

The review was led by the Evaluation Directorate of the Canadian International Development Agency (CIDA), with support from the Policy and Operations Evaluation Department of the Netherlands Ministry of Foreign Affairs. The methodology applied in the review is a structured meta-synthesis of the content of 52 evaluations published by WFP's Office of Evaluation using a common set of assessment criteria derived from the DAC evaluation criteria. It was conducted by a team from CIDA's Evaluation Directorate (Michelle Guertin) and the consulting firm, Goss Gilroy Inc. (Sheila Dohoo Faure, Ted Freeman, Danielle Hoegy, Molly McCreary and Kofi Kobia). The team is grateful to WFP for its helpfulness and useful, practical advice.

This report is designed for use by all members of the DAC-EVALNET and other interested stakeholders. The findings of this review are also presented in a separate report prepared for CIDA, which includes information on CIDA's relationship with WFP. A similar review of the development effectiveness of the United Nations Development Programme has also been conducted and we hope that these two reviews represent the beginning of a series of reviews of multilateral organization effectiveness, based on their own evaluations.

We would like to thank everybody who contributed to this review for their valued input and support.

Ted Kliest   Goberdhan Singh
Senior Evaluator   Director General
Policy and Operations Evaluation Department   Evaluation Directorate
Netherlands Ministry of Foreign Affairs   Canadian International Development Agency

Table of Contents

Figures

Tables

List of Abbreviations

CIDA	Canadian International Development Agency
DAC	Development Assistance Committee
DAC-EVALNET	DAC Network on Development Evaluation
EMOP	Emergency Operation
GHD	Good Humanitarian Donorship
MOPAN	Multilateral Organization Performance Assessment Network
OECD	Organisation for Economic Co-operation and Development
PRRO	Protracted Relief and Recovery Operation
UN	United Nations
WFP	World Food Programme

Executive Summary

Background

This report presents the results of a review of the humanitarian and development effectiveness of the United Nation's (UN) World Food Programme (WFP). WFP is both a humanitarian and development UN agency. It is the world's largest humanitarian agency fighting hunger worldwide. In emergencies, it distributes food where it is needed to save the lives of victims of war, civil conflict and natural disasters. After the cause of the emergency passes, it focuses on development, using food to help communities rebuild their lives. It was established in 1961 by the Food and Agriculture Organization and the United Nations General Assembly.[1] WFP's Strategic Plan for 2008 – 2013 identifies five strategic objectives for the organization:[2]

- Save lives and protect livelihoods in emergencies (Strategic Objective 1);
- Prevent acute hunger and invest in disaster preparedness and mitigation measures (Strategic Objective 2);
- Restore and rebuild lives and livelihoods in post-conflict, post-disaster or transition situations (Strategic Objective 3);
- Reduce chronic hunger and under-nutrition (Strategic Objective 4); and
- Strengthen the capacities of countries to reduce hunger, including through hand-over strategies and local purchase (Strategic Objective 5).

Purpose

The purpose of the review is to provide an independent and evidence-based assessment of the humanitarian and development effectiveness of WFP operations (hereafter referred to as "programs") for use by external stakeholders, primarily bilateral development agencies.

Methodology

The approach and methodology used in this review were developed under the guidance of the Organisation for Economic Co-operation and Development (OECD)'s Development Assistance Committee (DAC) Network on Development Evaluation (DAC-EVALNET). It was developed to fill an information gap regarding the development effectiveness of multilateral organizations. Although these organizations produce annual reports for their management and/or boards, bilateral shareholders were not receiving a comprehensive overview of the performance on the ground of multilateral organizations. This approach complements the organizational effectiveness assessment of the Multilateral Organization Performance Assessment Network (MOPAN).

The methodology does not rely on a particular definition of development effectiveness. The Management Group and the Task Team that were created by the DAC-EVALNET to develop the methodology had previously considered whether an explicit definition of development

---

[1] WFP Website: http://www.wfp.org/about Accessed January 2012
[2] WFP Strategic Plan 2008 – 2013, WFP, p. 1

effectiveness was needed. In the absence of an agreed upon definition of development effectiveness, the methodology focuses on the essential characteristics of developmentally effective multilateral programming, as described below:

- Programming activities and outputs would be relevant to the needs of the target group and its members;
- The programming would contribute to the achievement of development objectives and expected development results at the national and local level in developing countries (including positive impacts for target group members);
- The benefits experienced by target group members and the development (and humanitarian) results achieved would be sustainable in the future;
- The programming would be delivered in a cost efficient manner;
- The programming would be inclusive in that it would support gender equality and would be environmentally sustainable (thereby not compromising the development prospects in the future); and
- The programming would enable effective development by allowing participating and supporting organizations to learn from experience and use of performance management and accountability tools, such as evaluation and monitoring to improved effectiveness over time.

Based on the above-mentioned characteristics, the review's methodology uses a common set of assessment criteria derived from the DAC's evaluation criteria (Annex 1). It was pilot tested during 2010 using evaluation material from the Asian Development Bank and the World Health Organization. In June 2011, the approach and methodology were endorsed by the members of the DAC-EVALNET as an acceptable approach for assessing the development effectiveness of multilateral organizations.

Since WFP is the first humanitarian assistance agency examined using the common approach, prior to beginning the review process, the review team examined the criteria for this review to ensure that they would capture the effectiveness of all of WFP programming. As a result, the definitions used in the criteria were expanded to capture elements of emergency and protracted relief and recovery operations.

The methodology used in the review is a structured meta-synthesis of the content of 52 evaluations conducted by the Office of Evaluation and published by WFP between 2006 and 2011 (Annex 2).[3] The sampling process is described in further detail in Annex 3. After being screened for quality (the approach and criteria used are described in Annex 4), each evaluation was reviewed to identify findings relating to six main criteria for assessing the humanitarian and development effectiveness:

- The Achievement of Humanitarian and Development Objectives and Expected Results;
- Cross Cutting Themes (Environmental Sustainability and Gender Equality);

---

[3] It should be noted that because the evaluations were conducted between 2006 and 2011, they reflect programming that occurred before or during that period. WFP staff indicated that the organization had made a number of changes to address the issues raised in its evaluations conducted and their inputs are provided in Annex 8.

- Sustainability of Results/Benefits;
- Relevance of Interventions;
- Efficiency; and,
- Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness.

These six criteria were assessed using 19 sub-criteria that are considered to be essential elements of effective humanitarian assistance and development (Annex 1). Findings for each of these were classified, by the review team, using a four-point scale: "highly satisfactory," "satisfactory," "unsatisfactory" and "highly unsatisfactory." Classification of findings was guided by a grid with specific instructions for each rating across all sub-criteria (Annex 5). In order to contextualize the results of the review, a separate analysis of major WFP corporate documents was conducted (see Annex 5 for a list of documents consulted).

The percentages shown in this report are based on the total number of evaluations that addressed the sub-criterion. However, coverage of the sub-criteria in the evaluations reviewed varies from strong to weak. When the coverage is weak, the results have not been presented.

Key Findings

Evaluations Report that WFP is Achieving Most of its Objectives and Expected Results

The review of WFP's evaluations since 2006 indicates positive results with respect to the achievement of humanitarian and development objectives and expected results, based on generally strong coverage of the sub-criteria in the evaluations. Over three-quarters of WFP evaluations reported findings of satisfactory or better for all four sub-criteria (77% for sub-criterion 1.1 "Programs and projects achieve stated objectives"; 92% for sub-criterion 1.2 "Positive benefits for target group members"; 82% for sub-criterion 1.3 "Substantial numbers of beneficiaries/contribution to national humanitarian and development goals"; and 81% for sub-criterion 1.4 "Significant changes in national humanitarian and development policies/programs"). WFP's most cited achievements were in the distribution of food aid and other programming, which contributed to improvement of food consumption, prevention of acute hunger and reduction of the risk of chronic hunger and malnutrition for target populations. A key factor contributing to WFP effectiveness is its strong logistics capacity. WFP also achieved its objectives and expected results when it engaged in policy dialogue with and provided support to developing country governments for the development of national humanitarian and development policies and programs. Program interruptions due to financial issues were most likely to detract from the achievement of objectives and expected results.

WFP Faces Challenges in Promoting Gender Equality

WFP has not been particularly effective in addressing the cross cutting theme of gender equality. WFP's programming, in just over half (56%) of the 32 evaluations that addressed this sub-criterion, was given a rating of satisfactory or better with respect to its effectiveness in addressing gender equality (2.1). Successes, however, were noted in the areas of the provision of sex-disaggregated data and some improvements for women and girls. Few common themes

with respect to contributing factors on gender equality were identified. However, key factors in explaining the results were identified in the "End-of-Term Evaluation of WFP's Gender Policy (2003-2007)."[4] The evaluation indicates that the policy had contributed to enhancing "visibility of women and girls not just as vulnerable beneficiaries but as critical contributors to household and community food security." However, WFP's evaluation also noted that there were issues with reporting on the impact of gender equality on effectiveness since the indicators used measured outputs rather than outcomes. WFP adopted a new gender policy in 2009 that includes the following goals: to address the challenges of restoring the gender mainstreaming mandate, complete the transition from women's programming to gender mainstreaming, re-orient the institutional approach to enable context-led approaches and marshal and allocate resources at the operational level.

Environment Sustainability – No Results to Report

Only five evaluations reported on environmental sustainability (2.2) as a cross cutting theme, which prevented the review from identifying any results in this area. However, WFP recently prepared a note for its Executive Board that reported on its consultations with experts and partners and outlined how WFP has addressed the impact of climate change on hunger to date. It describes its plans to address climate change in the future and reflects the commitment to finalize a climate change policy by the end of 2012.[5] Although WFP had already prepared a paper for the Board in 1998, which recognized the importance of environmental issues and the potential impact of WFP programming on the environment,[6] the assessment of environmental impacts is not yet identified in WFP's Monitoring and Evaluation Guidelines and, as a result, evaluations are not assessing the environmental impact of WFP programming.[7]

Sustainability of Benefits/Results is a Challenge

Some care is needed in interpreting the results for sustainability of benefits/results. Only one sub-criterion received strong coverage (3.1). One received moderate coverage (3.2) and the results for the third (3.3) are not presented because of weak coverage. Overall, the findings on sustainability of the benefits/results of WFP programming represent an area for improvement for the organization. Fewer than half (43%) of the evaluations reported findings of satisfactory or better with respect to likelihood that benefits would continue (3.1). In two-thirds (64%) of the evaluations, WFP's contribution to building institutional and community capacity for sustainability was rated as satisfactory or better (3.2). Sustainability faced particular challenges when WFP: lacked time and resources for interventions; did not engage with developing country governments; and was short of early attention to planning the shift from relief to recovery.

WFP Programs are Highly Relevant to their Context

---

[4] "Summary Report of the End-of-Term Evaluation of WFP's Gender Policy (2003-2007)", WFP, 2008, p.10
[5] "Climate Change and Hunger: Towards a WFP Policy on Climate Change", WFP, 27 April 2011
[6] "WFP and the Environment: Issues and Priorities", WFP, 3 September 1998
[7] "Monitoring and Evaluation Guidelines", WFP http://www.wfp.org/content/monitoring-and-evalutation-guidelines Access January 2012

The relevance of WFP's interventions was rated best among the six humanitarian and development effectiveness criteria and was well covered in WFP's evaluations. Eighty-six percent (86%) of the evaluations reported findings of satisfactory or better with respect to the relevance of WFP's programming to target group needs (4.1) and goals of the developing country governments (4.2). Additionally, 84% of evaluations reported positive findings on effective partnerships (4.3). The important factors contributing to the positive results related to the relevance of WFP's interventions are the quality of WFP's needs assessments and the nature of the arrangements of the partnerships, created by forming strong links with developing country governments, other multilateral agencies, civil society and, to some extent, donors.

Evaluations Report Mixed Results in Efficiency

There was moderate to strong coverage of the efficiency of WFP's programming in its evaluations. However, efficiency represents another challenge for the organization in some areas. Nearly two-thirds (63%) of the evaluations reported findings of satisfactory or better with respect to WFP's systems and procedures for efficient program implementation (5.3). Fifty-eight percent (58%) of the evaluations reported findings of satisfactory or better on cost efficiency (5.1), but only 42% of the evaluations reported positive findings for the timely achievement of objectives (5.2). Frequent financial issues (e.g., lack or unpredictability of donor funding) contribute to negative findings in the area of efficiency and timeliness of WFP's programming. Furthermore, efficiency is affected by logistics and transportation setbacks, and the need to serve very dispersed populations. Effective management of logistics and the WFP's greater use of local resources were identified as factors contributing to efficiency.

Good Use of Evaluation, but Inadequate Performance Frameworks and Weak Monitoring

Results in this area should be interpreted with care, given that coverage for only one sub-criterion (6.2) was strong. The coverage for two sub-criteria (6.1 and 6.4) was moderate. The coverage for one sub-criterion (6.3) was weak and, as a result, no results have been reported. While WFP receives positive ratings for systems for evaluation (6.1) its use of evaluation (6.4) (66% and 91% of the evaluations reported findings of satisfactory or better, respectively), monitoring and reporting on results (6.2) is a challenge (32% of the evaluations reported findings of satisfactory or better). The factors that detracted from the achievement of effective evaluation and monitoring included insufficient capacity both within WFP and within its partner organizations for monitoring and inadequate performance frameworks for programming. To some extent this is also affected by the lack of sufficient resources for monitoring – both for adequate staff training and travel to the field.

Conclusions: Humanitarian and Development Effectiveness of WFP

Evaluations carried out since 2006 indicate that WFP's programming is relevant to the needs of the target groups and developing country governments. The evaluations also reflect effectiveness in the achievement of humanitarian and development objectives and expected results. This is an important achievement considering the complex environment in which WFP

operates. While programming has been effective, WFP has not been as effective in addressing the cross cutting theme of gender equality and ensuring the sustainability of its development results. The cost efficiency and timeliness of its programming also represent areas for improvement for the organization. While WFP made effective use of evaluations to improve program effectiveness, these evaluations are often confronted with issues arising from the quality of the monitoring information available and the adequacy of the performance frameworks. A broad challenge to WFP's effectiveness is the issue of program interruptions, which are caused either by the lack of predictable donor funding, poor contingency planning and/or inefficiencies in WFP operations.

Areas for Attention

This analysis of the strengths and weaknesses of the effectiveness of WFP's humanitarian and development programming led to the identification of areas that required ongoing attention from WFP, in collaboration with its bilateral partners, including:

- Addressing the interruptions of programming that detract from WFP's performance;
- Promoting gender equality;
- Following through on WFP's identified commitments to address environmental issues;
- Strengthening the sustainability of its programming; and
- Strengthening results performance measurement.

1.0 Introduction

1.1 Background

This report presents the results of a review of the humanitarian and development effectiveness of the United Nation's (UN) World Food Programme (WFP). The common approach and methodology were developed under the guidance of the Organisation for Economic Co-operation and Development (OECD)'s Development Assistance Committee (DAC) Network on Development Evaluation (DAC-EVALNET). The review relies on the content of published evaluation reports produced by WFP supplemented with a review of WFP corporate documents.

The method uses a common set of assessment criteria derived from the DAC's evaluation criteria (Annex 1). It was pilot tested during 2010 using evaluation material from the Asian Development Bank and the World Health Organization. The overall approach and methodology were endorsed by the members of the DAC-EVALNET as an acceptable approach for assessing the development effectiveness of multilateral organizations in June 2011. This review is among the first to be implemented following the DAC-EVALNET's endorsement.

From its beginnings, the process of developing and implementing the reviews of development effectiveness has been coordinated with the work of the Multilateral Organization Performance Assessment Network (MOPAN). By focusing on development effectiveness and carefully selecting assessment criteria, the reviews seek to avoid duplication or overlap with the MOPAN process.

1.2 Why Conduct this Review?

The purpose of the review is to provide an independent, evidence-based assessment of the humanitarian and development effectiveness of WFP operations (hereafter referred to as "programs") for use by external stakeholders, primarily bilateral development agencies.

The current approach to assessing the development effectiveness of multilateral organizations was developed in order to address a gap in the information available to bilateral development agencies. While the Multilateral Organization Performance Assessment Network (MOPAN) provides regular reports on the organizational effectiveness of multilateral organizations, it does not yet report on development effectiveness and has not addressed the information gap this review is meant to fill. Other options such as large-scale, joint donor-funded evaluations of a given multilateral organization are much more time-consuming, more costly and result in a significant management burden to the organization being evaluated before, during and after the evaluations. The current approach is intended to work in a coordinated way with initiatives such as the DAC-EVALNET/United Nations Evaluation Group (UNEG) Peer Reviews of multilateral organization evaluation functions. It also recognizes that multilateral organizations continue to make improvements and strengthen their reporting on development effectiveness. The ultimate aim of the approach is to be replaced by regular, evidence-based, field-tested reporting on development effectiveness provided by multilateral organizations themselves.

1.3 WFP: A Global Organization Committed to Fighting Hunger

1.3.1 Background and Objectives

The WFP is both a humanitarian and development UN agency. It is the world's largest humanitarian agency fighting hunger worldwide. In emergencies, it distributes food where it is needed to save the lives of victims of war, civil conflict and natural disasters. After the cause of the emergency passes, it focuses on development, using food to help communities rebuild their lives.[8] It was established in 1961 by the Food and Agriculture Organization and the United Nations General Assembly.[9] In recent years, WFP has shifted its delivery mechanisms from "food aid" (focus on the shipment of food) to "food assistance" (including cash transfers, rather than physical food, where appropriate).[10] WFP has also taken a leadership role in the global humanitarian coordination cluster system for global logistics and emergency telecommunications.[11]

1.3.2 Strategic Plan

WFP's Strategic Plan for 2008 – 2013 identifies five strategic objectives for the organization:[12]

- Save lives and protect livelihoods in emergencies (Strategic Objective 1);

- Prevent acute hunger and invest in disaster preparedness and mitigation measures (Strategic Objective 2);

- Restore and rebuild lives and livelihoods in post-conflict, post-disaster or transition situations (Strategic Objective 3);

- Reduce chronic hunger and under-nutrition (Strategic Objective 4); and

- Strengthen the capacities of countries to reduce hunger, including through hand-over strategies and local purchase (Strategic Objective 5).

1.3.3 Work and Geographic Coverage

According to its most recent annual report, in 2010, WFP provided food to 109.2 million beneficiaries in 75 countries.[13]

---

[8] WFP website, http://www.wfp.org/about Accessed January 2012
[9] WFP website, http://www.wfp.org/about/corporate-information/history Accessed January 2012
[10] WFP website, http://www.wfp.org/content/revolution-food-aid-food-assistance-innovations-overcoming-hunger, Accessed January 2012
[11] WFP Executive Board Agenda Item 5, January 2010, *WFP's Role in the Humanitarian Assistance System*, WFP/EB.1/2010/5-C.
[12] *WFP Strategic Plan 2008-2013*, WFP, p. 1
[13] "Annual Performance Report for 2010", WFP, WFP/EB.A/2011/4, 13 May 2011, p.75

Table 1: Direct Expenses, by Category, 2007 – 2010

Category	2007		2008		2009		2010	
	US$ 000	%	US$ 000	%	US$ 000	%	US$ 000	%
Development	309,318	11	292,112	8	275,906	7	287,842	7
Relief	2,005,656	73	2,733,744	77	3,239,887	81	3,220,081	80
Emergency**	716,411	(36)	944,581	(35)	1,418,385	(44)	1,660,195	(52)
PRRO**	1,289,245	(64)	1,789,163	(65)	1,821,502	(56)	1,559,885	(48)
Special Operations	166,244	6	200,252	6	176,364	4	221,510	6
Other***	272,090	10	309,639	9	293,457	7	270,898	7
Total	**2,753,308**	**100**	**3,535,746**	**100**	**3,985,613**	**100**	**4,000,330**	**100**

Source: "Annual Performance Report for 2010", World Food Programme
*   Excludes programme support and administrative costs.
**   Percentages are as percent of total Relief.
*** Operational Expenses, including General Fund, Special Accounts and Trust Funds that cannot be apportioned by project/operation.

