

Peer Review of the Evaluation Function of the UN World Food Programme (2008-2013)

OECD DEVELOPMENT ASSISTANCE COMMUNITY – UN EVALUATION GROUP

April, 2014

Vol. I – Main Report



World Food Programme

wfp.org

Table of Contents

1. EXECUTIVE SUMMARY	I
1.1. Background to the Peer Review	i
1.2. Main Findings	i
1.3. Strategic Choices for the Future	iv
1.4. Summary Assessment against the Peer Review Criteria	vi
1.5. Recommendations of the peer review	viii
1.6. Further recommendations	x
2. INTRODUCTION	1
2.1. Background to Peer Reviews	1
2.2. Peer Review in WFP	1
2.3. Purpose of the Peer Review	1
2.4. Timing and Duration of the Peer Review	2
2.5. Advance Preparations for the Peer Review	2
2.6. Inception and Interim Reports	2
2.7. Peer Review Methodology	2
2.8. Limitations	4
3. THE PURPOSE AND BENEFITS OF EVALUATION	5
4. THE ENVIRONMENT FOR EVALUATION IN WFP	6
4.1. Organizational Change within WFP	6
4.2. Fit for Purpose	7
4.3. WFP Strategic Plan 2014-2017	7
4.4. Monitoring and Self-Evaluation (M&SE) Strategy	8
4.5. Business Process Review	9
4.6. Changes in the External Environment	10
4.7. External Assessments relevant to Evaluation	10
5. PROGRESS SINCE THE 2007 PEER REVIEW	11
5.1. Principal achievements	11
5.2. Unfinished Business	11
6. GOVERNANCE OF THE EVALUATION FUNCTION	12
6.1. The WFP Evaluation Policy	12
6.2. Independence of the Evaluation Function	12
6.3. Role of the Executive Board	14
6.4. Senior Management	16
6.5. Evaluation Committees	16
6.6. Internal and External Audit	16
6.7. Resources for Evaluation	17
6.8. Evaluation Strategy	18

7.	STRATEGIC CHOICES REGARDING THE WFP EVALUATION FUNCTION	18
7.1.	Model 1 – Centralized evaluation	19
7.2.	Model 2 - Centralized Evaluation plus Demand-led Decentralized Evaluation	19
7.3.	Model 3 – Centralized and decentralized evaluation function	19
7.4.	Selecting a Model	19
8.	THE OFFICE OF EVALUATION	20
8.1.	OEV Objectives and Strategy	20
8.2.	OEV Budget	21
8.3.	OEV Evaluation Planning and Selection	22
8.4.	OEV Evaluation Coverage	22
8.5.	Evaluation Types	23
8.6.	OEV Evaluation Portfolio	23
8.7.	Evaluation Criteria	24
8.8.	Joint Evaluations	26
8.9.	OEV Staffing	26
8.10.	OEV Team Management	27
8.11.	Vesting Independence in the Office of Evaluation	28
8.12.	Appointment of the Director Office of Evaluation	29
8.13.	Prioritising OEV Resources	29
9.	EVALUATION QUALITY	30
9.1.	The Evaluation Quality Assurance System (EQAS)	30
9.2.	Quality Assessment of WFP Evaluations	31
9.3.	Feedback from the field	32
9.4.	External Reviewers and Reference Groups	32
9.5.	Long Term Agreements with Evaluation Service Providers	33
9.6.	Potential areas for quality improvement	33
10.	MANAGEMENT RESPONSE TO WFP EVALUATIONS	34
10.1.	Management Response arrangements	34
10.2.	Quality of Management Response	35
10.3.	Implementation Follow-Up	35
11.	KNOWLEDGE MANAGEMENT	36
11.1.	No home in WFP for Knowledge Management	36
11.2.	Collation of evaluation lessons	36
11.3.	Publication of Evaluations	36
11.4.	Synthesis and dissemination of evaluation results	37
12.	APPLYING THE LEARNING FROM WFP EVALUATIONS	37
12.1.	OEV Knowledge Management Strategy	37
12.2.	Application of evaluations at country level	38
12.3.	Evaluation Training	39
12.4.	Application of Evaluation in WFP strategy and policy development	39
12.5.	Application of Evaluation in Country Strategy and Project Review	39

13. OPERATIONS EVALUATIONS	40
13.1. Deficit in Operations Evaluations 2008-2013	40
13.2. Progress in the OEV Operations Evaluations Series 2013-15	40
13.3. Evaluation of Emergency Operations	42
14. DECENTRALIZED EVALUATION	42
14.1. Activity levels and Expenditure for Decentralized Evaluation	42
14.2. Responsibilities for Supporting Decentralized Evaluation	43
14.3. Evaluation Capacity of regional bureaux	43
14.4. Roles and Responsibilities	44
15. EXTERNAL RELATIONS	45
16. SELECTED UN COMPARATOR EVALUATION FUNCTIONS	46
17. SUMMARY ASSESSMENT AGAINST THE PEER REVIEW CRITERIA	47
18. RECOMMENDATIONS OF THE PEER REVIEW	48
ABBREVIATIONS	53

LIST OF ANNEXES (AVAILABLE IN VOL. II)

Terms of Reference, Peer Review of the Evaluation Function of WFP	
Peer Review Schedule	
Assessment of Management Response to the 2007 Peer Review	
Summary of Quality Assessment of a sample of WFP evaluations	
Individual Interviews	
Documents consulted	
Summary results from the Staff Survey	
Evaluations completed by the Office of Evaluation - 2008-2013	
Long Term Agreements (LTAs)	
Evaluation Selection Criteria for OEV evaluations	
Management Response System	
Understanding Independent Evaluation, Decentralized Evaluation and Reviews	
Special Account for Operations Evaluations	
Comparator UN Evaluation Functions	
Comparative data on UN evaluation functions from the Joint Inspection Unit	
Alternative Models for the future of the WFP Evaluation Function	
(For context, refer to Section 7 of the main report)	
<i>Model 1</i> – Centralized evaluation	
<i>Model 2</i> - Centralized Evaluation plus Demand-led Decentralized Evaluation	
<i>Model 3</i> – Centralized and decentralized evaluation functions	
Summary of features by Model	
Provisional estimate of additional resources required by Model from PSA (or donor funds)	

1. EXECUTIVE SUMMARY

Background to the Peer Review

1. The first Development Assistance Committee/United Nations Evaluation Group (DAC/UNEG) peer review of the evaluation function at WFP was conducted in 2007. The Executive Director of WFP agreed that a second peer review should take place in 2013–2014, following the standards agreed by the Joint DAC/UNEG Task Force on Peer Reviews.
2. The review's core assessment question was: Are WFP's evaluation policy, function and products independent, credible and useful for learning and accountability purposes, as assessed by a panel of professional evaluation peers against United Nations Norms and Standards for Evaluation (2005) and the evidence base? A peer review provides an opportunity for evaluation peers to learn from each other, to exchange experience on good practice in evaluation, especially across the United Nations system, and to consider together how the evaluation function of the organization reviewed can be enhanced. This is the constructive spirit in which the panel offers its findings and recommendations.
3. The Executive Director requested an extended review period to allow the panel to consider the evolving organizational changes in WFP. The panel therefore made two visits to WFP Headquarters in Rome in May 2013 and January 2014. The Chair of the panel will present its final report at the Annual Consultation on Evaluation in May 2014, and the Board will consider the management response to the report at its 2014 Second Regular Session.

Main Findings

4. *Climate for evaluation.* There appears to be an increasingly favourable climate for evaluation in WFP. Evaluation enjoys a good reputation and receives considerable attention from senior management and the Board. The central Office of Evaluation (OEV) has been strengthened over the past six years, in terms of its resourcing, professionalism, and the relevance and quality of its evaluation products. Since the arrival of the Executive Director in 2012, evaluation has been given further impetus by her focus on demonstrating results and accountability and enhancing monitoring, reporting and evaluation in WFP, as set out in the Framework for Action. In "Fit for Purpose", the Executive Director's statement of intent for reforms in WFP, she committed herself to improving monitoring, reporting and evaluation, supported by a field-based evaluation function. WFP's transition from food aid to food assistance is reinforced in the new WFP Strategic Plan (2014–2017), and is moving WFP away from tried and tested programming approaches to innovative and less proven strategies, making evaluation even more important as a means of demonstrating what works.

5. *Progress since the 2007 peer review.* WFP has made significant progress since the first peer review, especially with the approval of the 2008 evaluation policy,¹ which provides clearer safeguards for the independence of evaluation, better definition of the roles of OEV and other actors in evaluation, and formalization of a balanced evaluation programme. The implementation of a system of eliciting a management response to each OEV evaluation has increased management accountability for evaluation results, while senior management and the Board recognize that the development of a quality assurance system has increased the quality, credibility and ownership of OEV evaluations. However, the management response to the 2007 peer review included other commitments against which limited or no progress has been made: no evaluation strategy has been developed; although the 2008 policy attempted to clarify evaluation terminology, staff are still confused by it; the implementation of monitoring systems to provide reliable data for evaluations has been slow; a Board sub-committee on evaluation has not been formed (and the panel does not support the establishment of such a committee); and a programme for increasing country-level evaluation capacity was launched, but has not been sustained. While resources for evaluation have increased since 2007, the 2008 evaluation policy did not establish a formula for safeguarding the allocation of these resources.
6. *Monitoring and self-evaluation strategy.* In 2012, WFP developed a monitoring and self-evaluation (M&SE) strategy, which reflects the revised approach to results-based management and addresses evaluation issues beyond the scope of the current evaluation policy. The strategy makes commitments to developing a menu of evaluation tools, and reiterates the WFP policy commitment to undertaking at least one self-evaluation during the life of every operation, with the assumption that the necessary funding will come from project resources. The strategy gives the regional bureaux a clear role in the oversight of decentralized evaluation but does not explain how decentralized evaluation capacity is to be built. Roll-out of the M&SE strategy did not begin until 2013 and the strategy is currently undergoing further development, led by the Performance Management and Monitoring Division (RMP). Dropping of the term “self-evaluation” is being considered, and OEV, regional bureaux and RMP are trying to delineate their roles, which the M&SE strategy has not clarified. A successful monitoring strategy is important in rectifying the widely recognized lack of reliable monitoring data, which is required to enhance evaluation evidence. The M&SE strategy is not a substitute for an evaluation strategy.
7. *Office of Evaluation.* OEV provides WFP with a strong central evaluation unit that produces high-quality evaluations. The Office has been professionalized since the 2007 peer review and has established its functional independence. It has delivered reliably and to schedule on its demanding programme of work, conducting the portfolio of evaluations agreed in consultation with management and the Board. OEV has improved its strategic targeting, quality and communication of evaluations through tailored products, and evaluation evidence has informed the reformulation of corporate policy and strategy. Currently, the Director of OEV has no responsibility for decentralized evaluation, putting WFP out of step with comparator United Nations organizations that have both centralized and decentralized evaluation functions. The WFP Circular on appointment of the Director of OEV allows the Director to return to employment with WFP at the end of his/her term, which is also out of step with other United

¹ WFP/EB.2/2008/4-A.