WFP's programming is divided into three main categories:

- Relief:

  o Emergency – emergency food aid provided through WFP's Immediate Response Account and, as warranted, a subsequent Emergency Operation (EMOP) that can assist populations in need by either food distributions or other projects such as food aid in exchange for reconstruction work;

  o Protracted Relief and Recovery Operations (PRRO) – carried out in protracted crises to help sustain affected communities as they re-establish livelihoods and stabilize food security. A PRRO is formed when it becomes clear that the 24-month assistance provided under an EMOP will be insufficient and includes such activities as food for education and training, extended relief, relief for refugees and food for recovery (e.g., food-for-assets programs); and

- Development – development food aid is used to help the chronically food insecure escape from the poverty trap. WFP ensures its aid is concentrated on pre-identified, food-insecure areas inside recipient countries – usually rural areas of low productivity, areas prone to natural disasters and areas vulnerable to periodic food shortages. It also works in towns and cities with high concentrations of malnutrition;

- Special Operations – short-term operations designed to speed-up the movement of food aid and involve logistics and infrastructure work (e.g., infrastructure repairs, intermittent airlifts and provision of common logistics) designed to overcome operational bottlenecks.[14]

As shown in Table 1, while WFP's expenses have remained fairly constant, with a slight increase in total expense over the past four years, there has been some shift in the category of these expenses. Development expenses have remained relatively stable. The majority of WFP's resources was disbursed in relief operations. Overall, the percentage of WFP funding addressing relief operations has increased from 73% in 2007 to 80% in 2010. However, the percent of relief expenses going to emergency responses has risen from 36% to 52% over the four years. The percent of relief funding expended for PRROs has correspondingly declined from 64% to 48%.

1.3.4 Evaluation and Results Reporting

The Office of Evaluation is headed by a Director and reports directly to the Executive Board. A new WFP Evaluation Policy, approved in 2008, focuses on the use of evaluation for both accountability and learning.[15] WFP's evaluation function conducts five different types of evaluations: strategic, policy, country portfolio, impact, and operations evaluations (Figure 1).

Figure 1: Types of WFP Evaluations[16]

Strategic Evaluations take a global perspective and focus on strategic issues that are central to strategic decision-making and related actions that affect more than one operation or one country office. An example of a recent strategic evaluation included in this review is "Strategic Evaluation of WFP's Role in Social Protection and Safety Nets."

Policy Evaluations examine how effective WFP policy has been in achieving its objectives, including the quality and results of a given policy and the factors that explain why these results occurred. An example of a recent policy evaluation included in this review is "End-of-Term Evaluation of WFP's Gender Policy (2003-2007)."

Country Portfolio Evaluations assist a country office to understand their portfolio and what difference it makes. They include all WFP's work in one country and normally cover a five-year period. An example of a recent country portfolio evaluation included in this review is "Yemen: An evaluation of WFP's portfolio (2006-2010)."

Impact Evaluations provide an in-depth analysis of the impact that WFP work has on beneficiaries. They focus on one program activity in one country at a time. An example of a recent impact evaluation included in this review is "WFP Cambodia School Feeding 2000-2010: A Mixed Method Impact Evaluation."

Operations Evaluations assess the relevance, efficiency and effectiveness of an operation and are normally undertaken by country office. The headquarters' evaluation unit does large operations evaluations and provides technical support and training to country office staff. An example of a recent operations evaluation included in this review is "occupied Palestinian territory: An Operation Evaluation."

---

[15] "WFP Evaluation Policy", WFP, 8 October 2008, p. 5
[16] WFP website, http://www.wfp.org/about/evaluation Accessed January 2012.

The results of WFP's evaluations are synthesized in an annual evaluation report that is presented to the Executive Board. The 2010 report summarized the results of twenty evaluations completed in 2010 that covered 13% of WFP's global program of work.[17] The report also emphasized the ongoing commitment of the Office of Evaluation to improve the quality, credibility and usefulness of evaluations, through increasing its human resource capacity and evaluation tools, notably the Evaluation Quality Assurance System, and the learning opportunities from evaluations.

A 2007 Peer Review of the Evaluation Function at WFP concluded that the credibility of the products of the evaluation function at that point was "uneven."[18] However, it also noted that the central evaluation office was "a strong unit, with committed, well trained and highly motivated staff."[19] The peer review panel also reported that if WFP were to implement the changes planned for its evaluation function, it would adequately address the findings of the peer review.

The review team conducted its own quality review of the evaluations included in this review. The results were positive, with 90% of the reviewed evaluations scoring 31 points or more out of a possible total of 45. (See Annex 3 for details of the review methodology and Annex 4 for the evaluation quality scoring grid.)

2.0 Methodology

This section describes briefly the main elements of the methodology used for the review. A more detailed description of the methodology is presented in Annex 3.

2.1 Rationale

The term "common approach" describes the use of a standard methodology, as implemented in this review, to assess consistently the development and, in this case, humanitarian, effectiveness of the multilateral organizations. It offers a rapid and cost effective way to assess the development effectiveness of the organization relative to a more time consuming and costly joint evaluation.[20] The approach was developed to fill an information gap regarding the development effectiveness of multilateral organizations. Although these organizations produce annual reports for their management and/or boards, bilateral shareholders were not receiving a comprehensive overview of the performance on the ground of multilateral organizations. MOPAN seeks to address this issue through organizational effectiveness assessments. This approach complements MOPAN's assessments.

---

[17] "Annual Evaluation Report 2010", WFP, 19 May 2011
[18] "Peer Review: Evaluation Function at the World Food Programme", Sida, 2007
[19] Ibid., p. 11
[20] "Joint evaluation" refers to a jointly funded and managed comprehensive institutional evaluation of an MO. It does not refer to DAC/ United Nations Evaluation Group Peer Reviews of the Evaluation Function.

The approach suggests conducting a review based on the organization's own evaluation reports when two specific conditions exist:[21]

1. There is a need for field tested and evidence-based information on the effectiveness of the multilateral organization; and

2. The multilateral organization under review has an evaluation function that produces an adequate body of reliable and credible evaluation information that supports the use of a meta-evaluation methodology to synthesize an assessment of the organization's effectiveness.

The evaluation function at WFP does produce enough evaluation reports of good quality to support an assessment of the humanitarian and development effectiveness of WFP. The second condition is also satisfied, as WFP's existing reporting mechanisms do not provide sufficient information on the organization's humanitarian and development effectiveness. Further details on the rationale are available in Annex 3.

2.2 Scope

The Office of Evaluation (OE) website at WFP (www.wfp.org/about/evaluation) identifies 52 evaluation reports conducted by the Office of Evaluation and published between 2006 and August 2011.[22] These included only evaluations conducted by the Office of Evaluation and did not include any decentralized evaluations conducted by Country Offices. This group of 52 evaluations forms the evaluation universe for this review. Rather than sample from a small population, the reviewers selected all 52 for the conduct of this review (described in more detail in the Annex 3).

While some earlier evaluations cover operations before 2006, most activities evaluated occurred during the period covered by this review (2006 – 2011). Annex 2 provides the list of evaluations included in the review sample.

The evaluations covered all priority areas of WFP programming, including development and relief (emergency and protracted relief and rehabilitation operations). However, not all types of WFP programming are covered equally by the evaluations. Only four evaluations in the sample covered, in whole or in part, emergency programming; whereas emergency relief has accounted for about one-third of WFP direct expenses between 2007 and 2010 (see Table 1 in Section 1.3.3). The evaluations reviewed covered countries accounting for 67% of all WFP program expenditures in the 2006 to 2009 period (the latest biennial budget period for which expenditure

---

[21] *Assessing the Development Effectiveness of Multilateral Organizations: Approach, Methodology and Guidelines*, Management Group of the Task Team on Multilateral Effectiveness, DAC EVALNET, 2011.

[22] The original list from WFP's website included 60 evaluations. However, on closer review, there were a number of duplications in the listing mainly caused by reports being published in different languages. In addition, some evaluations were pure process evaluations and did not include findings focused on humanitarian and development effectiveness.

data is available).[23] They also accounted for some programming in eight of WFP's 10 largest countries of operations.

The review of evaluation reports was supplemented by a review of WFP corporate documents.[24] These were done to contextualize the results of the review. A list of the documents consulted is provided in Annex 6. In addition, following the presentation of this review to WFP staff, they indicated that the organization had made a number of changes to address the issues raised in its evaluations conducted between 2006 and 2011. Their inputs are provided in Annex 8.

2.3 Criteria

The methodology does not rely on a particular definition of development effectiveness and, in this case humanitarian assistance effectiveness. The Management Group and the Task Team that were created by the DAC-EVALNET to develop the methodology had previously considered whether an explicit definition of development effectiveness was needed. In the absence of an agreed upon definition of development effectiveness, the methodology focuses on some of the essential characteristics of developmentally effective multilateral organization programming, as described below:

- Programming activities and outputs would be relevant to the needs of the target group and its members;
- The programming would contribute to the achievement of development objectives and expected development results at the national and local level in developing countries (including positive impacts for target group members);
- The benefits experienced by target group members and the development (and humanitarian) results achieved would be sustainable in the future;
- The programming would be delivered in a cost efficient manner;
- The programming would be inclusive in that it would support gender equality and would be environmentally sustainable (thereby not compromising the development prospects in the future); and
- The programming would enable effective development by allowing participating and supporting organizations to learn from experience and use of performance management and accountability tools, such as evaluation and monitoring to improve effectiveness over time.

The review methodology, therefore, involves a systematic and structured review of the findings of WFP evaluations, as they relate to six main criteria (described in text box below) and 19 sub-criteria that are considered to be essential elements of effective humanitarian and development (Annex 5). The main criteria and sub-criteria are derived from the DAC evaluation criteria, with

---

[23] It should be noted that this is based on the assumption, which is known to not always be the case, that an evaluation in a given country covers all the WFP programming in that country.
[24] The reviewers note that future reviews could benefit from interviews to provide context and additional information.

further refinements made during the pilot review of the Asian Development Bank and the World Health Organization to develop the common approach and methodology.

Since WFP is the first humanitarian assistance agency examined using the common approach, prior to beginning the review process, the review team examined the criteria and sub-criteria to ensure that they would capture the effectiveness of all three categories of WFP programming. Therefore, the definitions used in the criteria were expanded to capture elements of emergency and protracted relief and recovery operations. To the extent that this is the first application of this approach to humanitarian programming, it is, in fact, a pilot of these modified criteria and sub-criteria.

2.4 Limitations

As with any meta-evaluation, there are methodological challenges that limit the findings. For this review, the limitations include sampling bias, the challenge of assessing overall programming effectiveness when the evaluations cover multiple programming components or only a specific theme or project as part of a program area and the inability to draw conclusions by type of programming.

Any possible sampling bias was eliminated by including all relevant evaluations published during the period, although the coverage of WFP's evaluations may not be representative of the categories of its programming. In addition, there was adequate coverage of the criteria since 16 of the 19 sub-criteria used to assess humanitarian and development effectiveness are well covered in the evaluations reviewed (Annex 3). Three received a weak coverage rating.

A problem arises in interpreting the results of WFP's evaluations that covered multiple operations because these evaluations usually cover very different types of programming in the country. Arriving at an overall finding for the WFP program in a given country requires an overall assessment combining results from different program areas. In most instances, the evaluation report makes an effort to merge results across the main program areas into a single finding on each of the sub-criteria. Where these evaluations did not make such an overall judgment, the reviewers compiled evidence from all program areas and made their own assessment.

Finally, the review was not able to distinguish the effectiveness of WFP's programming by type of programming. Therefore, the report does not differentiate the effectiveness of humanitarian, compared to development, programming. It also does not distinguish the effectiveness reported in different types of WFP evaluations (i.e., impact, country portfolio, operational or

policy/strategic/thematic evaluations) for the following reasons: there were not sufficient evaluations available for inclusion in the meta-synthesis to allow for them to be analyzed separately; some evaluations cover multiple types of programming (e.g. emergency relief and country program) in the same evaluation; and WFP's programming does not always make a distinction between humanitarian and development programming (e.g., PRROs combine elements of extended humanitarian relief, food for work and early recovery).

3.0 Findings on WFP's Humanitarian and Development Effectiveness

This section presents the results of the review as they relate to the six main criteria and their associated sub-criteria (Table 2 and Annex 1). In particular, Table 2 below describes the ratings assigned by the review team of "satisfactory" or "unsatisfactory" for each of the six major criteria and their associated sub-criteria. The table also presents the number of evaluations that addressed each sub-criterion (represented by the letter *a*).[25]

Each of the following sections begins with a summary of the coverage and key findings and follows with the main factors contributing to these results. The importance of positive and negative factors contributing to results under each assessed area is described by quantifying how many evaluations identified a particular factor.

Table 2: Percentage of Evaluations Reporting Findings of Satisfactory and Unsatisfactory for each Sub-Criterion, and Number of Evaluations Addressing each Sub-Criterion

Criteria and Sub-Criteria	*a**	Satisfactory Ratings (%)**	Unsatisfactory Ratings (%)**
Achieving humanitarian and development Objectives and Expected Results			
1.1 Programs and projects achieve their stated humanitarian and development objectives and attain expected results.	52	77%	23%
1.2 Programs and projects have resulted in positive benefits for target group members.	50	92%	8%
1.3 Programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national humanitarian and development goals.	45	82%	18%
1.4 Programs contributed to significant changes in national humanitarian and development policies and programs (including for disaster preparedness, emergency response and rehabilitation) (policy impacts) and/or to needed system reforms.	31	81%	19%
Cross-Cutting Themes – Inclusive Humanitarian and Development which is Sustainable			
2.1 Extent to which multilateral organization supported activities effectively address the cross-cutting issue of gender equality.	32	56%	44%

[25] *a*= the number of evaluations that addressed the sub-criteria; *n*= the number of evaluations in the sample

Criteria and Sub-Criteria		a*	Satisfactory Ratings (%)**	Unsatisfactory Ratings (%)**
2.2	Extent to which changes are environmentally sustainable.	5	NA	NA
Sustainability of Results/Benefits				
3.1	Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations, to rehabilitation, reconstructions and, eventually, to longer term humanitarian and development results.	42	43%	57%
3.2	Projects and programs are reported as sustainable in terms of institutional and/or community capacity.	41	64%	36%
3.3	Programming contributes to strengthening the enabling environment for humanitarian and development.	10	NA	NA
Relevance of Interventions				
4.1	Programs and projects are suited to the needs and/or priorities of the target group.	51	86%	14%
4.2	Projects and programs align with national humanitarian and development goals.	46	93%	7%
4.3	Effective partnerships with governments, bilateral and multilateral development and humanitarian organizations and Non-governmental organizations for planning, coordination and implementation of support to development and/or emergency preparedness, humanitarian relief and rehabilitation efforts.	51	84%	16%
Efficiency				
5.1	Program activities are evaluated as cost/resource efficient.	38	58%	42%
5.2	Implementation and objectives achieved on time (given the context, in the case of humanitarian programming).	40	42%	58%
5.3	Systems and procedures for project/program implementation and follow up are efficient (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.).	41	63%	37%
Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness				
6.1	Systems and process for evaluation are effective.	39	66%	34%
6.2	Systems and processes for monitoring and reporting on program results are effective.	50	32%	68%
6.3	Results based management systems are effective.	11	NA	NA
6.4	Evaluation is used to improve humanitarian and development effectiveness.	35	91%	9%

*a = number of evaluations addressing the given sub-criterion
** Satisfactory ratings includes "satisfactory" and "highly satisfactory"; unsatisfactory ratings includes "unsatisfactory" and "highly unsatisfactory"

3.1 Evaluations Report that WFP is Achieving Most of its Objectives and Expected Results

3.1.1 Coverage

There was strong to moderate coverage of the various sub-criteria for this criterion (Figure 2). Sub-criterion 1.1 "Programs and projects achieve stated objectives" was addressed by all 52 evaluation reports in the sample. Similarly, sub-criteria 1.2 "Positive benefits for target group members" and 1.3 "Substantial numbers of beneficiaries/contribution to national humanitarian and development goals" were rated strong in coverage, with only two and seven evaluations, respectively, not addressing these sub-criteria. Sub-criterion 1.4 "Significant changes in national humanitarian and development policies/programs" was rated as moderate in coverage since 21 evaluations did not address this sub-criterion.

Figure 2: Number of Evaluations Addressing Sub-criteria for Humanitarian and Development Objectives and Expected Results
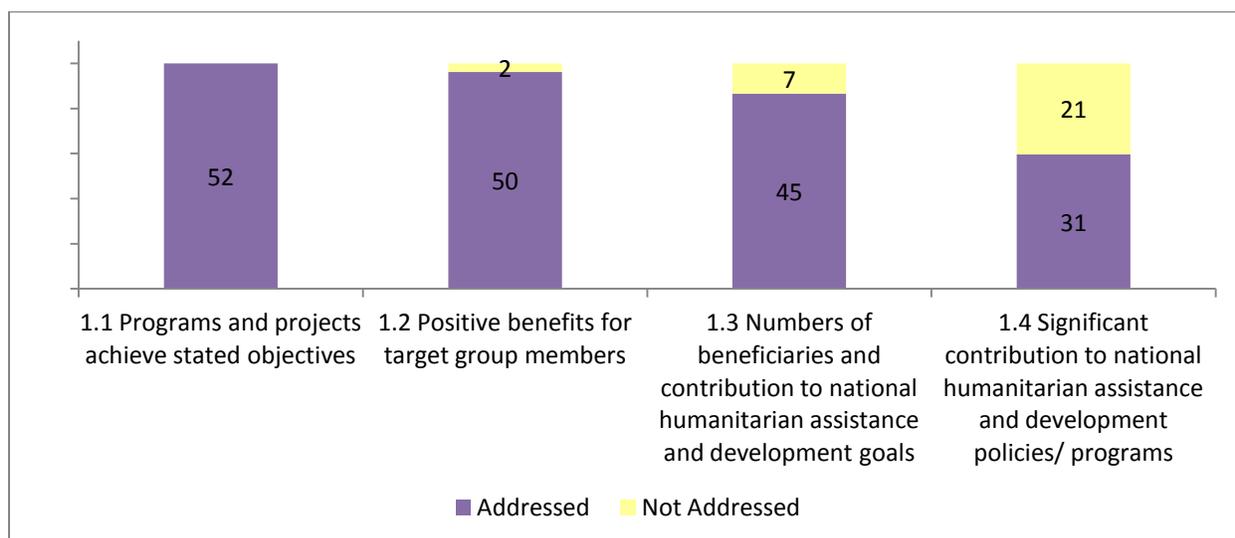
3.1.2 Key Findings

In terms of achieving humanitarian and development objectives and expected results, WFP programs were found to be effective, an important finding, given the challenging context of many of WFP's operations. While the rating for sub-criterion 1.1 "Programs and projects achieve stated objectives" are positive (77% satisfactory or better), the ratings for sub-criterion 1.2 "Positive benefits for target group members" were higher (92% satisfactory or better) (Figure 3). The rating for sub-criterion 1.3 "Substantial numbers of beneficiaries/ contribution to national humanitarian and development goals" and sub-criterion 1.4 "Significant changes in national humanitarian and development policies/programs" are also positive at 82% and 81% satisfactory or better, respectively, with 31% of evaluations reporting findings of highly satisfactory for sub-criterion 1.3. However, there may be some room for improvement regarding sub-criterion 1.1 "Program and projects achieve states objectives", where 23% of evaluations reported findings of unsatisfactory.

Figure 3: Achievement of Humanitarian and Development Objectives and Expected Results (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=52)
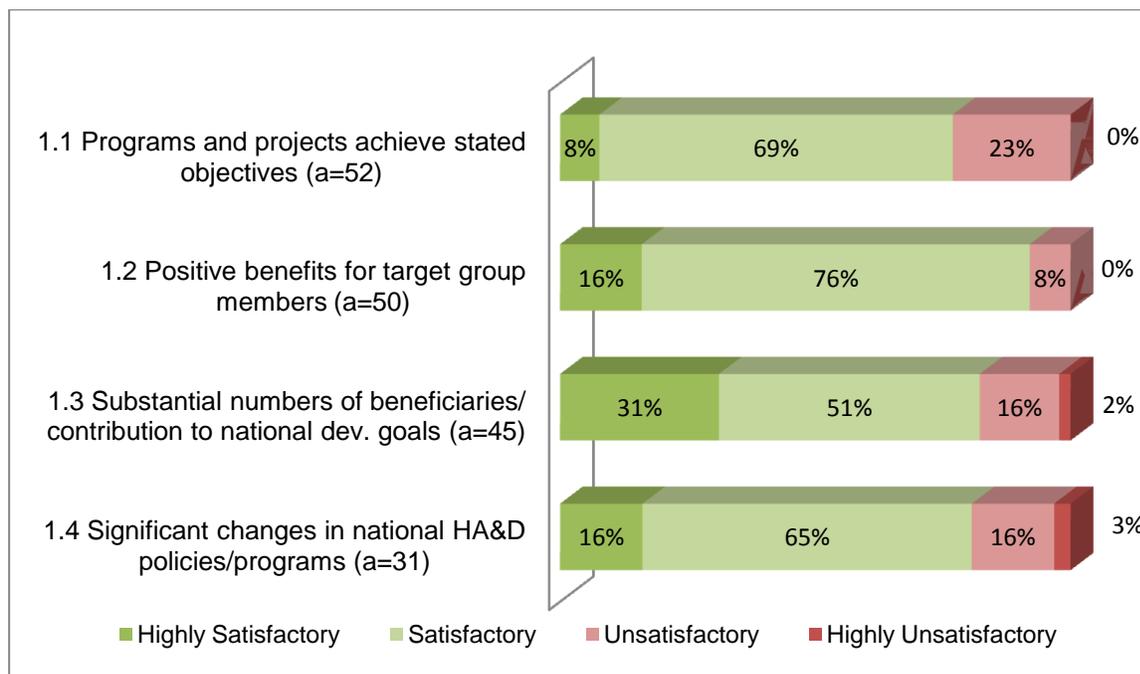
Given the number of evaluations included in the review, it was not possible to analyze separately the achievement of objectives and expected results in each of the different types of WFP evaluations. However, the objectives and expected results are different in nature across the four types of evaluations. For example, the expected results for impact or country portfolio and operational evaluations are defined at the level of beneficiaries within countries or regions. Results in the policy/strategic/thematic evaluations are more global – often providing a global view of WFP's activities. The qualitative information that accompanied the ratings provided a better sense of the achievement of objectives at these two levels.

Specifically, the evaluations emphasized the types of benefits accruing to target populations. These benefits include:

- Improvement of food consumption, prevention of acute hunger and/or reduction of the risk of chronic hunger, and reduction of malnutrition through the distribution of food aid (Highlight box 1) (20 evaluations);[26]

- Educational benefits (e.g., improved enrolment, attendance, retention and learning) and health benefits (e.g., improved nutrition, reduced morbidity) through the implementation of school feeding programs (18 evaluations);

---

[26] The numbers in brackets track the frequency with which analysts in the review team highlighted comments when reviewing evidence to support the findings in a given evaluation. The figures cited do not provide an exhaustive census of how many citations were made but, rather, an overall portrait of the emphasis given in different evaluation reports. They are meant to be illustrative, not a definitive count of occurrence of the factors.

- Creation or rehabilitation of community assets, restoration or protection of livelihoods, increased training from Food for Work/Food for Assets programs (10 evaluations); and

- Improvement of health benefits (e.g., prevention of severe acute malnutrition and excess mortality), including mention of specific populations (pregnant women, lactating mothers, children, people living with HIV/AIDS, tuberculosis patients) (9 evaluations).

The evaluations also reported on WFP's ability to reach all of its target population and/or to expand beyond the expect target population (12 evaluations). A similar number of other evaluations noted that the organization was able to reach a particularly large number of beneficiaries (10 evaluations).

WFP's contributions to national humanitarian and development policies and programs were also highlighted in a number of evaluations. The different areas in which WFP contributed to changes in policy and programs include:

- Increased commitment from developing country governments to food security issues through policy dialogue and WFP information (Highlight box 2) (8 evaluations);

- Improved humanitarian and development policies and programming in a number of areas, such as school feeding, education, nutrition, natural resources/land management, food security, integration of HIV nutrition and food security issues, social protection schemes and emergency management (Highlight box 3) (7 evaluations);

- Improved information systems and capacity building initiatives (e.g., vulnerability analysis and needs assessment) (8 evaluations); and

- Strengthened partnerships with a range of stakeholders, including developing country governments, civil society and other UN agencies (3 evaluations).

Highlight Box 1

Very successful food distribution

"WFP reached 96 percent of the number of beneficiaries specified in the EMOP and 99.5 percent of the number specified in the operational plans. This was a significant achievement given the difficulties of operating in Darfur. Another remarkable achievement was that the loss of cooperating partners in March 2009 had very little impact on the numbers of beneficiaries reached or the tonnages distributed. WFP launched into direct distribution very effectively."

EMOP Sudan: Food assistance to populations affected by conflict, 2010, p. iii

---

<div style="border: 2px solid blue;">

Highlight Box 2

Policy Dialogue

"WFP is recognised for its contribution to awareness raising and policy dialogue that led to the preparation and adoption of the NNP [National Nutrition Policy]. This advocacy work is now taken further by WFP's contribution to the REACH initiative that will help bring about a coordinated institutional response to nutrition issues."

Country Portfolio Evaluation of WFP Assistance to the Lao PDR, 2009, p. 35*.*

</div>

<div style="border: 2px solid blue;">

Highlight Box 3

Impact on Policy Environment

"WFP contributed to generating government commitment to addressing acute and chronic malnutrition and hunger, which created an enabling policy environment for developing organizational and individual capacities. This occurred at the regional and national levels, particularly in disaster preparedness, food fortification, nutrition, safety-net programmes and school feeding."

Capacity Development Policy and Operations (Summary Report), 2008 p. 7

</div>

3.1.3 Contributing Factors

The review cited several factors that either detracted from, or contributed to, WFP's achievement of objectives and expected results. In some cases, evaluations identified factors that contributed to the achievement of objectives as:

- Strong logistics capacity of WFP to procure, transport and distribute food aid, often in very difficult contexts (4 evaluations);

- Limited program interruptions, in one case facilitated by the availability of carryover stocks (4 evaluations); and

- Quality of partnerships with host country governments, other UN agencies or local communities, including the use of participatory approaches in working with communities (Highlight box 4) (4 evaluations).

Evaluations also identified factors that detracted from the achievement of objectives and the ability to provide positive benefits for target group members, such as:

- Financial issues, including shortage of commodities, lack or unpredictability of funding, and donor support, often resulting in program interruptions (18 evaluations);

- Challenges related to the local context, such as the geographic reach (often resulting in resources being spread too thinly), the state of security, and emergency situations (10 evaluations);

- Issues with program design (quality at entry), such as poor targeting, lack of a sufficiently customized design based on strategic choices, lack of clear feasible objectives or prioritization (Highlight box 4) (9 evaluations);

- Issues related to the capacity of developing country governments, including lack of ownership and skills and inability to meet objectives (8 evaluations), or capacity or choice of implementing partners, including the limited choice of partner organizations and lack of technical skills (6 evaluations);

- Delays in the delivery or distribution of food or delays in other components of the program (e.g., settlement of internally displaced people) – sometimes beyond the control of WFP (5 evaluations); and

- Failure to address adequately the long-term needs of affected populations and to integrate relief and recovery operations (4 evaluations).

Highlight Box 4

Factors Affecting Emergency School Feeding Programs

Factors affect the effectiveness of emergency school feeding projects include: appropriate targeting (based on solid information about specific educational needs and barriers to education); complementary inputs from others with respect to both educational and nutrition needs; and effective implementation.

Thematic Evaluation of WFP School Feeding in Emergency Situations (Full Report), 2007

3.2 WFP Faces Challenges with Gender Equality and Environmental Sustainability

3.2.1 Coverage

It is important to treat the results reported for the cross cutting themes with some care. For both the sub-criterion 2.1 "Effectively addresses gender equality" and the sub-criterion 2.2. "Changes area environmental sustainable", several evaluations did not report on these issues, resulting in a moderate (32 evaluations) and a weak (5 evaluations) level of coverage, respectively (Figure 4). Due to this weak level of coverage, results have not been presented for the sub-criterion 2.2.

Figure 4: Number of Evaluations Addressing Sub-criteria for Cross Cutting Themes – Gender Equality and Environmental Sustainability
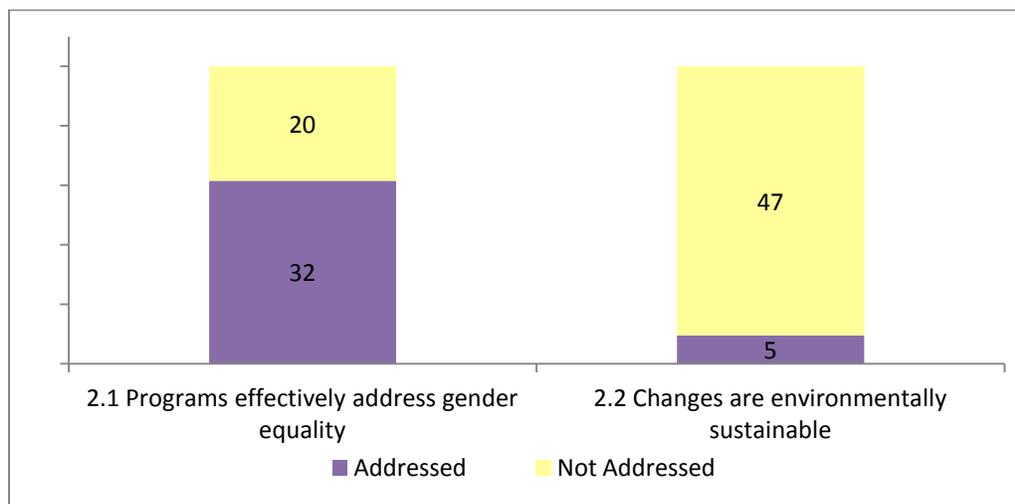
3.2.2 Key Findings

Even though environmental sustainability is not covered in many evaluations, WFP recently prepared a note for its Executive Board that reported on its consultations with experts and partners and outlined how WFP has addressed the impact of climate change on hunger to date. It describes its plans to address climate change in the future and reflects the commitment to finalize a climate change policy by the end of 2012.[27] Although WFP had already prepared a paper for the Board in 1998, which recognized the importance of environmental issues and the potential impact of WFP programming on the environment,[28] the assessment of environmental impacts is not yet identified in WFP's Monitoring and Evaluation Guidelines and, as a result, evaluations are not assessing the environmental impact of WFP programming.[29]

There is better coverage of gender equality in the evaluations. Of those evaluations for which it was possible to rate the sub-criterion 2.1, 56% reported findings of satisfactory or better (Figure 5) suggesting that the WFP programs have some challenges in promoting gender equality.

---

[27] "Climate Change and Hunger: Towards a WFP Policy on Climate Change", World Food Programme, WFP/EB.A/2011/5-F, 27 April 2011

[28] "WFP and the Environment: Issues and Priorities", WFP, 3 September 1998

[29] Monitoring and Evaluation Guidelines, WFP http://www.wfp.org/content/monitoring-and-evalutation-guidelines Accessed January 2012

Figure 5: Cross-cutting Themes – Gender Equality (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=52)
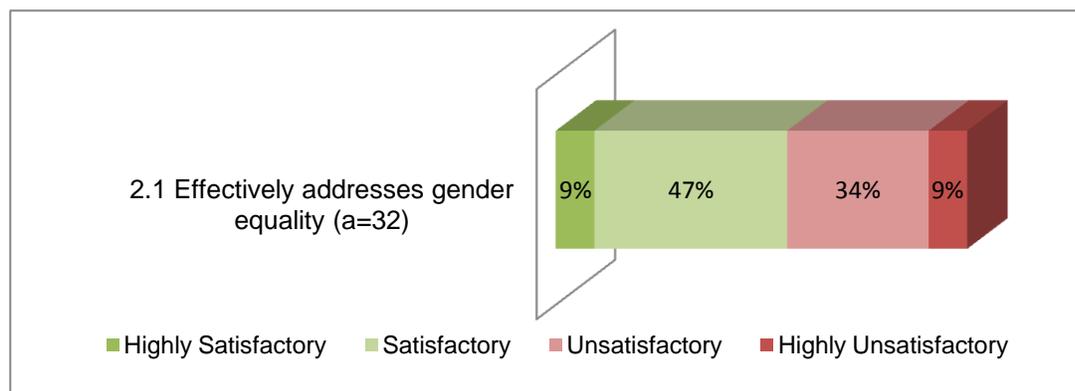
Conflicting observations were identified during the review, where some evaluations noted WFP's focus on gender as being strategic and a serious component of WFP's programming (e.g., referred to the commitment of WFP and/or the national government to make gender a key component of the programming) (12 evaluations), while others noted that gender equality was not integrated as a priority in WFP's programming (8 evaluations).

The most common benefit noted in the evaluations with respect to gender equality was improved results for women and girls (Highlight boxes 6 and 7) (15 evaluations) with respect to:

- Improved educational outcomes, primarily related to enrolment (9 evaluations);

- Increased participation, particularly in decision-making at the community level, including participation on committees responsible for development planning, food distribution, and watershed management (3 evaluations); and

- Increased access to food and improved awareness of health and nutrition issues (3 evaluations).[30]

Achievements in the area of provision of sex-disaggregated data and results of WFP programs were also highlighted in the evaluations (10 evaluations). The evaluations that scored highly satisfactory with respect to this sub-criterion, reported two themes: the promotion of gender issues in programming, and increases in participation, notably increases in enrolment rates for girls.

---

[30] Note that this list does not include all results and the categories are not mutually exclusive.

<div style="border: 2px solid blue; padding: 10px;">

Highlight Box 5

Serious Efforts in Promoting Gender Equality

"The promotion of gender equality and women's empowerment is a pivotal strategic element of the PRRO and echoes one of the national major concern of the Government of Afghanistan, as documented in the [Afghanistan National Development Strategy] and the National Action Plan for Women (NAPW), which aims at bringing about measurable improvements in women's status."

Evaluation of Afghanistan PRRO, 2009, p. 6

</div>

<div style="border: 2px solid blue; padding: 10px;">

Highlight Box 6

Tracking Results for Women and Girls

"In 2007, women formed 52 percent of beneficiaries of GFD [general food distribution], and in 2008 they formed 49 percent of a GFD caseload that was 45 percent more than planned so that the actual number of women beneficiaries was over 50,000 more than planned."

Occupied Palestinian Territory PRRO: September 2007 and August 2009, 2010, p. 51

</div>

WFP implemented a new gender policy in 2009.[31] Prior to the development of that policy, WFP had undertaken an evaluation of its previous Gender Policy (2003-2007).[32] The evaluation reported that, while the policy had some shortcomings, WFP had made progress on the implementation of gender equality by increasing the visibility of women, not only as beneficiaries, but also as contributors to the food security of households and communities. It had also increased the reporting of sex-disaggregated data and the consistent use of monitoring forms requiring sex-disaggregated data. However, it was noted that monitoring and reporting focused more on outputs of gender programming, than outcomes (Highlight box 5). This effectiveness review found similar results reflected in other WFP evaluations. Many of these evaluations covered the period prior to the implementation of the new policy.

---

[31] WFP Gender Policy 2009: Promoting Gender Equality and the Empowerment of Women in Addressing Food and Nutrition Challenges, WFP, 2009

[32] "Full Report of the End-of-Term Evaluation of WFP's Gender Policy (2003-2007): Enhanced Commitments to Women to Ensure Food Security", WFP, August 2008

<div style="border:2px solid blue; padding:10px;">

Highlight Box 7

Measuring the Impact of Gender Equality on Program Effectiveness

"Because the indicators measured outputs rather than outcomes, they do not support conclusions regarding the impact of [the policy] on the lives of women and their families. Some weaknesses stem from general issues: i) attention to outputs rather than outcomes; ii) failure to analyse and learn from collected data; iii) limited institutionalized resources for capacity development; and iv) a focus on delivering food to family representatives, with little attention to distribution in households."

Evaluation of WFP's Gender Policy (2003-2007) (Summary Report), 2008, p.10

</div>

3.2.3 Contributing Factors

Few common themes with respect to the contributing factors on gender equality were identified in the evaluations reviewed. However, a key factor in explaining the results was identified in the "End-of-Term Evaluation of WFP's Gender Policy (2003-2007)." The evaluation indicated that the policy had contributed to enhancing "visibility of women and girls not just as vulnerable beneficiaries but as critical contributors to household and community food security."[33] However, as detailed in the Highlight Box 5, there were issues with reporting on the impact of gender equality on effectiveness since the indicators used measured outputs rather than outcomes. The report concluded, "[t]here were initial political will and resources, and WFP's gender specialists demonstrated extraordinary dedication, but leadership and resources declined. Many believe that the focus on women has been accomplished and that gender mainstreaming is no longer important."[34]

Since the completion of the evaluation, WFP developed a new gender policy that focuses on identifying priorities and actions that build on WFP's strengths, such as its unique field presence and extensive partnerships, incorporating the findings of the recent evaluations, and indicating ways in which WFP can work more constructively to protect women.[35]

3.3   Sustainability of Benefits/Results is a Challenge

3.3.1 Coverage

Sub-criterion 3.1 "Benefits continuing after program completion" was addressed by 42 evaluation reports resulting in a strong level of coverage, while sub-criterion 3.2 "Institutional/community capacity for sustainability" was rated as moderate in coverage since 11 evaluations did not address this sub-criterion (Figure 6). The sub-criterion 3.3 "Strengthened enabling environment for development" was rated as weak in coverage since only 10

---

[33] "Summary Report of the End-of-Term Evaluation of WFP's Gender Policy (2003-2007)", WFP, 2008, p.10
[34] Ibid., p. 11
[35] "Gender Policy: Promoting Gender Equality and the Empowerment of Women in Addressing Food and Nutrition Challenges", WFP, 10 February 2009

evaluations addressed this sub-criterion. Therefore, no results have been presented for this sub-criterion.