Nations organizations. The panel proposes that the Director of OEV's appointment be limited to a single, non-renewable term of six years.

8. *Engagement in evaluation.* OEV is increasingly engaging WFP management and staff in the evaluation process, especially at the start and end of evaluations, to increase ownership and improve the quality of evaluation reports, including their recommendations. OEV is also increasing the use of evaluation evidence to inform WFP decision-making. Without compromising its independence, OEV could support efforts to apply evaluation learning more fully in project design. There is also scope for greater engagement of national stakeholders in WFP country-level evaluations.
9. *Evaluation resources.* It is difficult for OEV to strike a balance among the various roles assigned to it by the evaluation policy. OEV has enjoyed significant budget increases since 2008, both in its regular budget and – since 2012 – via a special account set up to support a three-year (2013–2015) programme for reviving operation evaluations. Even so, demand considerably outstrips OEV's capacity to deliver. In consultation with senior management and the Board, OEV needs to clarify how much of its effort should go into conducting evaluations – and of which types – and how much into communicating evaluation results, strengthening evaluation across WFP, contributing to organizational learning, and engaging in inter-agency networks and system-wide evaluation. With its current programme of work, OEV cannot fulfil all of these roles with the resources at its disposal.
10. *Evaluation quality.* This review assessed the quality of 20 evaluation reports, including those for seven decentralized evaluations. While OEV evaluations were highly rated, the quality of decentralized evaluations was more varied, but was satisfactory to good overall. The Evaluation Quality Assurance System (EQAS) has played a major role in ensuring the consistently high quality of OEV evaluations and gives WFP perhaps the most comprehensive evaluation management and quality assurance process in the United Nations system. Despite these obvious benefits however, OEV is aware that EQAS may have become too onerous and needs to be simplified. Most OEV evaluations are now carried out via long-term agreements. According to OEV, this has increased unit costs, but not the average cost of evaluations; brought time savings for evaluation managers, especially in initiating evaluations; and widened the pool of available consultants.
11. *Management response to and follow-up on OEV evaluations.* The introduction of a management response system that makes management responsible for agreeing on, implementing and tracking responses represents an important step forward since the last peer review. The WFP system for tracking the implementation of management responses is one of the most systematic in the United Nations system, although – unlike the parallel audit tracking system – it is not available online. Some management responses to evaluation recommendations are formulaic, especially when recommendations and/or responsibility for follow-up are not clear. Some Board members are concerned that the impact of follow-up actions is not reported or validated.
12. *Decentralized evaluation.* Decentralized evaluation – evaluation not conducted by OEV – was a major topic of discussion during the peer review because, in contrast to centralized evaluation, its development has been neglected. The peer review and the Business Process Review (BPR) have prompted discussion about how, or whether, decentralized evaluation is to be enhanced.

According to data collected by RMP, WFP country offices are conducting increasing numbers of evaluations, and the assessment of a small sample of decentralized evaluation reports showed that some evaluations managed by country offices are of a satisfactory or good standard. It is not known how well these reports are used, responded to by management, or followed up on, and WFP Headquarters has not systematically collected decentralized evaluations to analyse their quality, as some comparator United Nations agencies do. WFP does not track investment in evaluation at decentralized levels, and WFP financial reporting systems cannot provide this information. As part of the BPR, OEV and RMP have commissioned an assessment of decentralized evaluation capacity, which will show strengths and weaknesses in the planning, conducting and use of decentralized evaluations. WFP has seriously underestimated the effort required to establish and maintain a decentralized evaluation function and has invested much less in decentralized evaluation capacity than, for example, the United Nations Development Programme or the United Nations Children's Fund. In Fit for Purpose, the Executive Director set out her agenda for organizational change and gave her commitment to establishing a field-based evaluation function. There is increased resolve to find sustainable means of financing decentralized evaluation, but WFP senior managers are not yet fully agreed that strengthening decentralized evaluation is the right strategy for WFP. Establishing an effective decentralized evaluation function may be the most challenging element of WFP's future evaluation agenda.

13. *Operation evaluations.* Since 2008, WFP has not been meeting its annual policy targets for operation evaluations – 10 centralized and 20 decentralized. For 2013–2015, OEV has an ambitious programme of more than 60 operation evaluations, backed by a special account that draws on project funds, with the intention of progressively transferring the responsibility for operation evaluations to regional bureaux. Under the Executive Director's commitment to field-level evaluation, OEV has obtained additional resources for operation evaluations and has launched the first 12, working with regional bureaux and outsourcing much of the management to companies with long-term agreements. However, there is still insufficient planning of how regional bureaux can manage such evaluations after 2015, and most regional bureaux are not equipping themselves for the task.
14. *Regional bureaux.* The appointment of six regional monitoring and evaluation (M&E) advisers in 2013 is a positive step, but the advisers – and their Regional Directors – are clear that their main focus is on monitoring, although evaluation is part of their job description. The need to improve monitoring across WFP, including by rolling out the Country Office Monitoring and Evaluation Tool, will demand considerable efforts from country offices and regional bureaux. At least in the medium term, monitoring is likely to command far more attention than evaluation at regional bureaux, including from the regional M&E advisers. It may not be realistic to expect these advisers to manage operation evaluations using the OEV outsourced management model, and regional bureaux will require additional resources if they are to develop decentralized evaluation capacity in country offices. The BPR assessment of decentralized evaluation capacity will throw more light on these two issues.

Strategic Choices for the Future

15. Senior management should decide the most appropriate model for the evaluation function in WFP, and discuss this model with the Board. The panel proposes the following three alternative models as a basis for decision-making. These are discussed in detail in the full report:

- ❖ *Model 1 – Centralized evaluation.* In model 1, WFP continues to look to OEV as the source of high-quality and impartial evaluation, with any decentralized evaluations seen as being internal lesson-learning exercises for the office concerned and providing a measure of accountability to donors. Decentralized evaluations receive limited support, with no expectation that OEV or regional bureaux will provide consistent support or quality assurance. The special account continues to fund OEV operation evaluations beyond 2015, but at a lower level than for the 2013–2015 series.
 - ❖ *Model 2 – Centralized evaluation plus demand-led decentralized evaluation.* Model 2 extends model 1 by adding OEV quality standards and guidance for decentralized evaluation; technical support from OEV and regional bureaux provided on demand to country offices commissioning evaluation to improve evaluation quality; and OEV reporting on the quality of decentralized evaluation reports. Regional bureaux generate annual regional plans for evaluation based on country offices' intentions and coordinated with OEV's programme of operation evaluations. (As in model 1, the latter continues at a reduced level after 2015.) There is no WFP-wide plan for resourcing, strengthening or systematizing the decentralized evaluation function. OEV is given the resources to establish a help desk and enhance the roster of evaluation consultants from which country offices can draw, in partnership with regional bureaux. Resources will be required for two additional full-time professionals in OEV and one evaluation professional at each regional bureau.
 - ❖ *Model 3 – Centralized and decentralized evaluation.* In model 3, WFP adopts an eight- to ten-year roadmap for expanding and enhancing the evaluation function at all levels, with regional bureaux supporting evaluations at the regional and country levels, through country-level evaluation officers and evaluation focal points, with backup from OEV. OEV has two full-time staff members dedicated to supporting decentralized evaluation, each regional bureau has a full-time regional evaluation adviser and country offices with large operations – for example, those projected to exceed USD 150 million per year – have dedicated evaluation officers. OEV assesses the quality of evaluations from all organizational levels. Decentralized evaluations are conducted according to set criteria and on a planned cycle agreed by the Regional Director in consultation with the Director of Evaluation.
16. Models 2 and 3 recognize the added value of decentralized evaluation; model 1 does not. Model 1 can be achieved within the current allocations to evaluation from the Programme Support and Administrative budget and special account, but would require reducing OEV's evaluation output by 10 to 20 per cent. Models 2 and 3 cannot be delivered within the current resource framework.
 17. All three models have advantages and disadvantages, and a case can be made for each of them. Model 1 is the most easily achieved because it is closest to the current situation. The panel prefers model 2 because it recognizes the reality and potential added value of decentralized evaluation and is in line with the Executive Director's intent set out in Fit for Purpose for the establishment of a field-based evaluation function. In many respects, model 3 represents an ideal model but requires investment in evaluation capacity at a level well beyond anything WFP has so far contemplated and for which there does not seem to be an appetite among senior management. The panel believes that establishing model 2 as a first step could produce measureable improvements in the quantity and quality of decentralized evaluations within two to three years. At that point, WFP could review progress and decide whether the greater

investment in its decentralized evaluation function is generating sufficient value added to take the next major step in developing evaluation across WFP.

18. The panel proposes that the Executive Management Group decide which model to follow in mid-2014, after the BPR assessment of decentralized evaluation capacity is complete. This decision will underpin WFP's response to the peer review, which will be considered by the Board at its 2014 Second Regular Session.

Summary Assessment against the Peer Review Criteria

19. The terms of reference for the peer review asked the panel to make an independent assessment of the WFP evaluation function against the three principal criteria of independence, credibility and utility.

Independence

20. Overall, the functional independence of centralized evaluation in WFP is high, without full structural independence. Intentionality² is well established for centralized evaluation and WFP is making well-planned and rational decisions about the selection of OEV evaluations. Functional independence of the centralized evaluation function is expressed as follows:
 - ❖ The Director of OEV consults WFP managers when drawing up the OEV work plan but retains the final choice over evaluation subjects in the work programme submitted to the Board.
 - ❖ The Director has full discretion in the selection of subjects for evaluation.
 - ❖ The Director has full authority over the management of OEV's human and financial resources for evaluation, but not over the level of those resources.
 - ❖ OEV has not generally sought funding for its activities directly from donors.
 - ❖ OEV is independent in supervising evaluators.
 - ❖ The Director of OEV submits evaluation reports to the Board and publishes evaluations on WFP's public website without interference from management.
21. The Director of OEV issues an annual evaluation report on progress against the OEV work programme, summarizing the main OEV evaluations for the year and making new recommendations based on a synthesis of OEV evaluation findings, which require a management response that is presented to the Board.
22. The panel considers that independence and utility could be strengthened by OEV assuming full ownership of its evaluation reports, as OEV is better placed than evaluation consultants to determine the final shape of each report and its recommendations.
23. There is a risk that WFP may overemphasize the need for independence in its decentralized evaluations, where the primary aims are impartiality and quality. Ensuring a degree of separation between the project manager and those organizing the evaluation will increase the level of confidence in being able to achieve a balanced, impartial result.