Figure 6: Number of Evaluations Addressing Sub-criteria for Sustainability of Results/Benefits
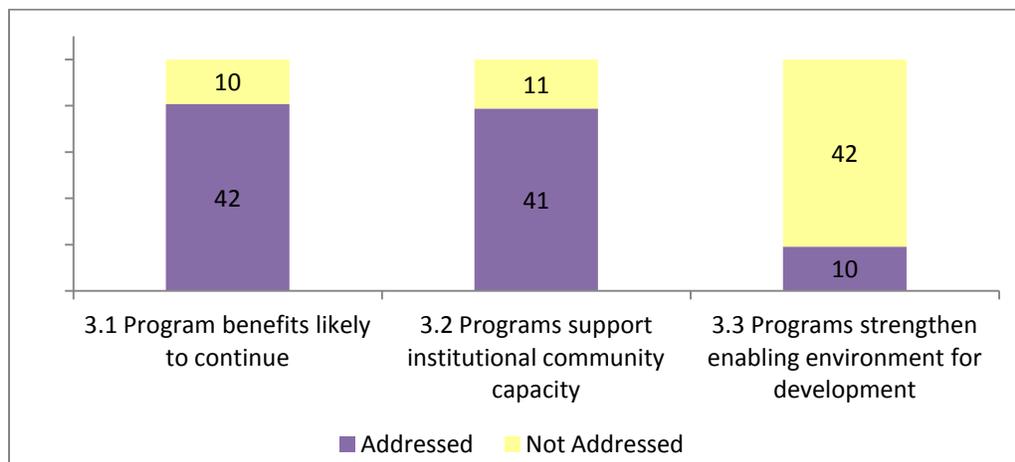

3.3.2 Key Findings

The findings regarding sustainability[36] reflect a mixed level of performance (Figure 7). Only 43% of evaluations reviewed reported findings of satisfactory or better for the sub-criterion 3.1 "Benefits continuing after program completion." The evaluation reports were more positive for

---

[36] This criterion addresses the overall sustainability of programming, as defined by the DAC – "The continuation of benefits from a development intervention (such as assets, skills, facilities or improved services) after major development assistance has been completed." (*Glossary of Evaluation and Results Based Management (RBM) Terms*, Working Party on Aid Evaluation, Development Assistance Committee, DCD/DAC/EV(2000)6, 30 October 2000 p. 16). The United Nations Development Group (UNDG) defines environmental sustainability as "the longer term ability of natural and environmental resources and ecosystem services to support continued well-being." (Mainstreaming Environmental Sustainability in Country Analysis and the UNDAF: A Guidance Note for United Nations Country Teams and Implementing Partners Teams, United Nations Development Group, 2009, http://www.undg.org/docs/10662/ES_GuidanceNote_FINAL.pdf) Although the DAC definition of sustainability refers specifically to development interventions, the review accepts the emerging view in the humanitarian community that sustainability is also relevant to humanitarian programming, since support for only short term interventions focused on the saving lives "…makes it hard to look for more cost-effective and sustainable approaches to addressing not only needs, but the factors that contribute to vulnerability … [and means that] many activities that could potentially support a gradual transition to recovery are often unfunded due to the bureaucratic divide, trapped between humanitarian funding and development funding, but not eligible for either." Sustainability in humanitarian action", Ross Mountain, Philip Tamminga, DARA, 2010 http://healthg20.com/wp-content/uploads/2011/11/107-111-Sustainability-in-humanitarian-action-DARA_2010.pdf

---

sub-criterion 3.2 "Institutional/community capacity for sustainability", with 64% reporting findings of satisfactory or better, where 15% of the evaluations reported results at the highly satisfactory level.

Figure 7: Sustainability of Results/Benefits (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=52)
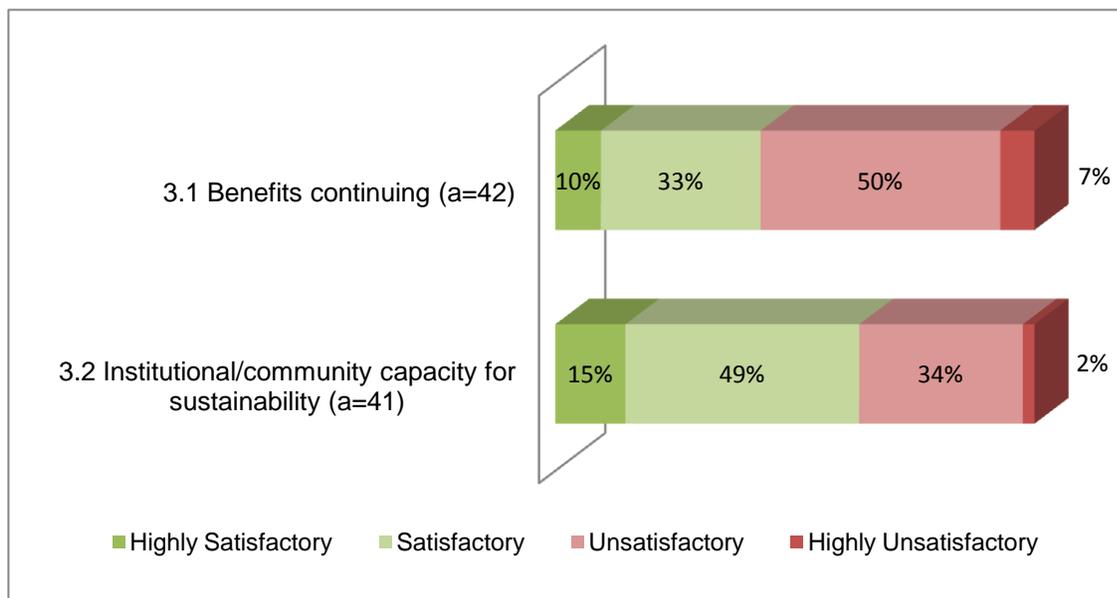
The evaluations reported on the following key success areas for sustainability:

- Sustainability of school feeding programs because governments have been able to ensure the long-term funding of the activities (5 evaluations); and

- Sustainability of physical assets that have been built or rebuilt as a result of a high level of community ownership (3 evaluations).

The use of capacity building and, specifically, training, was highlighted in the evaluations as a means to develop institutional and community capacity for sustainability (13 evaluations).

In addition, WFP's thematic evaluation of its capacity development policy and operations identified, based on 15 country case studies, areas in which WFP contributed to strengthening the enabling environment. These include: influencing policy through analysis and advocacy; supporting regional networks; and contributing to national policy development in the areas of nutrition food fortification, school feeding, safety net programs and disaster preparedness (Highlight box 8).[37]

---

[37] "Evaluation of WFP's Capacity Development Policy and Operations, World Food Programme", WFP, 2 May 2008, p. 24 – 26

> **Highlight Box 8**
>
> **Contribution to Capacity Development**
>
> "… WFP funded "The Cost of Hunger Study" in the Latin America region which shows to decision makers what the economic cost of hunger is. … Combined with advocacy work at various levels, including inter-ministerial meetings in the region, it resulted in political commitment and agreements, which form the basis for adopting national policies and strategies, allocating national resources, and developing organizational and technical capacities for policy implementation."
>
> *Capacity Development Policy and Operations (Full Report),* 2008, p. 24

3.3.3 Contributing Factors

Since the level of sustainability of WFP programming is relatively low, most evaluations reflected on the factors that hindered sustainability, including:

- Insufficient time and resources to implement programs to their full potential (i.e., short-term and under-funded interventions), including inadequate resources for the phase-out period (17 evaluations);

- Lack of involvement and ownership of government or government commitments in terms of resources and food distribution (10 evaluations);

- Weak exit strategies and the failure to build sustainability into the program design (Highlight box 9) (10 evaluations);

- Lack of support for government capacity building, in terms of conceptual understanding, technical and managerial skills, and accountability requirements (7 evaluations);

- Lack of attention early in the emergency response to the eventual shift from disaster relief to development, including the engagement of development actors and donors (7 evaluations);

- Challenges in the local context, such as the lack of security or socio-economic conditions, such as food prices, the status of agricultural production capacity or the capacity for individual ownership and rights (7 evaluations); and

- Lack of engagement of other stakeholders, including other UN agencies (4 evaluations) or communities (4 evaluations).

Factors that contributed to sustainability were also reported:

- Engagement with communities, through collaboration with WFP and strong community participation in program planning, and development of community cohesion through the implementation of projects and capacity through working with community organizations to develop managerial and technical capacity (8 evaluations);

- Engagement and commitment of governments to food security and nutrition issues reflected, for example, in the development of national frameworks and close collaboration with WFP (7 evaluations); and

- Capacity in developing country governments, in areas such as technical skills for food security assessments, logistics and monitoring (6 evaluations).

Highlight Box 9

Importance of Exit Strategies

"A critical dimension of sustainability is how well WFP is developing and implementing exit strategies from its livelihood recovery related activities. In several of the case study countries, WFP was forced through funding constraints to abruptly cease rather than phase out activities. There is a need to undertake stronger contingency planning to avoid cutting off activities at short notice and to provide more planned and gradual transitions when faced with funding or pipeline constraints. Exit strategies need to include advocacy with development donors and government to develop policies and programmes to address needs previously covered by WFP."

Effectiveness of WFP Livelihood Recovery Interventions (Full Report), 2009, p. vi

---

3.4 WFP Programs are Highly Relevant to their Context

3.4.1 Coverage

All three sub-criteria for relevance have high levels of coverage, with most evaluations addressing these issues. As illustrated by Figure 8, very few evaluations were coded as "not addressed" for each of the three sub-criteria in this area.

Figure 8: Number of Evaluations Addressing Sub-criteria for Relevance
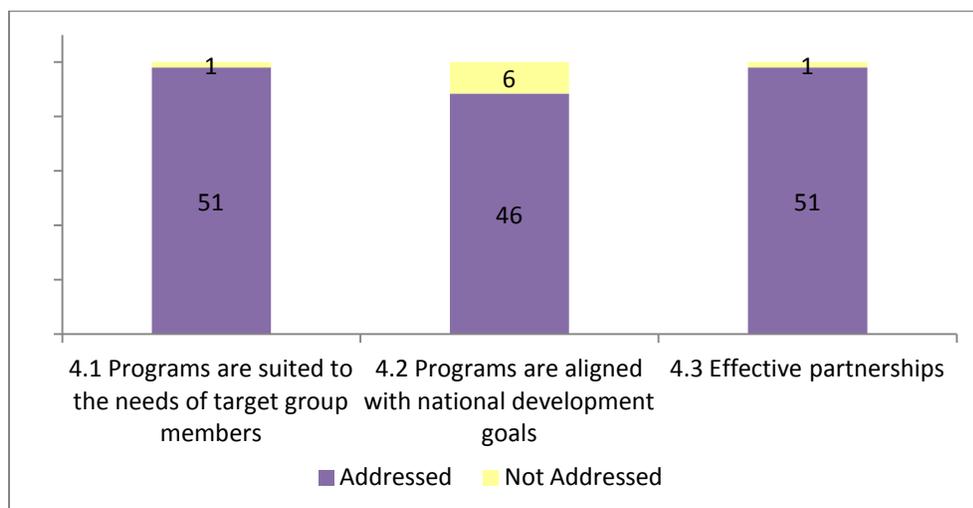
3.4.2 Key Findings

Relevance represents one of the strongest elements in WFP's humanitarian and development effectiveness. Eighty-four percent (84%) or more of the evaluations reported findings of satisfactory or better for all three relevance sub-criteria (Figure 9). The rating for sub-criterion 4.1 "Suited to the needs of target group members" was 86% satisfactory or better; sub-criterion 4.2 "Aligned with national humanitarian and development goals" was 93% satisfactory or better;

and sub-criterion 4.3 "Effective partnerships" was 84% satisfactory or better. The performance on sub-criterion 4.2 "Aligned with national humanitarian and development goals" was particularly good, with 50% of the evaluations reporting findings of highly satisfactory.

Figure 9: Relevance (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=52)
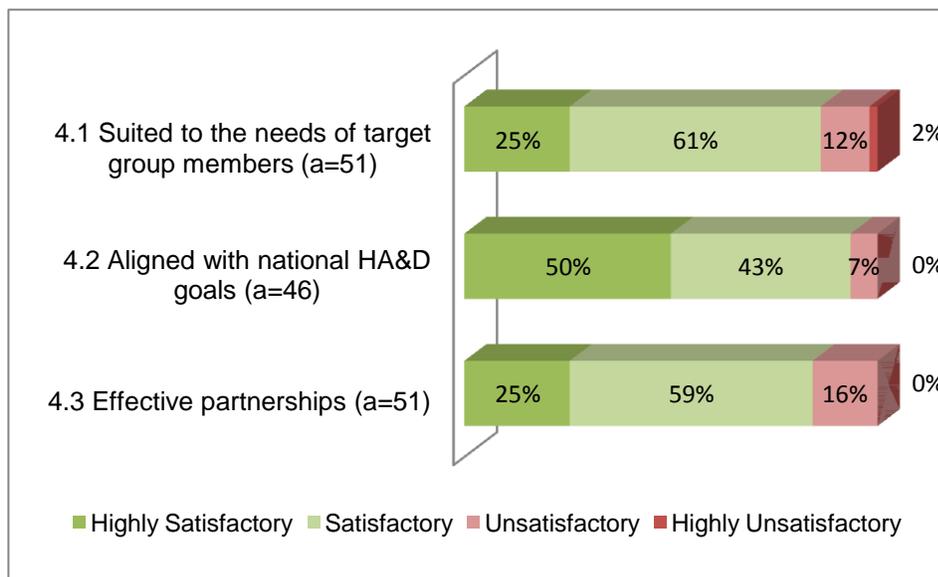
The evaluation reports highlighted a number of areas in which programming was seen as being particularly relevant. These areas include:

- Use of food aid for both relief and recovery activities and the combining of these activities, particularly identified in PRRO evaluations (Highlight box 10) (7 evaluations); and

- Targeted programming to specific populations based on location (e.g., remote areas of the country or vulnerable communities) or specific needs (e.g., babies, infants and pregnant/lactating women for Mother and Child Health and Nutrition programming or returning internally displaced persons) (11 evaluations).

Highlight Box 10

Integration of Relief and Recovery

"…the tendency to view relief and recovery as two separate categories of support, each associated with distinct food assistance activities, is not helpful. Combinations of relief assistance to meet basic food needs and recovery assistance to restore key household and productive assets have proved to be effective in helping people to recover livelihoods."

Livelihood Recovery Interventions (Full Report), 2008 p. 47

The evaluations provided additional evidence of the nature of WFP's partnerships, which contributed to its overall effectiveness in this area. The types of partnerships that were identified include:

- Partnerships with government and national institutions, including ministries responsible for health, education, social services (including women and children, social security, refugees and repatriation), agriculture and economic/rural development (Highlight box 11) (28 evaluations);

- Partnership with multilateral partners, including UN agencies and international financial institutions (Highlight box 12) (28 evaluations). The most commonly mentioned were the United Nations Children's Fund (17 evaluations), Food and Agriculture Organization (10 evaluations), Office for the Coordination of Humanitarian Affairs and World Health Organization (5 evaluations each) and United Nations High Commissioner for Refugees (4 evaluations). The nature of the relationships with the UN agencies ranges from common UN planning and joint programming to the delivery of materials and services;

- Partnerships with civil society, including both national and international non-governmental organizations and local community associations (25 evaluations). The nature of the relationships with non-governmental organizations and community associations includes operational planning, implementation and monitoring; and

- Partnerships with donors (8 evaluations). Some evaluations identified that WFP is working with donor representatives in common approaches and pooled resources (3 evaluations).

> **Highlight Box 11**
>
> **Joint Government/UN Programming**
>
> As part of its Food for Work programming, WFP supports "the Green Afghanistan Initiative (GAIN), a joint programme of action of the Government and six UN agencies, which started in 2005 and aims at environment and natural resource protection." WFP provides food-based support to various activities under the programme, mainly to nursery workers.
>
> *Afghanistan PRRO (Full Report),* 2009, p. 28

3.4.3 Contributing Factors

The review noted that the following factors positively contributed to relevance:

- The nature of the arrangements of the partnerships (Highlight box 12) (8 evaluations). Partnerships are strengthened with the formalization of arrangements by letters or memoranda of understanding, collocation in the same geographic area, regular meetings, joint programs of action and maintaining flexibility;

- Close partnership with developing country governments (7 evaluations) that allow WFP to influence ongoing debates about food security, future directions of government policy and develop common approaches;

- Good use of needs assessments and the integration of needs information in program design (7 evaluations). This includes examples of the use of needs information to adapt programs to changing needs (3 evaluations); and

- WFP staff's vision, accessibility and commitment, which contributed to effective partnerships (3 evaluations).

On the other hand, the review noted examples from evaluations where the following factors detracted from relevance:

- Lack of available partners and/or their limited commitment (11 evaluations). A few evaluations noted that partnerships with other partners are even more important in areas where there are limitations in some partner groups (3 evaluations). For example, when there is a new government in place in a country, and limited UN agency or non-governmental organization capacity or reduced donor support, the evaluations noted the importance of building on other possible partnerships; and

- Weaknesses in conducting needs assessments during the planning of some WFP programs and poor links between needs and programming (6 evaluations). This contributed to a reduced fit between programs and the needs of target populations. Specifically, the review identified programming where school feeding programs were not adequately based on nutritional and/or education needs (4 evaluations).

Highlight Box 12

Working in Partnership

"In efforts to conduct capacity development at a policy-level, WFP seems to have greater impact when working in partnership. … In West Africa, for example, the collective voices of WFP and UNICEF are combined in an effort to influence governments to take longer-term, more sustainable approaches to nutrition; similar examples abound in other regions as well. In addition to partnerships with other UN organizations, WFP works with non-governmental organizations for community level interventions, creating synergies between short-term food intervention to alleviate hunger and longer term capacity development to mitigate the effects of future disaster."

Evaluation of WFP's Capacity Development Policy and Operations, 2008, p. 22

3.5 Evaluations Report Mixed Results in Efficiency

3.5.1 Coverage

All sub-criteria for efficiency were rated moderate in coverage (Figure 10). Fourteen evaluation did not address sub-criterion 5.1 "Programs cost efficient", while 12 did not address sub-criterion 5.2 "Programs implemented/objectives achieved on time" and 11 evaluations failed to address sub-criterion 5.3 "Systems/procedures for program implementation efficient."

Figure 10: Number of Evaluations Addressing Sub-criteria for Efficiency
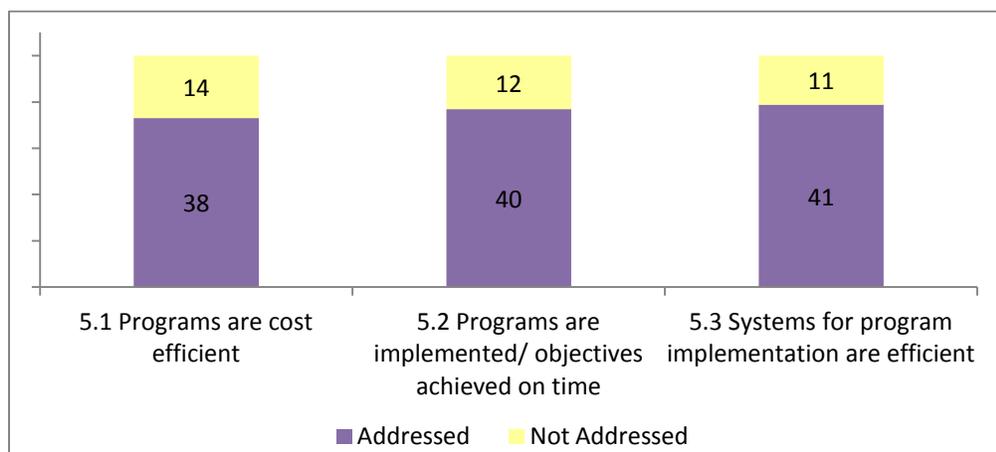
3.5.2 Key Findings

The efficiency of WFP's programming represents another challenge for the organization in some areas. The findings with respect to efficiency reflect more positive findings for WFP's systems and procedures and cost efficiency than they do for the timeliness of WFP's activities (Figure 11). For the sub-criterion 5.3 "Systems/procedures for program implementation efficient", 63% of the evaluations reported findings of satisfactory or better and 58% for sub-criterion 5.1 "Programs cost efficient." However, only 42% of the evaluations reported findings of satisfactory or better for the sub-criterion 5.2 "Programs implemented/objectives achieved on time."[38]

Figure 11: Efficiency (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=52)

---
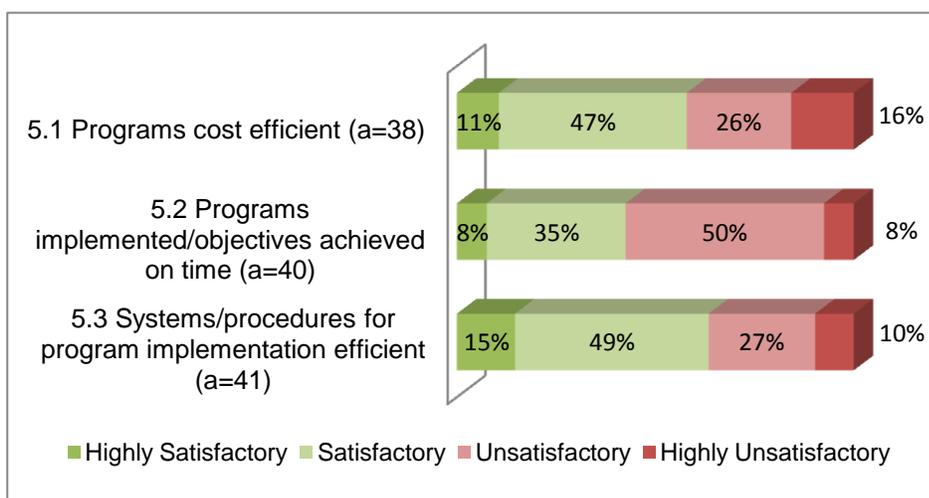
[38] Similarly to the other sub-criteria, revisions to the efficiency sub-criteria were made to improve their applicability to HA operations. However, further development of these sub-criteria may be necessary for future reviews to strengthen their applicability to this type of operations.

Many technical issues were raised with respect to the efficiency of WFP's programming, such as the cost of importing supplies, losses due to infestations, issues in using international tenders, high cost of logistics, WFP's approach to assessing support cost. A couple of evaluations suggested that WFP staff is aware of issues of cost efficiency, has the necessary skills to manage difficult logistics and is attentive to costs in planning and procurement (2 evaluations).

The most common concern with respect to the timeliness of WFP's interventions was delays in the delivery or distribution of food or the inability to carry out activities as planned, including reduced rations (Highlight box 13) (12 evaluations).

The evaluations reviewed also revealed observations on WFP's systems and procedures, including inadequate financial and/or human resources (3 evaluations) and weaknesses in the organizational structure and culture (3 evaluations). These few comments on organizational weaknesses suggest that the organizational structure and culture do not optimally support flexible planning, design and implementation approaches to respond to rapidly changing situations and that opportunities are missed for analysis, reflection and learning.

Highlight Box 13

Timeliness of Emergency School Feeding

"In all of the observed programmes, the delivery of food commodities to schools had been interrupted at least once. Sometimes the disruption affected only specific geographic areas during certain times of the year, or particular commodities, in other cases, the programme overall was affected and sometimes delayed significantly because the food did not reach the schools."

School Feeding in Emergency Situations, 2007, p. 46

3.5.3 Contributing Factors

Positive factors contributing to efficiency identified in the evaluations reviewed include:

- WFP's strength in the management of logistics (Highlight box 14) (7 evaluations);

- Greater use of local resources – both staff and procurement – that, at times, represented a creative solution to challenging situations, such as deliveries to remote areas (5 evaluations); and

- New management systems, sharing facilities with partners, reducing the number of staff, and the high level of staff awareness of the importance of efficiencies (5 evaluations).

The key factors that limited program efficiency were:

- Issues with resources, often resulting in program interruptions, at times specifically related to the lack or unpredictability of donor funding, negatively affected both the efficiency and timeliness of WFP's operations (Highlight box 15) (18 evaluations);

- Staff capacity issues, timeliness of the interventions, and the efficiency of systems and procedures (12 evaluations). There was recognition of the strong capacity of WFP staff as generalists, but gaps in technical expertise were identified. Four evaluations noted particularly the problem of high staff turnover and the lack of continuity in staffing;

- Logistics and transportation costs, including inefficiencies in the set-up of logistics and issues with particularly complex logistics (10 evaluations);

- Challenges with the scope of the programming, including very dispersed target populations, high levels of insecurity and particularly large-scale complex programs (9 evaluations); and

- Limited partner capacity, including a limited number of non-governmental organization partners in the country and limited staff capacity in government or non-governmental organizations partner organizations (3 evaluations).

Highlight Box 14

Effective Logistics

"Pipeline management handled the difficulties of resource availability very well, local procurement initiatives maximized the limited opportunities available as well as helping to overcome the key problem of the timely availability of wheat flour and the intricate demands of even the pilot school feeding scheme. The logistics department has functioned well in an environment subject to frequent disruptions."

occupied Palestinian territory PRRO, 2010, p. 35

Highlight Box 15

Implications of Program Interruptions

"Reductions in food commodities have hampered the CP [Country Program] in achieving desired results in terms of geographic coverage and the number of beneficiaries, particularly in MERET. The natures of output and outcome achievement are commendable in their own right but, because of these commodity and related budget reductions, fewer households are benefiting from these results."

Mid-Term Evaluation of the Ethiopia Country Programme, 2009, p. 43

3.6   Good Use of Evaluation, but Inadequate Performance Frameworks and Weak Monitoring

3.6.1 Coverage

Some care is required in treating the results reported regarding the use of evaluation and monitoring to improve humanitarian and development effectiveness. There were issues with the coverage in the evaluations of three sub-criteria, namely the sub-criterion 6.1 "Systems and processes for evaluation are effective", the sub-criterion 6.4 "Evaluation results used to improve

humanitarian and development effectiveness", and the sub-criterion 6.3 "Results based management systems are effective." Several evaluations did not report on these issues, resulting in a moderate (39 and 35 evaluations) and a weak (11 evaluations) coverage, respectively, of these sub-criteria (Figure 12). In fact, since the coverage is low for the sub-criterion 6.3, no results have been reported. As for the sub-criterion 6.2 "Systems and processes for monitoring are effective", it was rated strong in coverage, as only two evaluation reports did not address this sub-criterion.

Figure 12: Number of Evaluations Addressing Sub-criteria for Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness
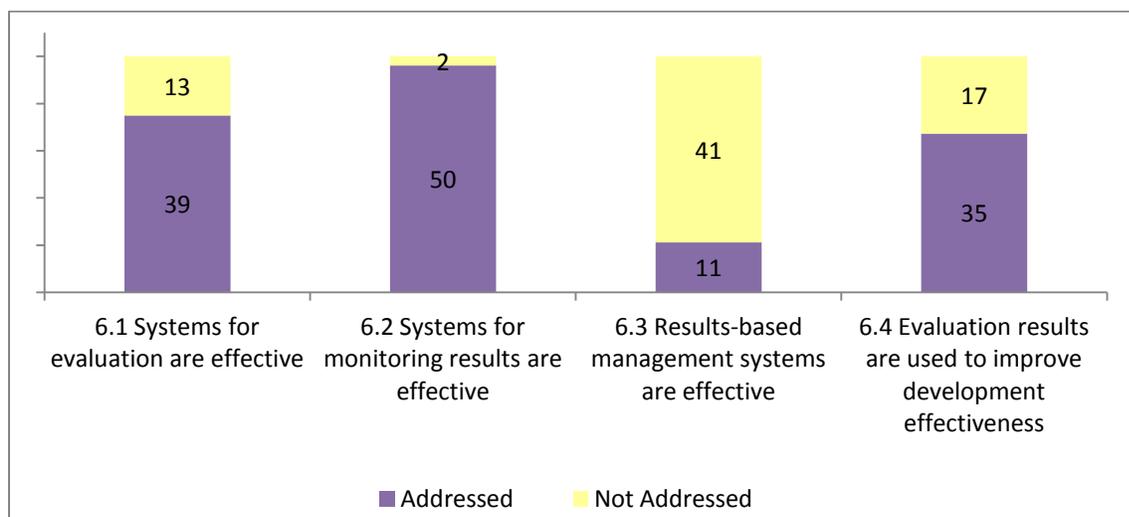
3.6.2 Key Findings

WFP's use of evaluation and monitoring to improve humanitarian and development effectiveness was assessed as positive, especially in respect to the systems for, and use of, evaluation, but less so for monitoring and reporting on results (Figure 13). Two-thirds of the evaluations (66%) reported findings of satisfactory or better for the sub-criterion 6.1 "Systems for evaluation effective", while only 32% reported findings of satisfactory or better for the sub-criterion 6.2 "Systems and processes for monitoring are effective."

Ninety-one percent (91%) of the evaluations reported findings of satisfactory or better for sub-criterion 6.4 "Evaluations used to improve humanitarian and development effectiveness," of which 88% reported findings of highly satisfactory.[39]

---

[39] For sub-criteria 6.4, the main indicator used by the review team was the frequency of a clear management response to the evaluations under review as well as the adequacy, specificity and clarity of that response, especially where it includes an action plan with time bound commitments for taking action on recommendations. A "highly satisfactory" finding was coded when the management response contained clear actionable items that were likely to result in the organizational and programmatic changes needed to achieve the intent of the recommendation.

Figure 13: Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=52)
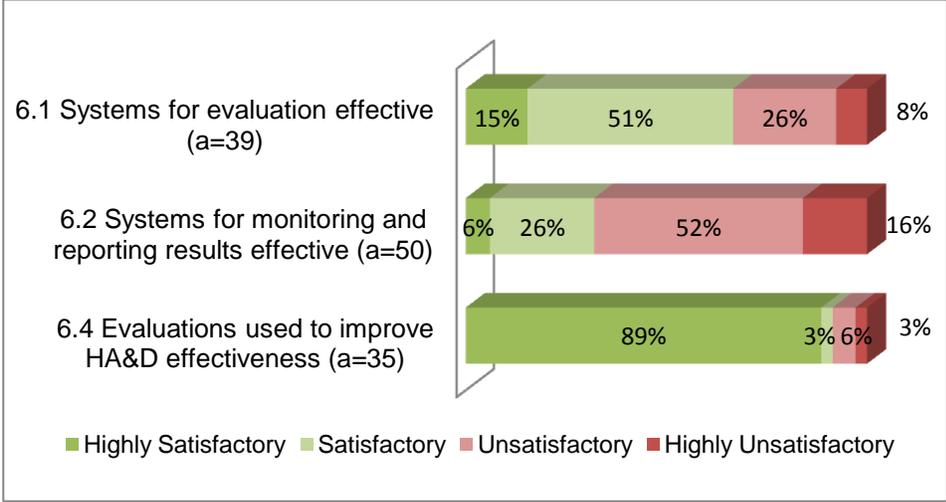
The evaluation reports highlighted a few issues with respect to WFP's use of evaluation and monitoring to improve humanitarian and development effectiveness, including:

- Evaluation and monitoring of WFP's programs focus more on measuring outputs rather than outcomes[40], which has an impact on WFP's ability to assess the effect of its programming on beneficiaries' health and nutritional status and report on WFP's humanitarian and development effectiveness (17 evaluations);[41]

- Findings from previous evaluations and studies were used in subsequent program design and other uses are made of evaluative information and monitoring data at the country-level (e.g., stakeholder workshops) (Highlight box 16) (13 evaluations); and

- Progress in WFP's evaluation and monitoring processes has been observed over the past few years (8 evaluations). Progress has been achieved through the development of new systems and databases, creation of monitoring and evaluation positions at WFP, development of tools and staff and partner training.

---

[40] It is recognized that the measurement of outcomes of humanitarian programming is challenging. However, this observation applied also to evaluations of WFP's development programming.
[41] In 2009, WFP began conducting impact evaluations specifically to evaluate outcomes and impact-level results. Although some were included in this review, the contribution of one type of evaluation could not be isolated in the ratings.

<div style="border: 2px solid blue; padding: 10px;">

Highlight Box 16

Importance of Monitoring Information

"What is missing is a monitoring mechanism providing an overall view of the situation, especially in the context of an EMOP."

Évaluation de la réponse du PAM à la crise alimentaire au Niger en 2005, 2006, p. 16

</div>

3.6.3 Contributing Factors

Positive factors that contribute to achieving effective evaluation and monitoring to improve WFP's effectiveness include:

- Conduct of studies, including evaluations, reviews, and assessment baseline studies that supported the implementation and evaluation of program activities (18 evaluations); and

- Planning for the measurement of humanitarian and development effectiveness, including specific references to monitoring and evaluation plans and strategic frameworks for accountability (7 evaluations).

Factors that detracted from the achievement of effective evaluation and monitoring include:

- Lack of adequate capacity for evaluation and monitoring in WFP or partner organizations, including government and non-governmental organizations (15 evaluations). This included limitations on financial resources and staff time and lack of relevant experience (Highlight box 17);

- Limitations with respect to the development of effective frameworks for monitoring and reporting (8 evaluations). This included limitations with: the performance frameworks that were not sufficiently detailed or did not include methods for measuring impacts; log frames that were weak and did not include relevant indicators for some components of programs (e.g., livelihood recovery); and monitoring approaches that did not include all components of programs, and indicators that were imprecise or did not adequately reflect the expected results; and

- Poor quality of inputs to monitoring systems from community-level service providers, which affected WFP's monitoring data quality (4 evaluations).

4.0 Conclusions

The 52 evaluations reviewed provided a good coverage of WFP's programming, and offer a good level of confidence for the findings reported. However, during the period under review (2006-2011), WFP has continued to improve its programming, with many changes responding to the findings of the evaluations used in this review. A comprehensive survey of the improvements initiated by WFP in the recent past is outside the scope of this review. However, WFP has provided an overview of some of the most important of these recent changes (see Annex 8).

Based on the identified findings and related contributing factors, this review concludes that:

1. Evaluations carried out since 2006 indicate that WFP is effective in the achievement of humanitarian and development objectives and expected results. Over three-quarters of WFP evaluations reported that program performance in achieving objectives and expected results was satisfactory or better. WFP's most cited achievements were in the distribution of food aid and other programming, such as school feeding programs, which contributed to: improvement in food consumption, prevention of acute hunger, reduction of risk of chronic hunger, and reduction malnutrition for target populations. WFP's policy dialogue with, and support to, developing country governments for the development of national humanitarian and development policies and programs was also noted positively in the evaluations reviewed. A key factor contributing to WFP effectiveness in achieving humanitarian and development objective and expected results was its strong logistic capacity. However, there are also factors that detract from the achievement of objectives, notably with respect to financial issues, such as program interruptions and the lack of predictable donor funding;

2. Addressing the issue of program interruptions is the most critical challenge for WFP. These may be caused by the lack of predictable donor funding, poor contingency planning and/or inefficiencies in WFP operations. They represent factors that detract from WFP's performance in a number of areas – achievement of objectives, sustainability, efficiency and monitoring.

3. WFP has not been as effective in addressing the cross cutting themes of gender equality and the environment.

i. WFP's programming in only just over half of the evaluation was given a rating of satisfactory or better with respect to its effectiveness in addressing gender equality. Successes, however, were noted in the area of provision of sex-disaggregated data and some improvements for women and girls. WFP's challenge in promoting gender equality was also reflected in its own evaluation of the former Gender Policy (2003-2007), which noted the need for restoring the gender mainstreaming mandate, completing the transition from women's programming to gender mainstreaming, re-orienting the institutional approach to enable context-led approaches and marshalling and allocating resources at the operational level.

ii. The cross cutting theme of environmental sustainability was not adequately addressed in the evaluation reports, which prevented the review from identifying any results in this area.

4. The sustainability of the results of WFP programming presents an area for improvement for the organizations. While the sustainability of EMOP benefits represents a significant challenge, a number of evaluations indicated that programs could be better designed to assist in the transition to more sustainable access to food in the post-program period. Key factors that detracted from sustainability include: lack of time and resources for long-term, well-funded interventions; lack of engagement with developing country governments; and the lack of attention early in the intervention to plan the shift from relief to recovery and, eventually, development. In order to strengthen the sustainability of program benefits, WFP will need to increase efforts to ensure strong program ownership by developing country partners. It will also need to address the absence, in some programs, of an explicit exit strategy and the failure to integrate sustainability into program design.

5. The area in which WFP's humanitarian and development effectiveness is the strongest is the relevance of its programming in developing country contexts. The programming is both highly relevant to the needs of target groups and is aligned with the priorities of the national governments. The important factors contributing to such a high level of relevance are the quality of WFP's needs assessments and its strong partnerships. For an agency with food assistance at the centre of its mandate, effectively matching programming to the needs of target group members is one of the most critical elements of effective programming. A positive conclusion in this area is especially significant for WFP. Additionally, WFP has been successful in developing effective partnerships by forming strong links with developing country governments, other multilateral agencies, civil society and, to some extent, donors;

6. The efficiency of WFP's programming represents another area for improvement for the organization. Efficiency is reduced because of logistics and transportation setbacks, the need to serve dispersed populations, and frequent financial issues (e.g., lack or unpredictability of donor funding). On the sub-criterion related to cost efficiency, three out of five evaluations reported findings of satisfactory or better, leaving significant room for improvement. The sub-criterion regarding timeliness was rated as satisfactory or better in less than half the evaluations, often affected by program interruptions. WFP's systems and processes were rated better (63% satisfactory or better) – when they work, WFP's strength

in the management of logistics and its greater use of local resources lead to greater efficiency; and

7. While WFP receives high ratings for its systems and use of evaluation (66% and 91% rated satisfactory or better, respectively), the underpinning of evaluation – monitoring – is a challenge. Only 32% of the evaluations reported findings of satisfactory or better for the adequacy of the systems for monitoring and reporting on results. There is a sense of insufficient capacity both within WFP and within its partner organizations for monitoring and inadequate performance frameworks for many of its programs. To some extent this is also affected by the lack of sufficient resources for monitoring – both for adequate staff training and travel to the field.

Examining WFP's humanitarian assistance and development effectiveness across the six criteria used in this report builds on and further develops WFP's own reporting of its results. WFP's own summary of its evaluation findings as reflected in its annual evaluation report for 2010 indicated:[42]

- WFP's strongest area of performance is its contribution to the life saving component of Strategic Objective 1 – relief delivered through general food distribution, food for work in lieu of general distributions and contributions to grain banks;

- The extent to which it contributes to Strategic Objectives 1, 2 and 3 (protecting and rebuilding lives and livelihoods and disaster preparedness work) depends largely on the performance of its Food for Work activities. The evaluations show shortfalls in this area, particularly as a result of underfunding;

- WFP's evaluations of its school feeding programming reconfirmed its contribution to Strategic Objective 4 (addressing chronic hunger). While it was shown to have a positive impact on enrolment, attendance and attainment rates, impact evaluations suggest that "school feeding's effectiveness levelled off when children reached an age where their income-earning potential outweighed the value of schooling and the school meal;"[43] and

- Strategic Objective 4 is also supported by nutrition programs. However, they struggle to demonstrate results, in part because of small size of the programs related to overall needs, and because of difficulties in measuring outcomes.

While this review's findings are broadly consistent with WFP's 2010 annual evaluation report, they also differ. As noted in the observations related to WFP's programming in this review, WFP demonstrated its ability to deliver in difficult operating environments. However, many evaluations covered by WFP's summary highlighted areas for improvement that lie within its control: increased targeting of programming and better planning and performance management. In addition, as noted in this review, the predictability and timeliness of funding explained, in part, the limitations of WFP's operational performance and results.

---

[42] "Annual Evaluation Report 2010", WFP, 2011
[43] Ibid., p. 3

5.0 Areas for Attention

The previous section summarized the strengths and weaknesses in WFP's humanitarian and development effectiveness. This section highlights the areas in which WFP, working with its bilateral partners, needs to continue to focus attention.