² Intentionality is an evaluation norm referring to a clear organizational intention to plan and use evaluation to inform decision-making and improve performance.

Credibility

24. The evaluations managed by OEV are highly credible and are respected by the Board, senior management and external parties. The panel finds that OEV evaluation reports are of high quality, backed by a highly structured set of standards and procedures in the EQAS.
25. Compared with other United Nations evaluation functions, WFP's central evaluation function performs well in terms of spending, numbers of evaluations, transparency and quality assurance. The Director and staff of OEV are seen as credible in inter-agency fora and evaluation networks and are often asked to participate because they make valuable contributions. OEV has competent staff but insufficient capacity for some of the functions it is expected to perform, especially in helping to ensure that evaluation results are utilized and in developing evaluation capacity across WFP. OEV's reputation remains strong, based on its generation of high-quality reports.
26. The good use that WFP makes of evaluation reports is further evidence of their perceived reliability and relevance. The low coverage of operation evaluations since 2010, compared with WFP's policy commitment, is being addressed. Future credibility depends on OEV ensuring that its portfolio of evaluations remains relevant to WFP as it evolves.
27. The credibility of the evaluation function is somewhat reduced by shortcomings in the management response system and by field staff's perception that evaluations are not always sufficiently realistic and rooted in the country context.

Utility

28. Over the review period, evaluation has become more accepted; the past tendency for it to be marginalized has been reduced. All types of OEV evaluation are seen to be relevant and useful and are used across WFP. OEV evaluations are a useful accountability tool as their reports are discussed by the Board, with the exception of those for single operation evaluations, which are presented in synthesis reports.
29. The review found several examples of WFP staff using evaluation to benefit country programmes and global policy and strategy development. In general, the learning from evaluation is limited by the lack of a corporate knowledge management function. However, OEV-managed evaluations are used well, not only for accountability to the Board, but also in influencing policy and strategy. They are also well used at the field level for immediate programme improvements.
30. OEV intends to make greater use of evaluability assessments, especially for WFP strategies. Evaluability assessments provide not only a check on whether the proposed intervention can be evaluated, but also a systematic analysis of its coherence, feasibility and metrics. This should help to increase the quality of assessed strategies. The panel considers that evaluations can be made even more influential by selectively targeting evaluation lessons towards WFP project design, evaluability and evaluation planning.

Recommendations of the peer review

31. The panel presents the following recommendations for consideration by WFP management and the Board.
32. The panel's recommendations are divided into two sets. Recommendations 1 to 10 apply regardless of the evaluation model that management selects under Recommendation 1. Recommendations 11 to 15 are relevant only if WFP selects an evaluation model that includes the intentional strengthening of decentralized evaluation, i.e. model 2 or 3.
33. **Recommendation 1: Evaluation model for WFP.** WFP management should take decisions concerning the most appropriate model for the evaluation function in WFP, using the three models set out in paragraph 15 and making sure that the necessary human and financial resources are made available to implement the preferred model. These decisions should be made when the results of the BPR assessment of decentralized evaluation capacity become available and before the management response to the peer review is submitted to EB.2/2014.
34. **Recommendation 2: Evaluation policy.** WFP should revise the 2008 evaluation policy in line with the selected evaluation model to ensure that financial resources for evaluation are protected; criteria for evaluation selection and coverage are clearly specified; the roles of the Board, OEV and management at different levels of the organization are clarified; evaluation terminology is revised; and the typology of evaluations is brought into line with current WFP practice. The policy should also specify WFP's role in developing national evaluation capacity, and the involvement and role of partners in country-level evaluation.
35. **Recommendation 3: Oversight of the evaluation function.** The Board should request the development of a set of key performance indicators to support its oversight of evaluation across WFP, giving OEV adequate time to establish the necessary systems. The panel suggests the following as a possible set of indicators:
 - 1) the numbers, types and coverage of evaluations taking place across WFP;
 - 2) the human and financial resources used for evaluation;
 - 3) progress in the development of WFP's capacity and competence in evaluation; and
 - 4) ratings of the quality of evaluations, including decentralized evaluations if model 2 or 3 is selected.
36. The Board should ensure that WFP management has systems and processes in place to maximize the use of evaluation results in policy and strategy development, as well as in project and programme design. The strategic use of evaluation results should be the Board's main focus in its deliberations with WFP management, over and above discussion of individual evaluations.
37. **Recommendation 4: Management response.** WFP management should improve the quality and effectiveness of management responses to evaluations, in particular by giving due attention to the ownership of follow-up. This requires the active engagement of relevant senior managers and other stakeholders during the evaluation process and beyond. Specifically:

- ❖ OEV should continue to improve the quality of evaluation recommendations by ensuring that they are clear, realistic and relevant to WFP's context and structures. OEV should also strengthen its dialogue with evaluation consultants and management at the draft report stage, including through its current good practice of holding workshops with major stakeholders so that, as far as possible, recommendations are agreed before the evaluation is complete – although OEV may retain recommendations with which management does not agree.
 - ❖ WFP management should nominate a member of the Executive Management Group to be responsible for engaging management in each OEV evaluation from its outset through to its presentation to the Board, with RMP in a supporting and advisory role.
 - ❖ WFP should provide staff with online access to the RMP database for tracking progress in implementing the management responses to recommendations, so that the staff responsible can enter completed actions online, as they already do for audit recommendations.
 - ❖ WFP should replace the current management response template with a more detailed format for setting out action plans in which management can specify more clearly how it intends to act in response to each evaluation.
 - ❖ WFP management may request OEV's informal comments on how well a draft management response corresponds to the findings, conclusions and recommendations of an evaluation. The management response will still be fully owned by management, and OEV will in no sense be approving or taking responsibility for its content.
38. **Recommendation 5: Vesting evaluation independence in OEV.** For each evaluation it manages, OEV should take ownership of all aspects of the evaluation report, including the recommendations. Independence is vested in OEV, not in evaluation consultants.
39. **Recommendation 6: Evaluation quality assurance.** OEV should commission an external consultant to review the EQAS and identify evaluation steps and elements that can be simplified, reducing the time inputs of staff and consultants, the number of steps and the duration of evaluations. OEV should also take account of the panel's proposals on improving evaluation quality.
40. **Recommendation 7: Utilization of evaluation.** In redesigning its project and programme planning and approval process, WFP management should ensure that evaluation evidence is taken into account, not only from any evaluations of the project under consideration, but also from other relevant evaluations; management should request OEV's assistance in identifying such evaluations if necessary.
41. OEV should strengthen its inputs to WFP's revision of project and programme planning, design and approval processes to encourage the use of evaluation evidence and improve arrangements for evaluation within projects, in particular by promoting the design of projects that can be evaluated effectively (evaluability); the use of prior evaluations of a project, and other relevant evaluations; and planning to facilitate evaluation from the outset.

42. **Recommendation 8: Evaluation training.** OEV should develop evaluation training in partnership with the Human Resources Division, including modules for evaluation planning and management in WFP management training; and WFP induction courses covering the essential elements of evaluation concepts and purposes, and the evaluation policy.
43. **Recommendation 9: Roles and responsibilities.** WFP management should ensure that regional monitoring and evaluation advisers have a reporting line to OEV for technical oversight and support on evaluation, and review of their performance with regard to evaluation.
44. WFP management should clearly delineate the roles of OEV and RMP, giving OEV the lead responsibility for evaluation standards and guidance at all levels of WFP, with RMP taking the lead on monitoring and all forms of project/programme review by management.
45. WFP management should request Regional Directors to take responsibility for receiving and reviewing management responses to the single operation evaluations and decentralized evaluations in their regions.
46. Internal Audit and OEV should agree how to identify audit risks regarding compliance with the WFP evaluation policy, and should develop standard questions for testing the compliance of individual country offices, regional bureaux or Headquarters divisions where risks of non-compliance are identified.
47. **Recommendation 10: Monitoring and evaluation guidance.** OEV should work with RMP to make sure that WFP guidance on evaluations and, separately, on monitoring and review are distinct but mutually compatible and understood by country offices; that any regional and global training of M&E officers is coordinated; and that management reviews are used as key inputs to operation evaluations, country portfolio evaluations and other types of evaluation.

Further recommendations

48. Should WFP select an evaluation model that includes further enhancement of decentralized evaluation – evaluations managed by regional bureaux and country offices – the following recommendations would also apply.
49. **Recommendation 11: Evaluation strategy.** OEV should develop an evaluation strategy in line with the selected model for evaluation. The evaluation strategy, separate from the monitoring and review strategy, should set out how WFP will develop evaluation capacity, resourcing, selection, coverage, and utilization across the Organization.
50. **Recommendation 12: Role and designation of the Director of Evaluation.** To support the Board's governance of the function, the Board should request the Director of OEV to oversee and report on the evaluation function across WFP. In its annual report, OEV should include an assessment of the quality of decentralized evaluations and the salient issues emerging from these evaluations. In line with this change in role, the Board should request WFP management to redesignate the Director of OEV as the Director of Evaluation.

51. To avoid any possible conflicts of interest, the Board should limit the Director's term to a single period of six years, non-renewable, and without the right of return to WFP. This would replace the current arrangements – a term of four years renewable once, with the possibility of re-employment in WFP.
52. The Executive Director should redesignate the Director of OEV as the Director of Evaluation. The Director should retain direct responsibility for OEV-managed evaluation and should also be responsible for standard-setting, oversight and support of evaluation across WFP.
53. **Recommendation 13: Decentralized evaluation standards and guidelines.** OEV should develop appropriate and realistic standards for decentralized evaluations based on a simplified version of its EQAS and reflecting the norms and standards of the United Nations Evaluation Group. Once these standards are agreed, OEV should issue guidelines to country offices on the management of decentralized evaluations.
54. **Recommendation 14: Evaluation expertise.** WFP management should ensure that the WFP People Strategy includes the development of a staff cadre for assessment, monitoring and evaluation in WFP, so that regional bureaux and country offices have the human resource capacity and expertise to implement the evaluation strategy.
55. **Recommendation 15: Evaluation database.** OEV should develop an online database for all centralized and decentralized evaluations into which country offices and regional bureaux can upload their evaluation reports. OEV can use this database to make an annual assessment of the quality of evaluation reports, with a summary included in the Annual Evaluation Report. WFP may be able learn from the database solutions developed for this purpose by the United Nations Development Programme and the United Nations Children's Fund.