1. Considering the fact that interruptions of programming detract from WFP's performance in a number of areas, WFP needs to continue to work with its bilateral partners to address these interruptions. The review notes that WFP is undertaking reviews of its resourcing and financial frameworks. WFP needs to continue to work with other donors to encourage them to provide more predictable funding. It may also be possible to address the issue of program interruptions by improving internal practices in targeting WFP's programming and improving cost efficiency so that the organization can, to the extent possible, avoid these interruptions;

2. Promoting gender equality is a weakness if WFP's performance. While WFP did develop and approve a new gender policy in 2009, there is a continuing need to ensure the effective implementation of this new policy, including improved delivery on gender equality results, so that it contributes to WFP's effectiveness as a humanitarian and development organization;

3. WFP identified commitments to address environmental issues. However, these need to be integrated into the planning, implementing, monitoring and reporting, and evaluating of its programs. This would include addressing, among other things, the environmental impact of its choice of food commodities and its food for work/assets programming. The review suggests that these are not currently addressed by WFP, at least not by WFP evaluations. WFP could benefit from assessing its programs' environmental sustainability to ensure that they make a positive contribution;

4. Given the review's conclusions on sustainability, WFP should continue to work on strengthening the sustainability of its programming. It should take into account the need for sustainability of its development activities and the more effective transition of its humanitarian activities to development programming; and

5. The review identified weaknesses in WFP's performance frameworks. They need to be strengthened to ensure they reflect adequately the expected results of its programming. Furthermore, it should address the underlying issues affecting its monitoring and reporting systems, by strengthening its internal monitoring capacity and working with its partners to strengthen their own capacity.

Annex 1: Criteria Used to Assess humanitarian and development Effectiveness

1. **Achievement of humanitarian and development Objectives and Expected Results**
1.1 Programs and projects achieve stated objectives 1.2 Positive benefits for target group members 1.3 Substantial numbers of beneficiaries/contribution to national humanitarian and development goals 1.4 Significant changes in national development policies/programs
2. **Cross Cutting Themes: Inclusive humanitarian and development Which can be Sustained (Gender Equality and Environmental Sustainability)**
2.1 Programs effectively address gender equality 2.2 Changes are environmentally sustainable
3. **Sustainability**
3.1 Program benefits are likely to continue 3.2 Programs support institutional and community capacity 3.3 Programs strengthen enabling environment for humanitarian and development
4. **Relevance**
4.1 Programs are suited to the needs of target group members 4.2 Programs are aligned with national humanitarian and development goals 4.3 Effective partnerships with government, civil society and humanitarian and development partners
5. **Efficiency**
5.1 Program activities are cost efficient 5.2 Programs are implemented/objectives achieved on time 5.3 Systems for program implementation are efficient
6. **Using Evaluation and Monitoring to Improve humanitarian and development Effectiveness**
6.1 Systems and processes for evaluation are effective 6.2 Systems and processes for monitoring are effective 6.3 RBM systems are effective 6.4 Evaluation results used to improve humanitarian and development effectiveness

Annex 2: Evaluation Sample

#	Year	Title	Type
1	2011	Summary Evaluation Report Rwanda Country Portfolio	Country Portfolio
2	2011	WFP School Feeding Programmes in Cote d'Ivoire (1999-2009): A Mixed Method Impact Evaluation – Vol I – Full Report	Impact
3	2011	School Feeding Programmes in The Gambia 2001-2010: A Mixed Method Impact Evaluation – Vol I – Full Report	Impact
4	2011	WFP's Role in Social Protection and Safety Nets: A Strategic Evaluation	Strategic
5	2010	WFP Cambodia School Feeding 2000-2010: A Mixed Method Impact Evaluation	Impact
6	2010	Mali : Une évaluation du portefeuille d'activités du PAM (2003–2009)	Country Portfolio
7	2010	Mid-Term Evaluation of the Ethiopia Protracted Relief and Recovery Operation 10665.0 (2008-2010)	Operational
8	2010	Evaluation of WFP Country Programme 10418.0 Ghana (2006-2010)	Operational
9	2010	Impact Evaluation of WFP School Feeding Programmes in Kenya (1999-2008): A Mixed-Methods Approach – Vol 1	Impact
10	2010	Evaluation Report of the Egypt Country Programme 10450.0 (2007-2011)	Operational
11	2010	Nepal: An Evaluation of WFP's Portfolio – Vol I – Full report	Country Portfolio
12	2010	Sudan EMOP 10760.0: Food assistance to populations affected by conflict	Operational
13	2010	occupied Palestinian territory PRRO 10387.1: September 2007 and August 2009	Operational
14	2010	Chad: An Evaluation of WFP's Portfolio (2003-2009)	Country Portfolio
15	2010	Asistencia Alimentaria a Personas en Situación de Desplazamiento y a otros grupos con altos índices de inseguridad alimentaria afectados por la violencia en Colombia – OPSR 10588.0	Operational
16	2010	Evaluacion de la Operación Prolongada de Socorro y Recuperacion Guatemala 10457.0	Operational
17	2009	Evaluation of WFP Timor-Leste PRRO 10388.1 Assistance to Vulnerable Populations (September 2008-August 2010)	Operational
18	2009	Evaluation of Afghanistan PRRO 10427.0	Operational
19	2009	Joint Thematic Evaluation of FAO and WFP Support to Information Systems for Food Security	Thematic
20	2009	Mid-Term Evaluation of the Ethiopia Country Programme 10430.0 (2007-2011)	Operational

#	Year	Title	Type
21	2009	Country Portfolio Evaluation of WFP Assistance to Malawi – Final Report	Country Portfolio
22	2009	Country Portfolio Evaluation of WFP Assistance to the Lao PDR – Final Evaluation Report	Country Portfolio
23	2009	Strategic Evaluation of the Effectiveness of WFP Livelihood Recovery Interventions	Strategic
24	2009	Assistance to Populations Affected by the Cote d'Ivoire Protracted Crisis PRRO 10672.0	Operational
25	2009	Evaluation of the Mozambique Country Programme 10446.0 (2007-2009) – Full Report	Operational
26	2009	Evaluation of WFP response to Hurricane Felix in Nicaragua	Operational
27	2009	Country Programme Bangladesh CP 10410.0 (2007-2010)	Operational
28	2009	Strategic Evaluation of WFP's Contingency Planning 2002-2008	Strategic
29	2009	Summary Evaluation Report Republic of the Congo PRRO 103121	Operational
30	2009	Evaluation of the Liberia PRRO 10454.0 (July 2007-June 2009) – Full Report	Operational
31	2009	Evaluation of the Burkina Faso PRRO 10541.0 – Full Report	Operational
32	2009	Évaluation de l'Intervention Prolongée de Secours et de Redressement (IPSR) 10608.0	Operational
33	2009	Rapport d'évaluation de l'intervention prolongée de secours et de redressement à Madagascar (IPSR 10442.0)	Operational
34	2008	Thematic Evaluation of WFP's HIV and AIDS Interventions in Sub-Saharan Africa – Full Report	Thematic
35	2008	Full Report of the End-of-Term Evaluation of WFP's Gender Policy (2003-2007)	Policy
36	2008	Mid-Term Evaluation of the Colombia PRRO 10366.0	Operational
37	2008	Evaluation of WFP's Capacity Development Policy and Operations	Thematic
38	2008	Evaluation of Kenya Emergency Operation 10374.0 and Country Program 10264.0 (2004-2008)	Operational
39	2007	Evaluation of the WFP Strengthening Emergency Needs Assessment Implementation Plan –Vol 1	Operational
40	2007	Thematic Evaluation of the WFP School Feeding in Emergency Situations – Full Report	Thematic
41	2007	Mid-Term Evaluation of PRRO 10362.0 Ethiopia	Operational
42	2007	Mid-Term Evaluation of Special Operation No.10498 WFP Shipping Service in Aceh and Nias, Indonesia – Full Report	Operational
43	2007	Mid-Term Evaluation of the Southern Africa Protracted Relief and Recovery Operation (PRRO 10310.0) – Full Report	Operational

#	Year	Title	Type
44	2007	Évaluation de l'opération d'assistance alimentaire aux personnes vulnérable en situation de crise en Haïti (IPSR 10382.0)	Operational
45	2006	Mid-Term Evaluation of the Indonesia PRRO 10069 Assistance to Recovery and Nutritional Rehabilitation – Full Report	Operational
46	2006	Evaluation of EMOP 1-339.0/1: Assistance to populations affected by conflict in greater Darfur, West Sudan – Full Report	Operational
47	2006	Mid-Term Evaluation of the WFO India Country Programme (2003-2007)	Operational
48	2006	Evaluation of the WFP Yemen Country Programme (2002-2007) – Full Report	Operational
49	2006	Evaluation of the PRRO 10181.00 – Food Air for Relief and Recovery in Somalia – Full Report	Operational
50	2006	Evaluation of the Bhutan Country Programme 10133.00 (2002-2007) – Full Report	Operational
51	2006	Evaluation of the Tajikistan PRRO 10231.0 – Full Report	Operational
52	2006	Évaluation de la réponse du PAM à la crise alimentaire au Niger en 2005 – Vol 1	Operational

Annex 3: Methodology

This annex provides a more thorough explanation of the key elements of the methodology used for the review of WFP's humanitarian and development effectiveness. It is structured around the sequence of tasks undertaken during the review: determining the rationale for the review; drawing the sample of evaluations; undertaking the process of review and controlling for quality during the analysis phase; and assessing the level of coverage provided by the humanitarian and development effectiveness review.

The review of evaluation reports was supplemented w a review of WFP corporate documents. This supplementary information provided context for the reviewers and expanded the profile of the organization and its programming. A list of the documents consulted is provided in Annex 6.

Rationale for the Humanitarian and Development Effectiveness Review

The common approach and methodology offer a rapid and cost effective way to assess the development effectiveness and, in this case humanitarian assistance effectiveness of the multilateral organization, relative to a more time consuming and costly joint evaluation. The approach was developed to fill an information gap regarding the development effectiveness of multilateral organizations. Although these multilateral organizations produce annual reports to their management and/or boards, bilateral shareholders were not receiving a comprehensive overview of the performance on the ground of multilateral organizations. The Multilateral Organization Performance Assessment Network (MOPAN) seeks to address this issue through organizational effectiveness assessments. This approach complements MOPAN's assessments.

The approach suggests conducting a review based on the organization's own evaluation reports when two specific conditions exist:[44]

1. There is a need for field tested and evidence-based information on the effectiveness of the multilateral organization.

2. The multilateral organization under review has an evaluation function that produces an adequate body of reliable and credible evaluation information that supports the use of a meta-evaluation methodology to synthesize an assessment of the organization's effectiveness.

The evaluation function at WFP does produce enough evaluation reports of good quality to support an assessment of the humanitarian and development effectiveness of WFP. The second condition is also satisfied, as WFP's existing reporting mechanisms do not provide sufficient information on the organization's humanitarian and development effectiveness.

---

[44] *Assessing the Development Effectiveness of Multilateral Organizations: Approach, Methodology and Guidelines*, Management Group of the Task Team on Multilateral Effectiveness, DAC EVALNET, 2011.

WFP's Evaluation Function (Quantity and Quality)

Quantity of WFP Evaluations: WFP's evaluation function conducts five different types of evaluations. Figure 1 provides an overview of the difference evaluation types.

The Office of Evaluation website at WFP (www.wfp.org/about/evaluation) identifies 52 evaluation reports conducted by the Office of Evaluation and published between 2006 and August 2011.[45] These were all the Office of Evaluation-managed evaluations. Decentralized evaluations conducted by the Country Offices were not included. This group of 52 evaluations forms the evaluation universe for this review. Rather than sample from a small population, the reviewers selected all 52 for the conduct of this review (described in more detail in the annex).

Since early 2006, WFP's Office of Evaluation published evaluations covering WFP development, emergency and protracted relief and rehabilitation operations in 34 countries, including eight of 10 of the most important WFP countries in terms of direct expenditures in the 2006 to 2009 period. The evaluations also covered 67% of all direct program expenses in the same period.

In total, the evaluations published by WFP since early 2006 create a substantial pool of reports that is large and diverse enough to support a meta-evaluation approach assessing humanitarian and development effectiveness.

The Quality of WFP's Evaluations: A 2007 Peer Review of the Evaluation Function at WFP concluded that the credibility of the products of the evaluation function up that point was "uneven."[46] However, it also noted that the central evaluation office was "a strong unit, with committed, well trained and highly motivated staff."[47] The peer review panel also reported that, if WFP were to implement the changes planned for its evaluation function, it would adequately address the findings of the peer review.

The review team conducted its own quality review of a sample of 52 different evaluations (described in more detail in this annex). The results of that quality assessment were positive, with 90% of the reviewed evaluations scoring higher than 31 points out of a possible total of 45. The evaluations were scored against 11 different dimensions of quality derived from the United Nations Evaluation Group Norms and Standards for Evaluation. (Annex 4 for the evaluation quality scoring grid.)

Therefore, the first condition for carrying out a humanitarian and development effectiveness review using the common approach is met. There is a large enough body of evaluation reports of good quality to support the use of meta-evaluation to assess WFP's humanitarian and development effectiveness.

---

[45] The original list from WFP's website included 60 evaluations. However, on closer review, there were a number of duplications in the listing mainly caused by reports being published in difference languages. In addition, some evaluations were pure process evaluations and did not include findings focused on humanitarian and development effectiveness.

[46] "Peer Review: Evaluation Function at the World Food Programme", Sida, 2007

[47] Ibid., p. 11

WFP's Reporting on Humanitarian and Development Effectiveness

WFP reports to its Executive Board on humanitarian and development effectiveness mainly through three regular agency-wide reporting documents: annual evaluation reports, WFP annual reports, and the Executive Director's regular reports on the progress of the strategic plan.

The 2010 WFP annual report[48] highlighted global activities and provided a qualitative picture of humanitarian and development results, with some quantitative information, including WFP's financial reports. The qualitative picture is refined through case examples of work in specific countries and in response to specific emergencies. Quantitative information is mainly at the output level, with the exception of the links to the achievement of the Millennium Development Goals. The report described the extensive reach of WFP programming, as well as the number and type of beneficiaries reached. It also provided information on the challenges that the organization faced in the year.

At a more substantive level, WFP provides its Executive Board with regular reports on progress in implementing its strategic plan. The 2010 report documents results with respect to each of WFP's five strategic objectives and provides a rating on the level of achievement for each objective.[49] The ratings are based on specific output and outcome indicators associated with each strategic objective. The report includes an assessment of WFP's organizational performance by management result dimension, including securing resources, stewardship, learning and innovation, internal business processes, and operational efficiency. There is also a focus in the report on identifying lessons learned and reporting on innovative approaches.

This reporting approach represented a considerably more sophisticated use of output and outcome data than seen, for example, in the report for the 2005 year. The latest report is based on internal reporting of the achievement of the outputs and outcomes at the project level.[50] It is weakened, however, by the extent to which not all WFP projects report on the relevant indicators. In the 2010 report, the reporting rates[51] ranged from 12% to 100%. As a result, although WFP's regular reporting on humanitarian and development effectiveness is improving, it still lacks a strong overview of field-tested evidence on effectiveness.

Therefore, the second condition for carrying out a humanitarian and development effectiveness review using the common approach is met as there is still a need to improve the availability of effectiveness information of the organization that is based on field-tested and evidence-based reporting.

Selecting the Evaluation Sample

---

[48] "Fighting Hunger Worldwide: Annual Report 2010", WFP, 2011
[49] "Annual Report for 2010", WFP, 13 May 2011
[50] As long as one indicator is reported in ten or more projects. Ibid., p. 90
[51] Number of projects reporting on the indicator in the Standard Project Reports divided by the total number of projects that included the indicator in the logframe. Ibid., p. 87

Subsequent to a classification of WFP's universe of evaluations published between 2006 and 2011 based on type (Figure 15) and by year (Figure 16), the universe of evaluation was assessed based on the following principles:

- Maximize geographic coverage;
- Extent geographic coverage by adding a mix of thematic and strategic evaluations;
- Include a mixture of evaluation types; and
- Ensure that WFP's main program areas were adequately represented.

Based on this assessment, it was determined that the sample should contain the complete universe of evaluations. In addition to meeting the above principles, the use of the complete universe would allow for a sufficient number of evaluations for the conduct of the review and no sampling bias, unless the evaluations posted on the WFP web site are not representative of all WFP evaluations or the evaluations posted are not representative of all WFP's programming. The question then is the level of coverage of WFP activities in the available evaluations.

In all, the evaluations reviewed covered WFP development programs and emergency and protracted relief and rehabilitation operations in 34 countries, including some in which the largest operations were mounted during the period under review (e.g., Ethiopia, Somalia, Sudan, Cote D'Ivoire, Aceh Indonesia, and Colombia.). Geographic coverage is further strengthened by the thematic evaluations, which cover critical program sectors and initiatives on a global basis. A review of WFP annual expenditure reports shows that the countries in the sample accounted for 67% of all WFP direct expenditures in the 2006 to 2009 period (the latest period with available biennial budget reports).[52] It also shows that the evaluations covered some programming in eight of the 10 largest countries of WFP operations. While the period covered by the evaluations included six different years, the past three years account for 64% of the total expenditures (33 evaluations).

---

[52] It should be noted that this is based on the assumption, which is known to not always be the case, that an evaluation in a given country covers all the WFP programming in that country.

Figure 14: Type of Evaluation, as a Percentage of All Evaluations Reviewed

Legend:
- Operations Evaluations - 32 in 27 different countries - 62%
- Country Portfolio Evaluations - 6 in 6 different countries - 11%
- Impact Evaluations - 4 of school feeding programs at country level - 8%
- Strategic/Thematic Evaluations - 10 of global initiatives and sectors - 19%
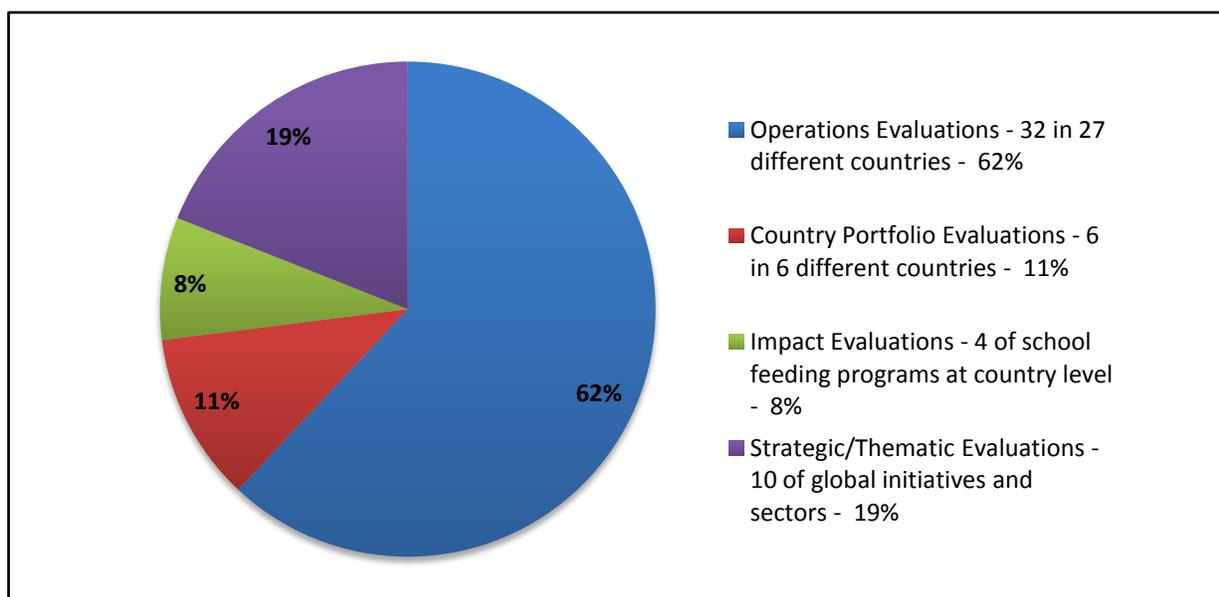
Figure 15: Year of Evaluation, as a Percentage of All Evaluations Reviews

Legend:
- 2011 - 4 evaluations - 8%
- 2010 - 12 evaluations - 23%
- 2009 - 17 evaluations - 33%
- 2008 - 5 evaluations - 10%
- 2007 - 6 evaluations - 11%
- 2006 - 8 evaluations - 15%
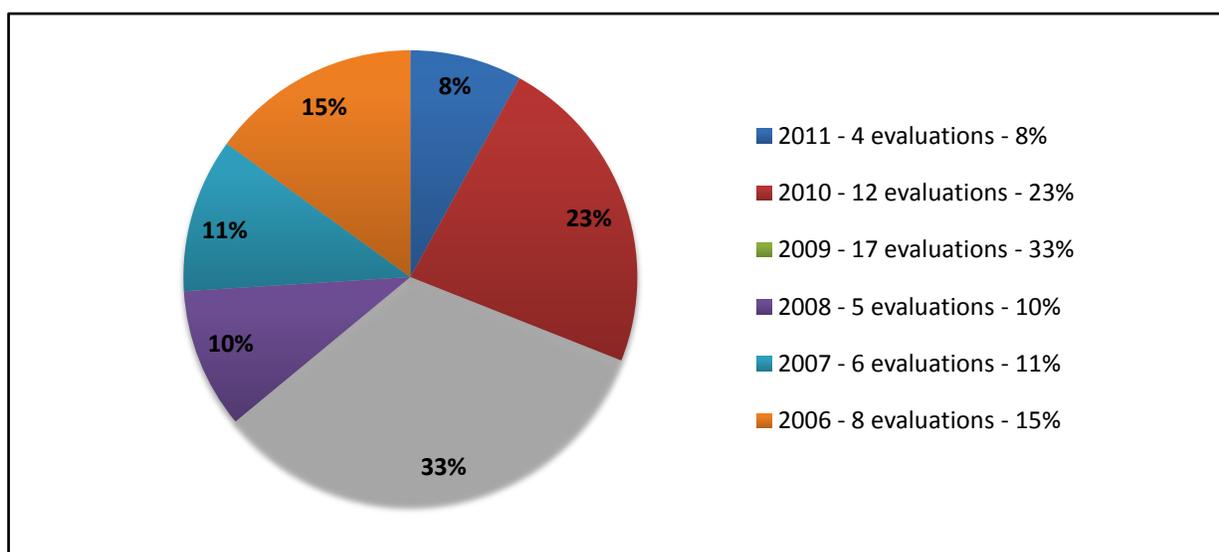
Process and Quality Assurance

The review itself was conducted by a team of four analysts and a team leader. A two-day training session was held for analysts to build a common understanding of the review criteria. Following, the team leader and analysts conducted a pre-test in order to independently review two evaluations. The team compared their ratings for these two evaluations and developed common agreement on the classification of results for all sub-criteria. This process helped to standardize classification decisions made by the analysts.

During the review of evaluations, analysts conferred regularly over any classification issues that arose. A second test occurred at the approximate mid-point of the review period, with all analysts independently rating a third evaluation. As previously, the analysts compared their respective classification decisions and resolved discrepancies.

Once the reviews were completed, the team leader reviewed the coded findings and carefully examined the cited evidence and contributing factors. Based on this examination, the team leader made a small number of adjustments to the coded findings. The process of training, testing and monitoring the review process minimized any inter-analyst reliability issues and controlled for bias on the part of any one reviewer.

All 52 evaluations in the sample were found to be sufficient quality to be included in the review. Of a possible maximum total quality score of 45, the mean score for all evaluations was 38. The distribution of total scores for all evaluations is described in Annex 4. Only five evaluations had a score of 30 or less. The quality criterion related to evaluation findings and conclusions being "relevant and evidence-based" is of primary importance. For this quality criterion, the mean score averaged across all evaluations in the sample was 3.8 out of a maximum score of four.

Coverage of Effectiveness Criteria

The review team developed a set of ranges in order to assess the level of coverage of a given sub-criterion. Strong coverage was assigned when the number of evaluations (*a*) addressing a particular sub-criterion was in the range of 45 to 52. Moderate coverage for a particular sub-criterion was assigned when *a* was between 30 and 44. Weak coverage for a particular sub-criterion was assigned when *a* was less than 30.

Of the 19 sub-criteria, eight received valid findings in 45 or more of the evaluation reports. Therefore, these eight sub-criteria received a rating of strong coverage. Another eight sub-criteria received valid findings in the moderate range. Three were rated weak in coverage (sub-criteria 2.2 "Changes are environmentally sustainable", 3.3 "Programs strengthen enabling environment for humanitarian and development" and 6.3 "Results based management systems are effective") (Figure 17). Given the weakness of coverage, results for these three sub-criteria were not included in the review of WFP's humanitarian and development effectiveness.

Figure 16: Levels of Coverage for Each Assessment Sub-Criteria

Criteria and Sub-Criteria		*a**	Coverage Level**	Satisfactory Ratings (%)	Unsatisfactory Ratings (%)
Achieving Humanitarian and Development Objectives and Expected Results					
1.1	Programs and projects achieve their stated humanitarian and development objectives and attain expected results.	52	Strong	77%	23%
1.2	Programs and projects have resulted in positive benefits for target group members.	50	Strong	92%	8%
1.3	Programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national humanitarian and development goals.	45	Strong	82%	18%
1.4	Programs contributed to significant changes in national humanitarian and development policies and programs (including for disaster preparedness, emergency response and rehabilitation) (policy impacts) and/or to needed system reforms.	31	Moderate	81%	19%
Cross Cutting Themes – Inclusive Humanitarian Assistance and Development which is Sustainable					
2.1	Extent to which multilateral organization supported activities effectively address the cross-cutting issue of gender equality.	32	Moderate	56%	44%
2.2	Extent to which changes are environmentally sustainable.	5	Weak	NA	NA
Sustainability of Results/Benefits					
3.1	Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations, to rehabilitation, reconstructions and, eventually, to longer term humanitarian and development results.	42	Strong	43%	57%
3.2	Projects and programs are reported as sustainable in terms of institutional and/or community capacity.	41	Moderate	64%	36%
3.3	Programming contributes to strengthening the enabling environment for humanitarian assistance and development.	10	Weak	NA	NA
Relevance of Interventions					
4.1	Programs and projects are suited to the needs and/or priorities of the target group.	51	Strong	86%	14%
4.2	Projects and programs align with national humanitarian and development goals implementation of support to development and/or emergency preparedness, humanitarian relief and rehabilitation efforts.	46	Strong	93%	7%

4.3	Effective partnerships with governments, bilateral and multilateral development and humanitarian organizations and non-governmental organizations for planning, coordination and	51	Strong	84%	16%
Efficiency					
5.1	Program activities are evaluated as cost/resource efficient.	38	Moderate	58%	42%
5.2	Implementation and objectives achieved on time (given the context, in the case of humanitarian programming).	40	Moderate	42%	58%
5.3	Systems and procedures for project/program implementation and follow up are efficient (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.).	41	Moderate	63%	37%
Using Evaluation and Monitoring to Improve Development Effectiveness					
6.1	Systems and process for evaluation are effective.	39	Moderate	66%	34%
6.2	Systems and processes for monitoring and reporting on program results are effective.	50	Strong	32%	68%
6.3	Results based management systems are effective.	11	Weak	NA	NA
6.4	Evaluation is used to improve humanitarian and development effectiveness.	35	Moderate	91%	9%

*n = number of evaluations addressing the given sub-criteria
** Strong: n=45 – 52
Moderate: n= 30 – 44
Weak: n = under 30

Annex 4: Evaluation Quality - Scoring Guide and Results

	Evaluation Quality Scoring Guide		
	Criteria to be Scored	**Points**	**Score**
A	**Purpose of the evaluation is clearly stated.** The report describes why the evaluation was done, what triggered it (including timing in the project/program cycle) and how it was to be used.	3	
B	**The evaluation report is organized, transparently structured, clearly presented and well written.** There is a logical structure to the organization of the evaluation report. The report is well written with clear distinctions and linkages made between evidence, findings, conclusions and recommendations.	3	
C	**Evaluation objectives are stated.** Evaluation objectives are clearly presented and follow directly from the stated purpose of the evaluation.	2	
D	**Subject evaluated is clearly described.** Evaluation report describes the activity/program being evaluated, its expected achievements, how the humanitarian and development problem would be addressed by the activity and the implementation modalities used.	4	
E	**Scope of the evaluation is clearly defined.** The report defines the boundaries of the evaluation in terms of time period covered, implementation phase under review, geographic area, and dimensions of stakeholder involvement being examined.	4	
F	**Evaluation criteria** used to assess program effectiveness are clearly identified in the evaluation report and cover a significant number of the Common Criteria for Assessing humanitarian and development Effectiveness.	5	
G	**Multiple lines of evidence are used.** The report indicates that more than one line of evidence (case studies, surveys, site visits, and key informant interviews) is used to address the main evaluation issues. One point per line of evidence to maximum of 5.	5	
H	**Evaluations are well designed.** The methods used in the evaluation are appropriate to the evaluation criteria and key issues addressed. Elements of good design include: an explicit theory of how objectives and results were to be achieved, specification of the level of results achieved (output, outcome, impact), baseline data (quantitative or qualitative) on conditions prior to program implementation, a comparison of conditions after program delivery to those before, and a qualitative or quantitative comparison of conditions among program participants and those who did not take part.	5	
I	**Evaluation findings and conclusions are relevant and evidence based.** The report includes evaluation findings relevant to the assessment criteria specified. Findings are supported by evidence resulting from the chosen methodologies. Conclusions are linked to the evaluation findings as reported.	4	
J	**Evaluation report indicates limitations of the methodology.** The report includes a section noting the limitations of the methodology. It indicates any limitations in the design as well as any problems in the implementation (low survey returns for example) and describes how their impact on the validity of results and any measures taken to address the limitations (re-surveys, follow-ups, additional case studies, etc.	5	
K	**Evaluation includes recommendations.** The evaluation report contains specific recommendations that follow on clearly from the findings and conclusions. Further, the recommendations are specifically directed to one or more organizations and are actionable and aimed at improving humanitarian and development Effectiveness. (Objectives achievement, cross cutting themes, sustainability, cost efficiency or relevance).	5	
	Total Possible Score	**45**	

The criteria used for assessing evaluation quality were assigned weighted possible scores (from 2 to 5). These weights (potential total scores for a criteria) were based on the relative importance and the complexity of a given criteria. The definition of each quality assessment criteria contains a single verifiable component, which corresponds to each available point. For example, if criteria E is worth a possible four points, then the analyst would need to verify that each of the four components of the definition were present in order to award four points for this criteria. The verified components minimized the amount of subjectivity involved in the process of quality review.

Evaluation Quality Scoring Results

During the Pilot Testing of the Common Approach, the Management Group of participating development agencies guiding the work on behalf of DAC-EVALNET suggested grouping quality score results for each evaluation into groups of five (in total score). This was seen as presenting the best level of "granularity" and transparency. It allows independent observers to reach their own conclusions on the distribution of quality scores.

Evaluation Quality Scores in Groups of 5 (Max = 45)	Evaluations in Each Bracket (#)	Evaluations in Each Bracket (%)
41-45	18	35%
36-40	23	44%
31-35	6	12%
26-30	5	10%
21-25	0	0%
16-20	0	0%
11-15	0	0%
6-10	0	0%
0-5	0	0%
Total	52	100%

Annex 5: WFP Corporate Documents Reviewed

WFP Corporate Documents

Annual Performance Report for 2009, WFP, WFP/EB.A/2010/4 20 May 2010

"Annual Evaluation Report 2010", World Food Programme, 19 May 2011

"Annual Performance Report for 2010", World Food Programme, WFP/EB.A/2011/4, 13 May 2011

"Contributions to WFP: Annual Contributions and Five-Year Aggregate Ranking" http://documents.wfp.org/stellent/groups/public/documents/research/wfp232961.pdf Accessed December 2011

"Fighting Hunger Worldwide: Annual Report 2010" World Food Programme, 2011

"Financial Framework Review, Second Regular Session of the Executive Board 2010", November 9, 2010

"Financial Framework Review Options: Financial Issues", Presentation to Executive Board, WFP/EB.A/2010/6-E/1, 19 May 2010

"Gender Policy: Promoting Gender Equality and the Empowerment of Women in Addressing Food and Nutrition Challenges", World Food Programme, 10 February 2009

Management Response to the Recommendations of the Annual Evaluation Report 2009 WFP/EB.A/2010/7-A/Add.1, 11 May 2010

Monitoring and Evaluation Guidelines, WFP http://www.wfp.org/content/monitoring-and-evalutation-guidelines Accessed January 2012

"Resourcing for a Changing Environment", WFP, WFP/EB.1/2010/5-B/Rev.1 12 February 2010

Steps Forward: Implementation of WFP Strategic Results Framework (2008 – 2013), WFP, WFP/EB.1/2010/5-D, 21 January 2010

"WFP and the Environment: Issues and priorities", WFP, WFP/EB.3/98/3, 3 September 1998

WFP Evaluation Policy, WFP/EB.2/2008/4-A, 8 October 2008

"WFP Financial Framework Review Options: Financial Issues", WFP/EB.A/2010/6-E/1, 19 May 2010

WFP Gender Policy 2009: Promoting Gender Equality and the Empowerment of Women in Addressing Food and Nutrition Challenges, WFP, 2009

"WFP's Role in the Humanitarian Assistance System", WFP/EB.1/2010/5-C, 12 January 2010.

WFP Strategic Plan 2008-2013, World Food Programme

Other Documents

Assessing the Development Effectiveness of Multilateral Organizations: Approach, Methodology and Guidelines, Management Group of the Task Team on Multilateral Effectiveness, DAC EVALNET, 2011

"Peer Review: Evaluation Function at the World Food Programme", Sida, 2007

Annex 6: Guide for Review Team to Classify Evaluation Findings

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
Common humanitarian and development Evaluation Assessment Criteria				
1. Achievement of Humanitarian and Development Objectives and Expected Results				
1.1 Multilateral organization supported programs and projects achieve their stated development and/or humanitarian objectives and attain expected results.	Evaluation finds that less than half of stated output and outcome objectives have been achieved including one or more very important output and/or outcome level objectives.	Evaluation finds that half or less than half of stated output and outcome level objectives are achieved.	Evaluation finds that multilateral organization supported programs and projects either achieve at least a majority of stated output and outcome objectives (more than 50% if stated) or that the most important of stated output and outcome objectives are achieved.	Evaluation finds that multilateral organization supported programs and projects achieve all or almost all significant development and/or humanitarian objectives at the output and outcome level.
1.2 Multilateral organization supported programs and projects have resulted in positive benefits for target group members.	Evaluation finds that problems in the design or delivery of multilateral organization supported activities mean that expected positive benefits for target group members have not occurred or are unlikely to occur.	Evaluation finds that multilateral organization supported projects and programs result in no or very few positive changes experienced by target group members. These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency.	Evaluation finds that multilateral organization supported projects and programs have resulted in positive changes experienced by target group members (at the individual, household or community level). These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency.	Evaluation finds that multilateral organization supported projects and programs have resulted in widespread and significant positive changes experienced by target group members as measured using either quantitative or qualitative methods (possibly including comparison of impacts with non-program participants). These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency.

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
1.3 Multilateral organization programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national humanitarian and development goals.	Evaluation finds that multilateral organization supported projects and programs have not contributed to positive changes in the lives of beneficiaries as measured quantitatively or qualitatively.	Evaluation finds that multilateral organization supported projects and programs have contributed to positive changes in the lives of only a small number of beneficiaries (when compared to project or program targets and local or national goals if established).	Evaluation finds that multilateral organization supported projects and programs have contributed to positive changes in the lives of substantial numbers of beneficiaries as measured quantitatively or qualitatively. These may result from development, relief, or protracted relief and rehabilitation operations and may include the avoidance of negative effects of emergencies.	Evaluation finds that multilateral organization supported projects and programs have contributed to positive changes in the lives of substantial numbers of beneficiaries. Further, they have contributed to the achievement of specific national development goals or have contributed to meeting humanitarian relief objectives agreed to with the national government and/or national and international development and relief organizations.
1.4 Multilateral organization activities contributed to significant changes in national humanitarian and development policies and programs (including for disaster preparedness, emergency response and rehabilitation) (policy impacts) and/or to needed system reforms.	Evaluation finds that national policies and programs in a given sector or area of development (including disaster preparedness, emergency response and rehabilitation) were deficient and required strengthening but multilateral organization activities have not addressed these deficiencies.	Evaluation finds that multilateral organization activities have not made a significant contribution to the humanitarian and development of national policies and programs in a given sector or area of development, disaster preparedness, emergency response or rehabilitation. (Policy changes in humanitarian situations may include allowing access to the effected populations).	Evaluation finds that multilateral organization activities have made a substantial contribution to either re-orienting or sustaining effective national policies and programs in a given sector or area of development disaster preparedness, emergency response or rehabilitation.	Evaluation finds that multilateral organization activities have made a substantial contribution to either re-orienting or sustaining effective national policies and programs in a given sector or area of development disaster preparedness, emergency response or rehabilitation. Further, the supported policies and program implementation modalities are expected to result in improved positive impacts for target group members.

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
2. Cross Cutting Themes: Inclusive Humanitarian Assistance and Development Which can be Sustained				
2.1 Extent multilateral organization supported activities effectively address the cross cutting issue of gender equality.	The evaluation finds multilateral organization supported activities are unlikely to contribute to gender equity or may in fact lead to increases in gender inequities.	The evaluation report finds that multilateral organization supported activities either lack gender equality objectives or achieve less than half of their stated gender equality objectives. (Note: where a program or activity is clearly gender focused (maternal health programming for example) achievement of more than half its stated objectives warrants a satisfactory rating).	Evaluation finds that multilateral organization supported programs and projects achieve a majority (more than 50%) of their stated gender equality objectives.	Evaluation finds that multilateral organization supported programs and projects achieve all or nearly all of their stated gender equality objectives.
2.2 Extent changes are environmentally sustainable.	Evaluation finds that multilateral organization supported programs and projects do not include planned activities or project design criteria intended to promote environmental sustainability. In addition, the evaluation reports that changes resulting from multilateral organization supported programs and projects are not environmentally sustainable.	Evaluation finds that multilateral organization supported programs and projects do not include planned activities or project design criteria intended to promote environmental sustainability. There is, however, no direct indication that project or program results are not environmentally sustainable. OR Multilateral organization	Evaluation finds that multilateral organization supported programs and projects include some planned activities and project design criteria to ensure environmental sustainability. These activities are implemented successfully and the evaluation reports that the results are environmentally sustainable.	Evaluation finds that multilateral organization supported programs and projects are specifically designed to be environmentally sustainable and include substantial planned activities and project design criteria to ensure environmental sustainability. These plans are implemented successfully and the evaluation reports that the results are environmentally sustainable.

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
		supported programs and projects include planned activities or project design criteria intended to promote sustainability but these have not been successful.		