Response to all Recommendations

In responding to and implementing these recommendations, WFP will need to adopt a sequenced approach to the development and revision of documents relating to evaluation, as follows:

- ❖ Develop the management response to the peer review. This will require inputs from management and the Board, with advice from OEV (OEV should not prepare the response). The response will need to take into account the BPR assessment of decentralized evaluation capacity, due by May 2014.
- ❖ OEV adjusts its work programme for 2015 taking into account the results of the peer review and the BPR assessment of decentralized evaluation capacity.
- ❖ Revise the evaluation policy.
- ❖ Develop an evaluation strategy, either in parallel with or after revision of the evaluation policy, in line with the selected model for evaluation.
- ❖ Update the references to evaluation in programme guidance.
- ❖ Make sure that the evaluation strategy and the monitoring and review strategy are consistent.

2. INTRODUCTION

2.1. Background to Peer Reviews

1. The OECD-DAC Network on Development Evaluation and the UN Evaluation Group (UNEG) maintain a Joint Task Force to support professional Peer Reviews of the evaluation function of UN organizations. Each Peer Review is intended to identify good practice and opportunities to further strengthen the evaluation function in the agency under review, with a view to contributing ultimately to improved performance in international development cooperation and humanitarian assistance. The Peer Review Panel will provide feedback on the Peer Review process to the DAC-UNEG Joint Task Force on Peer Reviews to contribute to the further development of the peer review instrument.

2.2. Peer Review in WFP

2. The first DAC-UNEG Peer Review of the evaluation function at the World Food Programme (WFP) was conducted in 2007³. The Executive Director (ED) of WFP agreed that a second Peer Review should take place in 2013-14, following the standards agreed by the Joint DAC-UNEG Task Force on Peer Reviews⁴. The Peer Review Panel comprises:
 - ❖ Colin Kirk, Director of Evaluation, UNICEF (Chair of the Panel);
 - ❖ Henri Jorritsma, formerly vice-chair of the DAC Evaluation Network and Deputy Director of the Policy and Operations Evaluation Department of the Netherlands Ministry of Foreign Affairs;
 - ❖ Susanne Frueh, Executive Secretary, UN Joint Inspection Unit, and member of the peer review panel, 2007.
3. The Panel was assisted by consultant advisers Simon Lawry-White and Jessica Alexander.

2.3. Purpose of the Peer Review

4. The Peer Review aims to provide an independent peer assessment of WFP's evaluation function, with the primary audiences being WFP's Senior Management and the Executive Board (the Board), as well as the Office of Evaluation (OEV). As per the Terms of Reference (TOR) for the Peer Review (see Annex 1), which was approved by the Panel members and shared with WFP's Office of Evaluation and WFP Management, the core assessment question is:

“Are the agency's evaluation policy, function and its products: independent; credible; and useful for learning and accountability purposes, as assessed by a Panel of professional evaluation peers against the UN Norms and Standards (2005) and the evidence base?”⁵.

³ Report of the Peer Review on the Evaluation Function at the World Food Programme, Jock Baker, Stefan Dahlgren, Susanne Frueh, Ted Kliet, Zenda Ofir, 2007.

⁴ See UNEG, 2011, UNEG Framework for Professional Peer Reviews of the Evaluation Function of UN organizations, UNEG/REF (2011)/1.

⁵ UNEG, 2011, UNEG Framework for Professional Peer Reviews in the Evaluation Function of UN organizations, UNEG/REF (2011)/1.

5. The review focuses on the quality, use and follow up of evaluations across WFP to promote accountability, learning, and performance improvement. The Peer Review has taken the WFP Evaluation Policy of 2008 and the Peer Review of 2007 as its starting points and includes an assessment of the relevance of the evaluation policy and the extent to which the recommendations from the 2007 Peer Review have been implemented. WFP senior management and the Board have been engaged in the review. The Panel and its consultant advisors have received excellent support from the ED and senior managers and from OEV in conducting the review, reflecting the high priority given to the exercise by the organization.

2.4. Timing and Duration of the Peer Review⁶

6. The ED requested that the peer review be conducted over an extended period from January 2013 to February 2014 to allow the Panel to be fully informed by the on-going organizational reforms in WFP. The Chair of the Panel will present its findings to Annual Consultation on Evaluation in May 2014, and the Executive Board will consider the management response to the Panel's report in November 2014.

2.5. Advance Preparations for the Peer Review

7. In 2012, the then newly appointed Director of OEV conducted an Internal Review of OEV generating a series of recommendations for the improvement of the work of the Office. Many of the recommendations have been implemented, and, as a first step in the Peer Review, OEV completed the UNEG Self-Assessment in January 2013⁷.

2.6. Inception and Interim Reports

8. An Inception Report was prepared and agreed with the Panel following the senior consultant adviser's first visit to WFP in February 2013. Following the Panel's visit in May 2013, an Interim Report was prepared, including a set of Interim Observations for the consideration of OEV and senior management and submitted in July 2013. Senior management discussed these observations and OEV made some adjustments to its 2014 work programme, based on the Interim Report.

2.7. Peer Review Methodology

9. The Peer Review methodology has comprised:
 - ❖ **Document Review.** WFP provided a very substantial set of background documentation for the consideration of the Panel. Further documents were identified during the review. (See Annex 6 for a full list of documents consulted).
 - ❖ **Quality review of evaluation reports.** 20 WFP evaluation reports were quality reviewed against a set of quality criteria agreed with OEV. Sampling of evaluations was purposive,

⁶ A schedule for the main elements of the Peer Review is included as Annex 2.

⁷ OEV completed the UNEG Self-Assessment, using a Normative Framework proposed by the Chair of the Panel. At the request of the Chair, the framework was completed twice, once for centralized evaluation and a second time for decentralized evaluation. The framework could be only partially completed for decentralized evaluation given that OEV has limited knowledge of, and responsibility for, decentralized evaluation function.

not random. Seven of these reports were from decentralized evaluations. Five OEV evaluations were subject to 'longitudinal assessment', to track the path from evaluation selection through to management response and, where possible, how the evaluation results have influenced decision-making and policy. The summary results of these assessments are included as in Annex 4. A separate report is available containing the full results of the quality assessment, together with the assessment methodology⁸.

- ❖ **Panel Visits.** The Panel visited WFP Headquarters in Rome in May 2013 and January 2014. The Panel met informally with the Board Bureau on its first visit and with interested Board members⁹ during its second visit, with the ED and other senior managers, including a teleconference with Regional Directors, the Director and staff of OEV, and other managers and technical staff at HQ.
- ❖ **Key Informant interviews.** In addition to the key meetings, the senior consultant held interviews in person and by phone with members of OEV, with a number of directors at WFP Rome, Regional Directors and with country directors and evaluation managers involved in the evaluations under longitudinal assessment.¹⁰ (See Annex 5 for a schedule of interviewees).
- ❖ **On-Line Surveys.** The Peer Review included two on-line surveys, one for Board members¹¹ and one for WFP staff. From a membership of 32 Board members, nine responses were received (28%). The WFP Chief of Staff sent a second survey to all country offices, regional bureaux, and HQ divisions, resulting in 46¹² responses from a possible 115 (40%). Two-thirds of these responses came from staff at D1 and P5 level, and a third of the respondents were HQ based. The summary results from the staff survey are included as Annex 7.
- ❖ **Benchmarking.** Some comparisons with UN organizations are provided, in particular the Rome-based agencies, FAO and IFAD, and for decentralized evaluations, with UNDP and UNICEF. (See Annexes 14 and 15). With permission of the relevant Inspector, provisional results from a UN Joint Inspection Unit (JIU) report on the evaluation functions in the UN system organizations have also been used. (The JIU report is due for completion in Q1 2014).

⁸ Evaluation reports were assessed using a template agreed by the Panel in discussion with the Director OEV. There is no agreed DAC/UNEG model for evaluation report quality assessment. The set used here were derived from current models from agencies, donors and networks, including UNEG, ALNAP, UNFPA, UNIFEM, OECD-DAC, UNICEF, and WFP. The criteria are included with the results of the assessment in Annex 4.

⁹ Afghanistan, Cameroon, Cyprus, Ethiopia, EU, Germany, Italy, Russia, South Africa, Sudan, Switzerland, USA.

¹⁰ Interviews were also conducted with HQ staff responsible for results based management, knowledge systems, programme appraisal, management response to evaluations, internal audit, and the WFP Board Secretariat. 4 out of 10 of OEV's Long Term Agreement suppliers of evaluation services were interviewed, as were external partners, including the Chair of UNEG, the Director of ALNAP, and the Director/Acting Directors of Evaluation at FAO and IFAD, and the former Director of Evaluation at WFP.

¹¹ The Director of OEV met with the Board Bureau to explain the purpose of the Peer Review and to request Board feedback to a set of questions agreed by the Panel. The Bureau requested that these questions be issued as an on-line survey.

¹² 6 responses were discounted as they provided minimal information, leaving 40 for analysis.

2.8. Limitations

10. By design, a DAC-UNEG Peer Review is not a full-fledged evaluation. However, the Panel is satisfied that it has gathered the information required to make an independent assessment of WFP's evaluation function. The most significant information gap is on the state of the decentralized evaluation function in WFP, for which a separate assessment under the BPR is due for completion in May 2014.

3. The Purpose and Benefits of Evaluation

11. According to the UN Evaluation Group Norms for Evaluation;

- ❖ “An evaluation provides evidence-based information that is credible, reliable and useful, enabling the timely incorporation of findings, recommendations and lessons into the decision-making processes of the organizations of the UN system and its members”¹³.
- ❖ “Purposes of evaluation include understanding why, and the extent to which, intended and unintended results are achieved, and their impact on stakeholders. Evaluation is an important source of evidence of the achievement of results and institutional performance. Evaluation is also an important contributor to building knowledge and to organizational learning. Evaluation is an important agent of change and plays a critical and credible role in supporting accountability.”

12. In summary, therefore, evaluation provides impartial and credible evidence of performance and results to feed decision-making and learning and to support accountability.

13. The importance of evaluation is being understood increasingly in WFP and some field staff and managers express a desire to see evaluation used and followed up more effectively. Nevertheless, several informants report that there is still a level of suspicion with regard to evaluation, with some seeing it as a threat and not relevant to an action-oriented organization like WFP.

14. The establishment of a well-respected, professional evaluation unit in WFP is certainly a considerable achievement. If there is a disadvantage, it is that OEV is seen to be primarily serving the Board, and WFP managers have not yet understood fully how they can also use evaluation to good effect at each level of the organization.

15. The importance of independence in decentralized evaluation should not be overstated. WFP should not compromise the hard-won independence of OEV, nor should it discard decentralized evaluation that only meets the independence test partially (see also Annex 12).

16. The benefits of evaluation include the following:

- ❖ Evaluation enhances the reputation of the organization with a wide range of stakeholders and has become an expected part of any UN agency’s holding itself accountable and showing that it is a learning organization;
- ❖ Evaluation maintains the confidence of the Executive Board that the organization is managing itself self-critically and making course corrections based on that learning;
- ❖ The organization can use evaluation to test the validity of its strategies, including its strategic plan, in achieving its goals;

¹³ UNEG Norms for Evaluation, 2005, updated 2012 http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=21

- ❖ Evaluation is an investment in donor confidence. Spending US\$50-250,000 on evaluation that helps to secure a further US\$5-25 million or more in funding makes good business sense;
- ❖ Evaluations create a pool of learning based on practical experience that the organization can use to improve its project and programme design and to make itself more effective;
- ❖ Evaluation can be used to influence government. Professional WFP evaluation reports, and evaluation teams, can provide governments with high quality analysis they might not receive any other way, and can be used to transmit findings to the government that the WFP country team could not do alone;
- ❖ The evaluation process allows WFP teams an opportunity to reflect, including with other parts of the organization, that seldom arises without an exercise that requires a 'pause for thought'.

17. OEV could consider the development of a communication strategy on the purpose and value of evaluation in general, in addition to communicating evaluation results.

4. The Environment for Evaluation in WFP

4.1. Organizational Change within WFP

18. The Peer Review is taking place at a time of major change within WFP, with significant implications for the evaluation function. Key events and developments include:

- ❖ The arrival of WFP's new Executive Director in April 2012;
- ❖ The arrival of the new Director of the Office of Evaluation in January 2012;
- ❖ Agreement of the WFP Strategic Plan 2014-2017, and the accompanying Strategic Results and Management Results Frameworks;
- ❖ Rollout of the Framework for Action¹⁴ intended to strengthen implementation of current strategic priorities and enhance WFP's efficiency, effectiveness and accountability;
- ❖ The Business Process Review, including the stream on Monitoring, Reporting and Evaluation;
- ❖ The agreement of a Monitoring & Self-Evaluation Strategy¹⁵ in 2012, consistent with organizational changes flowing from the Framework of Action.

¹⁴ WFP, June 2012, Strengthening WFP – A Framework for Action.

¹⁵ WFP, 2012, Improving Performance through the implementation of a corporate monitoring & self-evaluation strategy, Decision of the Executive Policy Council 15.

19. The Panel has noted that evaluation is seen at all levels of WFP as a very important function that receives considerable attention by senior management and the Board. The ED is supportive of the evaluation function, as her decisions to increase resources for evaluation indicate. With the current organizational changes, the internal climate for evaluative activities is becoming, if anything, still more favourable.

20. Evaluation has moved further to the fore in WFP since the 2007 Peer Review because:

- ❖ OEV products are taken seriously by WFP management and the Board as tools for accountability and learning;
- ❖ WFP is becoming more results-oriented and the Executive Director is focusing the organization on improving its ability to demonstrate impact and replicate good practice;
- ❖ Both donors and WFP managers are keen that evaluation provides evidence of WFP results that allow financial support to be maintained;
- ❖ WFP has become a more innovative organization, especially with its move away from its established modus operandi of food aid towards food assistance, and there is a recognised need to understand which new approaches work, and which do not¹⁶.

4.2. Fit for Purpose

21. It is widely accepted that a lack of reliable monitoring data is a long-term weakness in WFP organizational systems that constrains effective management. The lack of data also has the potential to undermine the credibility of evaluation. In her statement 'Fit for Purpose'¹⁷ released a few months after her arrival at WFP, the ED stated:

“My commitment to improving performance monitoring and reporting is absolute – significant improvements can and will be made with respect to monitoring, reporting and evaluation, and to the systems supporting performance management results and accountability. In the future we will integrate the corporate monitoring and reporting system, *supported by a field-based evaluation function* (italics added).”

22. WFP has since set aside funds and initiated systems to improve its situation monitoring and performance monitoring.

4.3. WFP Strategic Plan 2014-2017

23. Evaluation features prominently in the new WFP Strategic Plan for 2014-2017 approved by the Board in June 2013. For example:

¹⁶ The Monitoring and Self-Evaluation Strategy (M&SE) states that, “With the shift from Food Aid to Food Assistance ...and the introduction of different options for transfer modalities, WFP needs a more strategic approach to monitoring and evaluation”. In part it is these ‘different options’ that drive the need for evidence and learning, and hence evaluation.

¹⁷ “On my first day as Executive Director, I told WFP staff: “Every decision we make, every programme we launch, every dollar we spend – each must be judged by its impact on the people who depend on us”, Fit for Purpose — WFP’s New Organizational Design, August 2012 (page 7).

- ❖ “To meet rising demand for evidence of impact in the coming years, organizations will need to ... demonstrate returns and support continuous quality improvements” (Para 6);
- ❖ “WFP will enhance staff capacity to...evaluate results and demonstrate evidence of impact through monitoring and evaluation against established metrics” (Para 62);
- ❖ “To further support results measurement for the Strategic Plan, WFP will ... ensure best practice monitoring and evaluation and quality reporting for WFP operations; target key projects with additional corporate monitoring and evaluation analysis, support and guidance; support independent evaluations and build country office capacity to undertake self-evaluations of programmes for improved project design, implementation and risk management” (Para 66).

24. The Evaluation Policy could be amended to reflect a stronger role for evaluation in performance improvement, in line with the UNEG Norms and the new WFP Strategic Plan.
25. The strategic plan’s ‘Risk Matrix and Mitigating Actions’ includes as a current action “Building staff skill sets and scaling up of M&E staffing and funding at regional bureaux and Headquarters levels”. The Panel is not aware of any development plan for evaluation skills or budget development for regional bureaux, apart from the limited-term appointment of, and training for, Regional M&E Advisers, and, to some extent, OEV’s work in developing arrangements for contracted-out operations evaluation management designed for regional bureaux to access.
26. The Strategic Results Framework developed following the agreement of the Strategic Plan implies a still greater role for evaluation, in particular to be able to understand why outcome and output level results are or are not being achieved under WFP’s five strategic objectives.

4.4. Monitoring and Self-Evaluation (M&SE) Strategy

27. The responsibility for the implementation of the M&SE strategy lies principally with the Performance Management and Monitoring Division (RMP) and with the regional bureaux. In support of the strategy, RMP has been assigned US\$16 million over three years. A team to lead the strategy implementation was established in 2012-13, including a Chief Monitoring Officer. The strategy includes the progressive roll out of the COMET¹⁸ system for monitoring of strategic and project level outcomes. The strategy is currently being revised to take account of the new Strategic Plan and the accompanying Strategic Results Framework.
28. RMP has launched a peer-review of the draft Monitoring and Evaluation Strategy (2014-17). Standard Operating Procedures have also been drafted. The Panel understands that RMP is proposing a move away from the use of the term ‘self-evaluation’, as this is found to be confusing for staff. The Panel supports the use of the term ‘review’ to capture the body of evaluative, lesson learning and programme improvement exercises undertaken by management (for example, mid-term review, after action review) and proposes that the strategy be renamed Monitoring and

¹⁸ Country Office Monitoring and Evaluation Tool.

Review Strategy to make a clear delineation from evaluation (see Annex 12 for more on the meaning of 'review').

29. At field level, monitoring and evaluation are associated functions, so proposing a separate strategy for evaluation might be seen as creating an unhelpful division. The Panel's purpose in doing so builds on experience from across the UN system that the pressures from management to obtain the necessary monitoring and reporting data from the field tends to leave evaluation marginalised. The results of interviews for this review also pointed in this direction. A separate evaluation strategy for evaluation will encourage due attention to evaluation and evaluation results, although field colleagues will need help to understand how to plan for M&E based on both an 'M&R' strategy and an 'E' strategy.

4.5. Business Process Review

30. One of four streams in the on-going Business Process Review is Monitoring, Reporting and Evaluation (MRE). The BPR has drawn heavily on the 'view from the field'. Messages from the country level summarised during the global BPR workshop in July 2013 included:

- ❖ The Monitoring and Evaluation Strategy needs stronger communication;
- ❖ Establish a single performance management process and policy (including evaluation);
- ❖ Strengthen and protect monitoring and evaluation budgets;
- ❖ Recognise the need for proper evaluation follow-up;
- ❖ Evaluation is not seen as a corporate priority, but should be.

31. In May 2013, RMP concluded a Learning Needs Assessment for Monitoring and Evaluation, which identified Operations Evaluations as both the area of lowest skill and the highest priority learning need for M&E in country offices.

32. The July 2013 presentation of BPR results included 'Improve field level evaluations¹⁹' under the MRE heading. Phase 1 was proposed as a 'quick win' under the MRE heading, and a decision was taken to allocate US\$222,500 towards a joint OEV-RMP initiative to assess WFP's decentralized evaluation capacity and propose options for the future development of the decentralized evaluation function, reporting in May 2014. A further US\$351,500 for follow-up to develop standards, guidance, training and systems is currently proposed for 2014, but not yet funded at the time of writing.

33. The BPR noted that 'country offices are inadequately enabled and supported to evaluate their operations and regional bureaux face similar constraints', and that there was a 'Lack of clarity on requirements, standards and options; resource-intensive evaluation approach; limited evaluation capacity, guidance, training and support; and financial constraints.'²⁰

¹⁹ The Panel takes the BPR references to 'field level' evaluation to be synonymous with decentralized evaluation.

²⁰ These were all in line with the Panel's observations in its July 2013 Interim Report.