3. Sustainability

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
3.1 Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations, to rehabilitation, reconstructions and, eventually, to longer term humanitarian and development results.	The evaluation finds that there is a very low probability that the program/project will result in continued intended benefits for the target group after project completion. For humanitarian relief operations, the evaluation finds no strategic or operational measures to link relief, to rehabilitation, reconstruction and, eventually, to development.	The evaluation finds that there is a low probability that the program/project will result in continued benefits for the target group after completion. For humanitarian relief operations, efforts to link the relief phase to rehabilitation, reconstruction and, eventually, to development are inadequate. (Note, in some circumstances such linkage may not be possible due to the context of the emergency. If this is stated in the evaluation a rating of satisfactory can be given)	The evaluation finds it is likely that the program or project will result in continued benefits for the target group after completion. For humanitarian relief operations, the evaluation finds that strategic and operational measures to link relief to rehabilitation, reconstruction and, eventually, development are credible.	Evaluation finds that it is highly likely that the program or project will result in continued benefits for the target group after completion. For humanitarian relief operations, the evaluation finds that strategic and operational measures to link relief to rehabilitation, reconstruction and, eventually, development are credible. Further, they are likely to succeed in securing continuing benefits for target group members.
3.2 Extent multilateral organization supported projects and programs are reported as sustainable in terms of institutional and/or community capacity.	Evaluation finds that the design of multilateral organization supported programs and projects failed to address the need to strengthen institutional and/or community capacity as required. In the case of humanitarian operations,	Evaluation finds that multilateral organization programs and projects may have failed to contribute to strengthening institutional and/or community capacity or, where appropriate, to strengthen local capacities for delivery of relief	Evaluation finds that multilateral organization programs and projects may have contributed to strengthening institutional and/or community capacity but with limited success	Evaluation finds that either multilateral organization programs or projects have contributed to significantly strengthen institutional and/or community capacity as required or institutional partners and communities already had the required

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
	the design of programs and projects failed to take account of identified needs to strengthen local capacities for delivery of relief operations and/or for managing the transition to rehabilitation and/or development.	operations and/or for managing the transition to rehabilitation and/or development.		capacity to sustain program results.
3.3 Extent multilateral organization humanitarian and development programming contributes to strengthening the enabling environment for development.	The evaluation finds that there were important weaknesses in the enabling environment for humanitarian and development (the overall framework and process for national humanitarian and development planning; systems and processes for public consultation and for participation by civil society in humanitarian and development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures, service delivery and quality; and necessary improvements to supporting structures, such as capital and labour markets). Further, the multilateral organization activities and support provided to programs and projects failed to address	Evaluation finds that multilateral organization humanitarian and development activities and/or multilateral organization supported projects and programs have not made a notable contribution to changes in the enabling environment for humanitarian and development.	Evaluation finds that multilateral organization humanitarian and development activities and/or multilateral organization supported projects and programs have made a notable contribution to changes in the enabling environment for humanitarian and development including one or more of: the overall framework and process for national humanitarian and development planning; systems and processes for public consultation and for participation by civil society in humanitarian and development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures, service delivery and quality; and necessary improvements to	Evaluation finds that multilateral organization humanitarian and development activities and/or multilateral organization supported projects and programs have made a significant contribution to changes in the enabling environment for humanitarian and development including one or more of: the overall framework and process for national humanitarian and development planning; systems and processes for public consultation and for participation by civil society in humanitarian and development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures, service delivery and quality; and necessary improvements to

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
	the identified weakness successfully, further limiting program results.		supporting structures such as capital and labour markets.	supporting structures such as capital and labour markets. Further, the evaluation reports that these improvements in the enabling environment are leading to improved humanitarian and development outcomes.
4. Relevance				
4.1 Multilateral organization supported programs and projects are suited to the needs and/or priorities of the target group	Evaluation finds that substantial elements of program or project activities and outputs were unsuited to the needs and priorities of the target group.	Evaluation finds that no systematic analysis of target group needs and priorities took place during the design phase of developmental or relief and rehabilitation programming or the evaluation report indicates some evident mismatch between program and project activities and outputs and the needs and priorities of the target group.	Evaluation finds that the multilateral organization supported activity, program or project is designed taking into account the needs of the target group as identified through a process of situation or problem analysis (including needs assessment for relief operations) and that the resulting activities are designed to meet the needs of the target group.	Evaluation finds methods used in program and project humanitarian and development (including needs assessment for relief operations) to identify target group needs and priorities (including consultations with target group members) and finds that the program and project takes those needs into account and is designed to meet those needs and priorities (whether or not it does so successfully).
4.2 Multilateral organization supported humanitarian and development projects and programs align with national humanitarian and development goals:	The evaluation reports that significant elements of multilateral organization supported humanitarian and development program and project activity run counter to national humanitarian and	The evaluation reports a <u>significant portion</u> (1/4 or more) of the multilateral organization supported humanitarian and development programs and projects subject to the evaluation are not aligned	<u>Most</u> multilateral organization supported humanitarian and development programs and projects are reported in the evaluation to be fully aligned with national plans and priorities as expressed	<u>All</u> multilateral organization supported humanitarian and development projects and programs subject to the evaluation are reported in the evaluation to be fully aligned to national humanitarian and

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
	development priorities with a resulting loss of effectiveness.	with national plans and priorities but there is no evidence that they run counter to those priorities.	in national poverty eradication and sector plans and priorities. Wherever multilateral organization supported programs and projects are reported in the evaluation as not directly supportive of national plans and priorities they do not run counter to those priorities.	development goals as described in national and sector plans and priorities, especially including the national poverty eradication strategy and sector strategic priorities.
4.3 Multilateral organization has developed an effective partnership with governments, bilateral and multilateral development and humanitarian organizations and NGOs for planning, coordination and implementation of support to development and/or emergency preparedness, humanitarian relief and rehabilitation efforts.	Evaluation finds that the multilateral organization experiences significant divergence in priorities from those of its (government, NGO or donor) partners and lacks a strategy or plan which will credibly address the divergence and which should result in strengthened partnership over time.	Evaluation finds that multilateral organization has experienced significant difficulties in developing an effective relationship with partners and that there has been significant divergence in the priorities of the multilateral organization and its partners.	Evaluation finds that multilateral organization has improved the effectiveness of its partnership relationship with partners over time during the evaluation period and that this partnership was effective at the time of the evaluation or was demonstrably improved.	Evaluation finds that multilateral organization has consistently achieved a high level of partnership during the evaluation period.
5. Efficiency				
5.1 Program activities are evaluated as cost/resource efficient:	Evaluation finds that there is credible information indicating that multilateral organization supported programs and projects (development, emergency preparedness, relief and rehabilitation) are not	Evaluation finds that multilateral organization supported programs and projects under evaluation (development, emergency preparedness, relief and rehabilitation) do not have credible, reliable information on the costs of	Evaluation finds that the level of program outputs achieved (development, emergency preparedness, relief and rehabilitation) when compared to the cost of program activities and inputs are appropriate even when the program design	Evaluation finds that multilateral organization supported (development, emergency preparedness, relief and rehabilitation) programs and projects are designed to include activities and inputs that produce outputs in the most

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
	cost/resource efficient.	activities and inputs and therefore the evaluation is not able to report on cost/resource efficiency. OR The evaluation reports that multilateral organization supported programs and projects under evaluation present mixed findings on the cost/resource efficiency of the inputs.	process did not directly consider alternative program delivery methods and their associated costs.	cost/resource efficient manner available at the time.
5.2 Evaluation indicates implementation and objectives achieved on time (given the context, in the case of humanitarian programming)	Evaluation finds that less than half of stated output and outcome level objectives of multilateral organization supported programs and projects are achieved on time, there is no credible plan or legitimate explanation found by the evaluation, which would suggest significant improvement in on-time objectives achievement in the future.	Evaluation finds that less than half of stated output and outcome level objectives of multilateral organization supported programs and projects are achieved on time but the program or project design has been adjusted to take account of difficulties encountered and can be expected to improve the pace of objectives achievement in the future. In the case of humanitarian programming, there was a legitimate explanation for the delays.	Evaluation finds that more than half of stated output and outcome level objectives of multilateral organization supported programs and projects are achieved on time and that this level is appropriate to the context faced by the program during implementation, particularly for humanitarian programming.	Evaluation finds that nearly all stated output and outcome level objectives of multilateral organization supported programs and projects are achieved on time or, in the case of humanitarian programming, a legitimate explanation for delays in the achievement of some outputs/outcomes is provided.
5.3 Evaluation indicates that multilateral organization systems and procedures for project/program	Evaluation finds that there are serious deficiencies in agency systems and procedures for project/program	Evaluation finds some deficiencies in agency systems and procedures for project/program implementation but does	Evaluation finds that agency systems and procedures for project implementation are reasonably efficient and	Evaluation finds that the efficiency of agency systems and procedures for project implementation represent an important

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
implementation and follow up are efficient (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.)	implementation that result in significant delays in project start-up, implementation or completion and/or significant cost increases.	not indicate that these have contributed to delays in achieving project/program objectives.	have not resulted in significant delays or increased costs.	organizational strength in the implementation of the program under evaluation.

6. Using Evaluation and Monitoring to Improve humanitarian and development Effectiveness

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
6.1 Systems and process for evaluation are effective.	Evaluation finds that evaluation practices in use for programs and projects of this type (development, emergency preparedness, relief and rehabilitation) are seriously deficient.	Evaluation finds that there is no indication that programs and projects of this type (development, emergency preparedness, relief and rehabilitation) are subject to systematic and regular evaluations.	Evaluation finds that program being evaluated is subject to systematic and regular evaluations or describes significant elements of such practice. No mention of policy and practice regarding similar programs and projects. This may include specialized evaluation methods and approaches to emergency preparedness, relief and rehabilitation programming.	Evaluation finds that program being evaluated (along with similar programs and projects) is subject to systematic regular evaluations or describes significant elements of such practice.
6.2 Systems and processes for monitoring and reporting on program results are effective	Evaluation finds an absence of monitoring and reporting systems for the development and humanitarian programming. This would include the absence of adequate monitoring of outputs during the implementation of humanitarian programming.	Evaluation finds that while monitoring and reporting systems for the development and humanitarian programming exist, they either do not report on a regular basis or they are inadequate in frequency, coverage or reliability.	Evaluation finds that monitoring and reporting systems for development and humanitarian programming as appropriate are well established and report regularly.	Evaluation finds that monitoring and reporting systems for the program are well established and report regularly. The quality of regular reports is rated highly by the evaluation and results are reportedly used in the management of the program.

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
6.3 Results Based Management (RBM) systems are effective	Evaluation finds that there is no evidence of the existence of an RBM system for the program and no system is being developed.	Evaluation finds that while an RBM system is in place, or being developed, it is unreliable and does not produce regular reports on program performance.	Evaluation finds that RBM system is in place and produces regular reports on program performance.	Evaluation finds that RBM system is in place for the program and there is evidence noted in the evaluation that the system is used to make changes in the program to improve effectiveness.
6.4 Multilateral organization makes use of evaluation to improve development/humanitarian effectiveness	Evaluation report does not include a management response and does not have one appended to it or associated with it. There is no indication of how the evaluation results will be used. There is no indication that similar evaluations have been used to improve effectiveness in the past.	Evaluation report includes a management response (or has one attached or associated with it) but it does not indicate which recommendations have been accepted: or there is some, non-specific indication that similar evaluations have been used to improve program effectiveness in the past.	Evaluation reports includes a management response (or has one attached or associated with it) that indicates which recommendations have been accepted. Or there is a clear indication that similar evaluations in the past have been used to make clearly identified improvements in program effectiveness.	Evaluation reports includes a management response (or has one attached or associated with it) describes a response to each major recommendation which is appropriate and likely to result in the organizational and programmatic changes needed to achieve their intent.

Annex 7: Recent Changes at WFP

Following the presentation of this review to WFP staff, they indicated that the organization had made a number of changes to address issues raised in its evaluations conducted between 2006 and 2011. They provided this input to be included as an annex to the report.

Recent initiatives undertaken by WFP to address some of the issues identified in this report include:

Effective Financing, Resource Allocation and Programme Continuity

Prioritization of Resources

WFP has strengthened its resource allocation process through the establishment by the Executive Director of the Strategic Resource Allocation Committee (SRAC). The SRAC, composed of a small group of senior managers, is mandated to maintain strategic overview of operational needs and shortfalls; prioritise areas for major appeals and fundraising; develop and document procedures allowing for transparent comparative prioritisation among all operations; and oversee the strategic prioritisation and allocation of all corporate resources. In 2010 and 2011, the strengthened process of resources prioritization and formalized process of the SRAC led to larger and more meaningful resources being allocated to fewer projects with most critical and acute food assistance needs.

The SRAC is also responsible for the prioritization of non-directed/multilateral contributions, which make up approximately 10% of WFP's overall resources, to those projects with the greatest needs. The SRAC also oversees the prioritization and allocation of development resources, in an annual exercise at the end of each calendar year. This enables an element of predictability as early as possible for programme planning for the year ahead. Periodic SRAC reviews allow for monitoring and adjustments as and when required.

In addition, from December 2010 to May 2011, WFP worked with the Executive Board Membership through three organized sessions to address four key aspects of prioritization: 1) Out of the approximately 925 million hungry people in the world, how do we select those who require WFP food assistance?; 2) From WFP's overall Programme of Work (around US$5-6 billion in recent years); which projects get funded and why?; 3) What happens when a WFP project receives only partial funding?; 4) How does WFP choose its priorities to communicate to donors?

To address these questions, WFP presented the progress, achievements and re-designed processes in three relevant areas: 1) the continued work on the alignment of our programme categories and the links with our Strategic Objectives; 2) the criteria and model for resource prioritization and allocation; 3) consequences of prioritization at the country level. WFP strives to be transparent regarding the organization's activities and processes, and these discussions were designed to ensure that our business model and practices work to serve the greatest needs.

Forward Purchasing

The establishment of a Forward Purchase Facility, piloted since 2008 and mainstreamed in 2011, has enabled the implementation of a demand-driven global approach to purchase food ahead of individual project requirements, based on aggregated needs and resource projections. This has created a paradigm shift, making the supply chain more responsive and managing project demands more proactively.

Benefits registered and experience gained through the pilot phase of this initiative in the Horn of Africa and Southern Africa regions resulted in the decision to make forward purchase an integral part of WFP's supply chain process at WFP. In early 2011, staffing capacity augmentation allowed WFP to undertake a detailed historical requirements trend analysis by corridor (or group of countries sharing the same supply routes); analysis of procurement and shipping lead times; and historical commodity sourcing markets. These elements were critical in establishing the rolling 'corporate' inventory approach and in determining the optimum level for this inventory to mitigate two opposing risks: that of not being at a magnitude to have considerable impact; and that of holding high stock levels, leading to significant additional costs.

Since mid-2011, US$80 to US$100 million, out of a ceiling set by the Executive Board of US$150 million in corporate financing funds, have been used on a continuous basis and 'revolved' three times with purchase requests amounting to approximately US$300 million for over two-thirds, or 510,000 metric tons, of all food commodities purchased by WFP for project implementation. The combined average lead-time gained by projects through the early start of the supply chain process, was of about 60 days; this approach contributed significantly to WFP's response capacity to the Horn of Africa crisis in the late summer of 2011, and presently in terms of meeting increasing needs in the Sahel region.

Advance Funding and Innovative Procurement Mechanisms

The availability of advanced financing has allowed WFP to initiate the purchase of food in anticipation of a crisis, rather than post factum. The resultant time savings have had a significant impact in limiting the decline in malnutrition levels of affected populations, and are particularly relevant for the purchase of blended foods, where long lead-times often exist for production and delivery. Advance financing also allows WFP to agree to production schedules, and provide purchase guarantees for local producers. Both stratagems have helped reduce transport costs – using ocean instead of air transport - as well as deliver nutritious products at the outset of the emergency, thus increasing their efficacy.

WFP is establishing a 24/7 Procurement Help Desk which will ensure that procurement offices in any country can access immediate assistance, in and outside of regular HQ working hours. This support office, in conjunction with a new procurement manual, and a new training programme and run in conjunction with an international procurement institute – will ensure that greater procurement delegation can be made to country offices.

Supply Chain Optimisation Systems

With the focus on a more efficient and effective WFP Supply Chain, WFP's Logistics Division has embarked on a Supply Chain Optimisation Project. Under this umbrella are two specific innovative state-of-the-art systems that aim to enhance sound decision making via strategic supply chain information:

- Supply Chain Information Maps provide a single and unified overview of field supply chain operations (corridors, stocks, lead-time, etc.) within any given country, facilitating better planning and management of operations through quick and centralised access to key real time information.

- The Supply Chain Import Parity System (SCIPS) is a real-time web-based system that supports country offices in managing their food requests. By providing a cost and lead time comparison for the several sourcing options, combining food prices, transport costs and lead time information, SCIPS then identifies which commodity source is most cost effective, and allows the country office to track their orders.

In order to support food supply chain management, WFP Logistics launched a project to introduce the SAP-based Logistics Execution Support System (LESS) into the corporate WINGS II system. The LESS aims to standardize commodity inventory management and create clear accountability assigned to all logistics movements that will be traceable in the system. The pilot was launched in two country offices in November 2011 and full roll-out to all WFP field offices has begun in 2012. Simultaneously, WFP Logistics is supporting WFP corporate efforts to enhance cooperating partners' capacities in the tracking and accounting of food for onward distribution to beneficiaries. In this respect, the LESS project works closely with WFP Programme Division, which is developing a new Corporate Monitoring and Evaluation Tool, "COMET."

In addition, WFP Logistics, in collaboration with Programming and Procurement Divisions, are developing a set of operational Key Performance Indicators that emphasise performance measures dealing with reliability, responsiveness, agility and the cost effectiveness of the supply chain.

Gender

WFP is committed to addressing areas of improvement, including reporting on WFP's results relating to gender equality issues; and the full integration of gender into WFP's policies, projects and guiding documents. These are consistent with the evaluation recommendations of the WFP Gender Policy 2003 – 2007 and the subsequent priorities of the Gender Policy and Corporate Action Plan.

At present, WFP's gender equality results indicators are being reviewed and revised to reflect WFP's shift from food aid to food assistance, with a view to measuring the results and impact of WFP's gender mainstreaming work. This process includes field consultations; review and comment of proposed indicators through the WFP Policy Committee; and a planned finalisation and roll out of new indicators by the end of 2012.

Strengthening the capacities of WFP staff to carry out gender analysis is a priority in WFP's Corporate Action Plan. While many Rome-based staff have already been trained on the IASC gender marker and the gender audit tool; a more extensive capacity development plan targeting country office colleagues will be implemented between March and July 2012. This is expected to facilitate the adoption by WFP of the IASC gender marker which will be used to assess all projects and grant proposals for gender integration. A recently deployed GenCap advisor to the Food Security Cluster, coordinated by FAO and WFP and hosted by WFP, is supporting the WFP Gender Service to implement this plan.

The Gender Innovations Fund, established in 2010, encourages and creates opportunities for WFP country offices to be innovative in establishing partnerships, designing and implementing context-led projects based on gender analysis. The Fund has supported 29 WFP country offices and their partners to implement a wide range of activities, including training women and girls in agro-processing for food and nutrition security in Sierra Leone; supporting women-led small-scale enterprises to supply ready-made school meals in El Salvador; working with indigenous women in Colombia to promote food and nutrition security; supporting the government of Indonesia to mainstream gender in their food and nutrition security action plan; preventing and mitigating gender-based violence in South Africa; and involving boys and men in food and nutrition-related responsibilities in Bolivia.

Performance Management

WFP has made significant progress to improve performance management. A Performance Management Framework was developed and endorsed by the Executive Board and has been applied to support the continuous improvement of performance in WFP. The framework contributes to enhanced results-based reporting and informed decision-making throughout the organisation.

Also in place are a Strategic Results Framework, and a Management Results Framework. Updated annually, these frameworks identify strategic and management goals, expected outcomes and performance indicators on the corporate, regional, country and divisional levels. The Management Results Framework provides the basis for Annual Performance Plans that are developed by each WFP office and include standard performance indicators as well as customized indicators for particular units. All WFP operational projects are aligned to the Strategic Results Framework with the majority or projects now reporting on key outcome indicators. Project performance is reported annually through the Standard Project Reports (SPRs) and an analysis is published in the Annual Performance Report (APR). WFP is striving to strengthen its outcome reporting in SPRs, and in 2010 revised its APR to better demonstrate results, making it one of the primary accountability documents of the organisation.

In 2011, WFP made significant progress in institutionalizing a risk management approach throughout the organization, and most offices conduct regular risk self-assessment and maintain risk registers; critical components to an effective performance management framework. WFP is also implementing an entity-wide performance and risk management IT system, running on SAP, which will streamline and enhance performance and risk management functions at all levels of the organisation.

Also in 2011, WFP made progress in establishing a Country Strategy process. This process has become the principal planning tool to improve the focus of interventions and minimize the risk of overlap with other agencies.