4.6. Changes in the External Environment

34. As the new Strategic Plan acknowledges, “WFP is operating in an increasingly crowded field of state and non-state actors”; “Host governments are playing an even more active leadership role as partners align their food security and nutrition activities behind national plans and priorities”. Following the UN Quadrennial Comprehensive Policy Review (QCPR) 2012²¹, there is demand from Members States for accountability for UNDAF results to be strengthened and for streamlining and harmonization of UNDAF/country programming and processes, including system-wide reporting of results. Expectations are high for all UN operational agencies to be working more closely together in partnership. As a result, WFP should anticipate greater expectations for evaluations to be conducted with government or under government leadership, and further engagement with joint evaluation at country-level. The Panel noted interviews with WFP staff at all levels included very little discussion of the need for joint evaluation or government-led evaluation.
35. Since 2011, there has been considerable debate about ‘Resilience’ as a key concept for assistance to vulnerable situations, including in conflict and fragile states. WFP’s work on food assistance and food security fits well with this trend. WFP staff talk of trying to break down the artificial barriers between emergency and development work, which is essential for a resilience approach to succeed. In time, resilience programming may itself need to be evaluated.
36. Within evaluation circles, there is increasing emphasis on impact evaluation, for which WFP is ahead of the curve within the UN system and is recognised in evaluation professional circles in terms of impact methodology development.

4.7. External Assessments relevant to Evaluation

37. In 2011, the Canadian International Development Agency (CIDA) carried out a review of the humanitarian and development effectiveness in WFP, using WFP evaluations as its base data (discussed further under Quality of Evaluation below). An update to the 2011 DFID Multilateral Aid Review (MAR) was released in December 2013, in which, WFP was rated 3 out of 4 for progress, having been rated as ‘good value for money for UK aid’ in the original MAR report. The MAR noted that “WFP needs to achieve a results based culture’, and ‘accountability to partners and beneficiaries remains weak’. The 2013 update reports a ‘strong management commitment to embed a results culture across the organization’.
38. In parallel, the 2013 MOPAN²² assessment of WFP rated the use of evaluation, quality of evaluation and beneficiary engagement in evaluation as ‘adequate’, with independence of the evaluation unit ‘very strong’, and evaluation coverage of programme activities rated ‘inadequate’.
39. The UN Joint Inspection Unit has recently completed the first part of its review of evaluation functions across the organizations of the UN system, including WFP. The element of the review on decentralized evaluation is still ongoing. (JIU review results are referenced in 7.2, 9.3, 13.1 below).

²¹ <http://www.un.org/esa/coordination/2012qcpr.htm>

²² Multilateral Organisation Performance Assessment Network.

5. Progress since the 2007 Peer Review

5.1. Principal achievements

40. WFP has made good use of the 2007 Peer Review to strengthen evaluation in several areas, including:

- ❖ The issuing of a revised evaluation policy, which brought together various evaluation related documents, and provided; clearer safeguards for the independence of evaluation, better definition of the roles of OEV and various actors in evaluation, formalisation of a balance of evaluation types, and the establishment of a programme-specialist evaluation skills mix in the OEV team.
- ❖ A policy commitment to, and implementation of, a systematic process for management responses to all OEV evaluations, which accompany the evaluation reports presented to the Board, and with management being responsible for their completion and follow-up.
- ❖ The further development and refinement of a systematic quality assurance system (EQAS) that is recognised to have increased quality, credibility and ownership of evaluations.
- ❖ The initiation by OEV of a programme of work to increase the level of operations evaluations, backed by a Special Account established to ensure evaluation resources are available.
- ❖ Greater use of meta-evaluations and tailored reports that have allowed the learning from evaluations to be more widely used.
- ❖ Acquisition of the services of more skilled evaluation team, especially through the use of Long Term Agreements with established consultancy firms.

5.2. Unfinished Business

41. The management response to the last peer review also included commitments against which limited progress has been made:

- ❖ Evaluation resources have not been safeguarded: the evaluation policy provides no formula for defining or protecting the allocation of evaluation resources, even though evaluation resources have increased since 2008 and WFP has recently taken steps to make resources available on a more reliable basis for operations evaluations.
- ❖ The differences and relationship between self-evaluation, After-Action Reviews (AAR), decentralized evaluations and independent and external evaluations are not yet sufficiently articulated or understood in WFP.
- ❖ Progress has been slow in the implementation of monitoring systems to provide the reliable data on which evaluation depends.

- ❖ A Board sub-committee on evaluation was not formed (nor does the Panel support the establishment of such a committee).
- ❖ A programme for strengthening country level evaluation capacity was launched, but only as a temporary measure.
- ❖ OEV has a sophisticated process for its own work planning, but no evaluation strategy has been developed, as envisaged in the evaluation policy, that would allow for the clarification of evaluation roles and responsibilities and for the development of plans for evaluation capacity development across the organization.

42. All these areas remain relevant today. (Annex 3 provides a fuller assessment of progress against the 2007 Peer Review).

6. Governance of the Evaluation Function

6.1. The WFP Evaluation Policy

43. The development of the 2008 Evaluation Policy represented a major step forward for the evaluation function in WFP and, in the view of the former Director of OEV, was the most important outcome of the 2007 Peer Review. The policy is well aligned with the UNEG Norms and Standards, covers all key aspects, and shows no significant deviation from them²³. However, in several areas the 2008 policy falls short of WFP's current needs and ambitions. In the recommendations section below the Panel has suggested areas where the policy needs strengthening.

44. Areas requiring update include:

- ❖ The policy is weak with regard to safeguarding financial and human resources for evaluation, and hence ensuring adequate coverage;
- ❖ The policy will need to revise terminology related to evaluation, which has become confused;
- ❖ The policy focuses mainly on OEV and the roles of other parts of the organization in evaluation needs clear explanation, especially regional bureaux and country offices;
- ❖ The policy requires a clearer set of criteria for evaluation selection and coverage;
- ❖ The policy does not specify a role for WFP in national evaluation capacity development and the involvement and role of partners in evaluation at country level.

6.2. Independence of the Evaluation Function

45. The WFP Office of Evaluation enjoys a high level of functional independence in the conduct of evaluation and ranks highly on independence in comparison with other UN organizations. In the

²³ See http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=22

UN system, evaluation offices tend to report through the line management²⁴ and full structural independence is, *a priori*, not attainable²⁵ where an evaluation function reports to the chief executive, as in WFP. UNEG has recognised that a fully independent office reporting directly and exclusively to the governing body may not be possible or even appropriate for all UN organizations.

46. The Panel finds that all the following points from the evaluation policy regarding independence are being upheld:

- ❖ The Director of OEV consults WFP managers when drawing up the OEV work plan but retains the final choice over evaluation subjects in the work programme submitted to the Board²⁶. The Director has full discretion in the selection of subjects for evaluation.
- ❖ The Director has full authority over the management of OEV human and financial resources for evaluation, although the Director is subject to WFP staff rotation and other HR issues and does not control the level of financial resources for evaluation²⁷.
- ❖ OEV has not sought funding directly from donors for its activities (with a few exceptions).
- ❖ OEV is independent in supervising evaluators.
- ❖ The Director of OEV submits evaluation reports to the Board and publishes the evaluations on WFP's public website without interference from management²⁸.
- ❖ The Director of OEV issues an Annual Evaluation Report summarising the key OEV evaluations for the year and making fresh recommendations based on a synthesis of OEV evaluations findings, which themselves require a response from management that is presented to the Board.

47. Working relationships between OEV and other WFP stakeholders with a role in evaluation are generally good. The former Director of OEV maintained a greater separation of OEV from management than the current director at a time when an evaluation function that was independent, and seen to be independent, was being established. The current Director, building

²⁴ The WFP evaluation policy states that by reporting to the ED, "The Director and the Office of Evaluation are thereby independent of management functions responsible for the design, implementation and monitoring of WFP policies, strategies, operations and other work that may be subject to evaluation" (para 25).

²⁵ In FAO, the office reports both to the Director General and to the Programme Committee of the Council. Only in IFAD does the office report directly and exclusively to the Board, following the model favoured by the multi-lateral development banks, though as the IFAD President is both chief executive officer and chair of the Executive Board, even here complete independence is not achieved.

²⁶ In interviews, reference was made to one occasion where senior management tried to influence the content of a high profile evaluation report, but without the content of the report being materially affected.

²⁷ The Director of OEV has no influence over the overall envelope for evaluation financing and there are, in any case, no criteria for determining the size of the envelope.

²⁸ In the few cases where there have been disagreements between OEV and management over the way that evaluation findings have been expressed in evaluation reports, OEV has been able to determine the final wording. There have been rare occasions where managers have put inappropriate pressure on OEV managers to change the wording of evaluation reports.

on that platform, is engaging with management more in order that the function is as relevant and influential as possible, while not compromising the independence of the Office.

6.3. Role of the Executive Board

48. According to the Evaluation Policy, the Board has the following responsibilities with regard to evaluation (in summary, from para 15, 16):

49. The Board exercises an oversight function over evaluation in that it:

- ❖ Provides strategic guidance to the evaluation function;
- ❖ Reviews the work plan and budget as set out in WFP's Management Plan and Budget²⁹;
- ❖ Reviews the independence of evaluation.

50. The Board is also responsible for:

- ❖ Discussing selected evaluation reports;
- ❖ Taking decisions that guide management in its follow-up actions;
- ❖ Holding management responsible for ... management responses, and for follow-up including changes to policies and practices;
- ❖ Using evaluation findings and recommendations in its decision-making.

51. The Panel notes that the current evaluation policy does not spell out sufficiently clearly the oversight responsibility of the Board for the evaluation function across the organization, including both centralized OEV evaluations and decentralized evaluation. The nature of the Board's 'Strategic Guidance' is not explained in the policy.

52. The Executive Board has taken a very strong interest in evaluations in recent years and in 2010, asked that additional informal 'roundtable sessions' be held roughly two weeks before each Board meeting to allow for more extended discussions of evaluation reports than was possible in the formal sessions. The Board discusses all OEV evaluations in both formal and informal sessions. There are six sessions on evaluation per year with the Board, three formal and three informal, one of the informal occasions being the Annual Consultation on Evaluation held each May. This represents one of the most intense engagements of a governing board with evaluation reports anywhere in the UN system.

53. The Panel had some concerns regarding the role of the Board with regard to evaluation. In the Panel's view:

²⁹ Currently the Board only reviews OEV's workplan. It does not review plans for the evaluation function as a whole across the organisation and cannot be said therefore to be exercising oversight for the entire function.

- ❖ The Board should prioritise the maintenance of an effective evaluation function³⁰ over the consideration of individual evaluation reports, and proposes that a limited set of key performance indicators (KPIs)³¹ would help the Board fulfil its oversight role more fully³². It will take time for OEV to establish systems for data collection, supported by the regional bureaux, because information is not currently available on a reliable basis for any of the proposed measures.
- ❖ The Board should keep evaluation discussions at a strategic level and avoid focusing on specific details that risk losing sight of the ‘big picture issues’.
- ❖ The Board should pay attention to ensuring that management has systems in place for the use of evaluation results in policy and strategy development, and project and programme design, rather than demanding detailed information on the implementation of specific sets of recommendations. Moreover, in considering management responses, the Board should review the full set of responses to a given evaluation, without focusing narrowly on recommendations that the management did not accept.
- ❖ The Panel considers that the six Board discussions on evaluation each year could be used to better effect. The informal and formal sessions are too similar, with almost the same ground being covered in both. Consideration should be given to the informal and formal meetings having more distinct roles³³.

54. The Panel notes that the Board’s consideration of all OEV evaluations is not in line with current policy, which refers to the Board discussing selected evaluation reports.

55. Board members have a high level of confidence in OEV evaluations but are not satisfied with feedback they receive concerning the implementation of evaluation results. Currently, there is no obligation for management to compile progress reports for the Board on the results of

³⁰ An evaluation function includes:

- Resourcing arrangement;
- Quality standards and assurance;
- Guidance on the conduct of evaluation;
- Staffing levels, structure and reporting lines for evaluation (or monitoring and evaluation staff) ;
- Training and support for the commissioners of evaluation;
- Ensuring access to an adequate pool of external evaluators;
- The arrangements for the consideration and use of evaluations by the organization;
- How and where and by whom evaluations are published;

supported by an evaluation policy which sets out the organization’s intent to evaluate and the rules governing the application and resourcing of evaluation.

³¹ The proposed KPIs are: 1) the numbers, types and coverage of evaluations taking place across WFP; 2) the human and financial resources being used for evaluation; 3) progress in the development of WFP’s capacity and competence in evaluation; 4) the quality of evaluations, centralized and decentralized.

³² In UNICEF and UNDP, the inclusion of such KPIs in reports on evaluation to the Board has served to give the Board a more accurate understanding of the state of the function, and has also stimulated management to improve the quality and use of evaluations across the agency.

³³ The Director OEV has already started negotiating changes to the way the Executive Board reviews WFP’s evaluation work, for example, considering a selection of evaluations at the informal consultations, rather than all reports.

recommendations being implemented, other than percentage implementation rates (79% in 2013). However, as noted above, the Panel considers that the Board's efforts would be better targeted at ensuring processes are in place for evaluations results to be well used by the organization at all levels

6.4. Senior Management

56. The current WFP senior management, including the ED, Deputy Executive Director (DED) and Assistant Executive Directors (AEDs), support evaluation and see its value. They are also aware of the keen interest the Board takes in the function. In a meeting with the DED and AEDs, the following points were highlighted:

- ❖ The quality of OEV evaluations is improving and some are touching on issues at the core of the organization;
- ❖ Senior managers appreciate the advice on evaluation coming from OEV;
- ❖ Management response systems have weaknesses; bureaucratic management responses should be avoided.

57. The DED wants the organization to be aware that WFP will be expected to be part of nationally-led evaluation and joint evaluation with partners, especially UN agencies, as joint programming becomes more common but this does not seem to be a live discussion amongst managers at any level.

6.5. Evaluation Committees

58. For some years, WFP has had no internal evaluation committee, and has never had a Board level evaluation committee. The Policy Committee, which was an important venue for ensuring evaluation findings and recommendations were considered in the formulation of WFP Policy, has been discontinued.

59. In the current favourable environment for evaluation, the Panel sees no case for establishing either a Board committee on evaluation or an internal management committee on evaluation³⁴. There may be a case, however, for a time-limited and informal group of Board members to accompany the implementation of the recommendations from this Peer Review, and any further proposals on evaluation emerging from the BPR.

6.6. Internal and External Audit

60. The Directors of OEV and of Internal Audit both reported good communication between their offices. There is no systematic connection between audit and evaluation exercises³⁵. In preparing

³⁴ Several years back, there was an internal evaluation committee that served as a bridge between senior management and the Office of Evaluation at a time when there was considerable tension between the two. This no longer applies.

³⁵ Senior management raised the question of the overlap between evaluation and programme audit, prompted in part by the external auditors (currently India) performance audit on the 'Use of Cash and Vouchers', many of the recommendations from which could readily have appeared in an OEV evaluation report. The External Auditor and OEV have already discussed this report.

internal audits of country offices, the auditors ask for past evaluations from the country office and use them as background to inform the audit. Audit reports, which are now public, also provide source material for evaluations.

61. The Panel sees the potential for the promotion of evaluation by including 2-3 targeted questions on evaluation in internal audits. The internal auditor should review whether the country office is complying with the evaluation policy, as with all other WFP policies, and review whether evaluation recommendations accepted by country office management have been implemented.
62. WFP may also wish to consider 1) how to communicate the different or complementary roles of evaluation and audit³⁶, 2) how/whether OEV findings and recommendations can both draw from, and be synthesized within, the External Auditor's findings and recommendations.

6.7. Resources for Evaluation

63. Historically, the resourcing of evaluation, especially operations evaluations, has been a challenge for WFP. Since 2008, the budget for centralized evaluation has increased significantly in real terms. However, within operations, evaluation is one of the first items to be cut if funding falls below the budget. This reflects WFP's funding regime, where part-funding of projects is common, but also the weakness of the WFP evaluation policy in not setting criteria for evaluation coverage, funding and protection of evaluation resources.
64. Organizations use various ways of ensuring that evaluation expenditure is protected. Some have a guide for the level of programme expenditure that should go to evaluation (for example UNICEF 1%, USAID 3%), while others use triggers. For example, in both FAO and IFAD, projects/programmes over a certain size are automatically evaluated. The 2008 WFP evaluation policy set no criteria for automatic external evaluation of projects of a specified scale, although the previous 2003 evaluation policy did³⁷.
65. This is a significant gap in WFP's current evaluation policy and leaves the evaluation function exposed to arbitrary decisions about the allocation of resources for evaluation, both at the centre and at the decentralized levels. This needs to be addressed when revising the evaluation policy. Here, much will also depend on the model adopted for the future development of the evaluation function, as discussed in the next section.
66. The policy commitment to 30 operations evaluations per year (10 centralized, 20 decentralized) in the 2008 evaluation policy is arbitrary, not strategic, and is not related to the total number of operations or different types of operations, or their scale. The Panel proposes that a more rational set of criteria for targeting evaluation be added to the evaluation policy base (see 7.7 below).

³⁶ There are indications that some WFP staff do not understand fully the difference between audit and evaluation, and regard evaluation as a form of audit. (This is common in UN agencies).

³⁷ 2003 Evaluation Policy, para 14 b. "any operation if the cumulative budget of all phases exceeds US\$ 50 million and if the previous evaluation took place more than three years previously".

67. WFP's financial systems are not capable of generating financial expenditures on evaluation across the organization³⁸. The Director of OEV currently has no responsibility for tracking such expenditure but the Panel proposes that the Director should be enabled to do so as part of an expanded oversight role. The Panel further suggests that the responsibility for collecting decentralized evaluation reports, together with data on their costs, be passed to OEV.
68. A revised evaluation policy should have a clearer set of criteria for evaluation selection and evaluation coverage than the current policy. Together with an agreed set of evaluation types, the level of resources required to meet the policy commitments can then be determined. In this way, the Panel considers that WFP can provide itself with a transparent and systematic approach to defining, allocating and protecting resources for evaluation.

6.8. Evaluation Strategy

69. The evaluation policy commits WFP to develop an evaluation strategy, which has not been done. A strategy is required, *inter alia*, for the improvement of decentralized evaluation, to cover skills development, development and testing of standards, and the synthesising of learning from evaluations, none of which were covered by the 2012 Monitoring and Self-Evaluation (M&SE) strategy. The development of an evaluation strategy should come after the Executive Management Group has determined its approach to decentralized evaluation.

7. Strategic Choices regarding the WFP Evaluation Function

70. WFP has reached a point where strategic choices have to be made for the future direction of the evaluation function. Independent evaluation by OEV is now well established, but all other aspects of the function suffer from uncertainty, especially with regard to 1) the appropriate levels and mechanisms for resourcing human and financial for evaluation, 2) the value afforded by WFP to decentralized evaluation, 3) the respective roles of OEV, RMP, country offices and regional bureaux in operations and decentralized evaluation.
71. Due to a lack of information on WFP's decentralized evaluation function, the Panel is not in a position to recommend the most appropriate model for evaluation in WFP. The BPR assessment of decentralized evaluation capacity is gathering further information and will provide options for the decentralized function.
72. The Panel offers three models for WFP's consideration, which are set out in further detail in Annex 16, and included below in summary form only. These models are not intended to be rigid but can be seen as points along a scale. The question for WFP is where it wishes to position itself between Models 1, 2 and 3 and which elements of each it wishes to adopt.

³⁸ According to the Office of Budget, there are two project codes related to evaluation 1) Evaluation/ surveys, 2) Impact/Evaluation Monitoring. (One of these lines is used for the cost recovery element of the current operations evaluations). It appears that no reports have been run against these codes. They are only used when a country office contracts a commercial firm of consultants, which for decentralized evaluation is the exception. If the country office engages individual evaluation consultants, which is more typical, this goes into payroll, as does the consultants' travel, making these costs invisible as an evaluation cost.

73. Much of the rest of this report can be used to inform the choice of model and the Panel's only firm recommendation is that a clear choice should be made and followed through. The model selected will inform the revision of the Evaluation Policy and the development of an Evaluation Strategy.

7.1. Model 1 – Centralized evaluation

74. In Model 1, WFP determines that the organization should continue to look to OEV as the source of high quality and impartial evaluation, with any decentralized evaluations seen as internal lesson learning exercises for the office concerned and to provide a measure of accountability to donors. Decentralized evaluation receives limited ad hoc support with no formal expectation that OEV or regional bureaux should provide consistent support or quality assurance. The Special Account continues to fund OEV operations evaluations beyond 2015, though at a reduced level compared with the 2013-15 series.

7.2. Model 2 - Centralized Evaluation plus Demand-led Decentralized Evaluation

75. Model 2 extends Model 1 by adding: OEV quality standards and guidance for decentralized evaluation; technical support to country offices commissioning evaluation to improve quality, to be provided by OEV and regional bureaux based on country office demand; and OEV reporting on the quality of decentralized evaluation reports. Regional bureaux generate an annual regional plan for evaluation based on country office stated intentions, which is coordinated with OEV's programme of operations evaluations (at a reduced level compared with 2015, as in Model 1). There is no organization-wide plan to resource, strengthen or systematise the decentralized evaluation function. OEV establishes a help desk and is resourced to do so, and an improved roster of evaluation consultants from which country offices can draw, in partnership with regional bureaux. Additional resources will be required in both OEV and regional bureaux, an additional two full time professionals in OEV and one professional devoted to evaluation in each regional bureau.

7.3. Model 3 – Centralized and decentralized evaluation function

76. In Model 3, WFP adopts an 8-10 year roadmap for the expansion and strengthening of the evaluation function at all levels of the organization, with regional bureaux supporting evaluations at regional and country level and supporting country level evaluation officers and evaluation focal points, with backup support from OEV. OEV has two full time staff dedicated to supporting decentralized evaluation, all regional bureaux have a full time Regional Evaluation Adviser and all country offices with the largest operations (e.g. all those projected to exceed US\$150 million per year) have dedicated evaluation officers. OEV quality assesses evaluations from all levels of the organization. Decentralized evaluations take place according to set criteria and on a planned cycle, agreed by the Regional Director, in consultation with the Director of Evaluation.

7.4. Selecting a Model

77. Models 2 and 3 recognise the added value of decentralized evaluation, while Model 1 does not. Model 1 represents more or less the status quo and can be delivered, with some impact on OEV evaluation output, within the current Programme Support and Administration (PSA) and special

account allocations to evaluation. Models 2 and 3 require additional resources. Table 1 below gives a tentative estimate of the additional staff resources required for each model.

Table 1 – Estimated additional human resources required to support possible WFP evaluation models

	Model 1	Model 2	Model 3
OEV oversight role, senior officer	0.5	0.5	1
OEV oversight role, junior officer	0.5	0.5	1
OEV support role/help desk function, senior officer		1	0.5*
OEV support role/help desk function, junior officer		1	1*
Regional Bureau, evaluation adviser**	-	0.5	1 each
Largest country offices (US\$150m plus), evaluation officer	-	-	1 each

* In Model 3, regional bureaux provide most of the support to country offices: ** assumes that the regional M&E officers will continue to have perhaps 10% of their time available for evaluation.

78. All three models have advantages and disadvantages, and a case can be made for any of the three. Model 1 is the most easily achieved because it is closest to the status quo but it ignores the reality and potential added value of decentralized evaluation, and is not in line with the ED's intent set out in 'Fit for Purpose' for the establishment of a field-based evaluation function. Model 3 represents, in many respects, an ideal model but requires investments in evaluation capacity at a level well beyond anything WFP has so far contemplated and for which there does not yet seem to be an appetite amongst senior management. However, the Panel believes that WFP can see the models as a continuum evolving from Models 1 to 3. Model 2 could be established and produce measureable improvements in terms of better quality decentralized evaluations within 2-3 years. At that point, WFP could review progress and decide whether it sees sufficient added value from a greater investment in its decentralized evaluation function, and take the next major step in the development of evaluation across the organization, moving towards Model 3. For OEV to support Model 1 would require the Office to reduce its evaluation output by 10-20% if no additional resources are provided. Models 2 and 3 cannot be delivered within the current resource framework.

79. The Panel proposes that the Executive Management Group should make a decision on which model to follow in mid-2014, after the results of the current BPR assessment of decentralized evaluation capacity is complete. This decision, in turn, will underpin WFP's response to the Peer Review, due to be considered by the Board in November 2014.

8. The Office of Evaluation

8.1. OEV Objectives and Strategy

80. The Evaluation Policy describes the role of the OEV as follows (from para 19):

1. Setting direction and providing guidance, updating the Evaluation Policy and evaluation guidance when needed;

2. Selecting and preparing work programmes³⁹ with a budget that it is 'proportionate to programme costs' (although proportionate is not defined by the policy);
3. Conducting evaluation work: designing, planning, managing and undertaking (various types of) evaluation, while ensuring quality standards for evaluation are followed;
4. Fostering a corporate culture of accountability and learning and 'developing the evaluation capacities of WFP at various levels, in particular in the field';
5. Communicating evaluation findings to support decision-making, developing a repository of evaluation lessons, and organizing the annual informal consultation on evaluation;
6. Contributing to professional evaluation associations and groups.

81. Of these six roles, OEV puts the bulk of its effort into 2 and 3, planning and conducting of evaluations. OEV also invests in the other four roles but does not have sufficient resources to undertake them fully, given the demands of the agreed programme of evaluations. OEV does not set annual objectives for the other four roles. The Director of OEV is aware that unless more resources are provided, the Office will need to cut back on the number of OEV evaluations if it is to deliver fully on its other responsibilities, especially for oversight, increasing the utilization of evaluations, and capacity development.

8.2. OEV Budget

82. The PSA budget⁴⁰ for OEV's core work programme for the review period has been as follows⁴¹:

	Staff	Non-Staff	Total	Biennium	Total per year	Variance
2014	2,387,000	3,011,000	5,398,000		5,398,000	2%
2013	2,266,000	3,011,000	5,277,000		5,277,000	8%
2012	2,170,000	2,717,000	4,887,000	10,164,000	4,887,000	2%
2011					4,796,500	0%
2010	4,086,000	5,507,000	9,593,000	9,593,000	4,796,500	43%
2009					3,350,000	0%
2008	3,719,000	2,981,000	6,700,000	6,700,000	3,350,000	

83. The OEV budget increased by 61% between 2008 and 2014, reflecting the organization's increased commitment to evaluation, while OEV's budget as a percentage of total contributions to WFP has increased from 0.07% to 0.13% in 2014^{42,43}. According to recent figures from the JIU, this is in the

³⁹ Current practice is the approval of one year and, provisionally, two forward years.

⁴⁰ Excluding project funds for a series of single operations evaluations of US\$2 million in 2013 and US\$3.7 million in 2014.

⁴¹ Source: WFP Management Plans 2009-2013.

⁴² The percentage rises to 0.22% when the special account funding of US\$3.7 million for OEV operations evaluations in 2014 is added in.

⁴³ In practice, WFP's spending on evaluation is a higher percentage of turnover than these figures suggest because decentralized evaluation expenditure is included, for which data are not available.

same range as UNICEF and UNDP, but below normative UN organizations, such as FAO and ILO, that have a much smaller turnover while maintaining evaluation expenditure at broadly similar levels to WFP in US\$ terms. From 2010-2013, the average spend on WFP country portfolio evaluations was US\$194,667, on global evaluations (policy, strategy, thematic) US\$302,527 and on impact evaluations, US\$229,367. Figures from the JIU indicate that WFP appears to be in the mid to low range for the cost of individual evaluations, compared with its UN peers (see Annex 15).

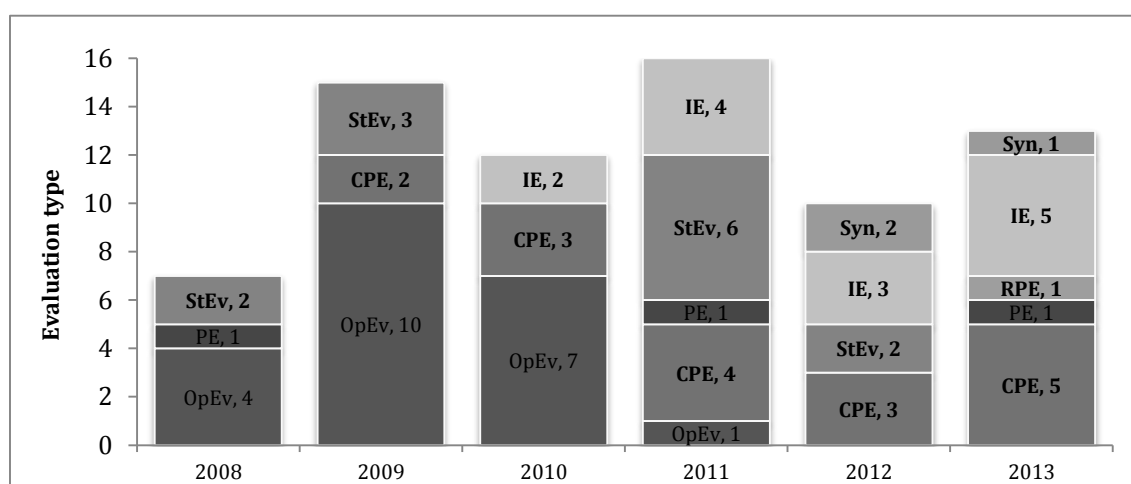
8.3. OEV Evaluation Planning and Selection

84. The UNEG standards require that the plans for evaluation be submitted to the Chief Executive and/or the Governing Body for approval. WFP is meeting this standard. The Director of OEV discusses the OEV work programme annually with Office of the ED, the DED and other directors, and also informally with the Board before its approval at the November Board session. The work programme is agreed for one year, and provisionally agreed for a further two years ahead.

8.4. OEV Evaluation Coverage

85. In terms of activity, on average, OEV is conducting a higher number of evaluations per year than most other UN central evaluation units. OEV has completed 70 evaluations, plus three evaluation syntheses since 2008, 12 per year on average. When the new operations evaluation series is included, by 2015 OEV will be amongst the highest performing evaluation units the UN by numbers of evaluations completed⁴⁴, although this does still not meet the internal demand. OEV's output in terms of evaluations completed for the period 2008-2013 is shown below⁴⁵.

Figure 1: Number of Evaluations in all Regions 2008-2013⁴⁶



⁴⁴ By 2015 total output could reasonably be anticipated to be at least 40 in the year. The latest relevant data from the JIU for evaluations completed by central evaluation units in 2011 was as follows: WHO-60, FAO-33, UN Secretariat-20, WFP-16, UNHCR-12, UNDP-11, UNICEF-6.

⁴⁵ Evaluations conducted by OEV in the period 2008-2013 are listed in Annex 7.

⁴⁶ Evaluations completed by OEV covering all Regions for 2008-2013. CPE: Country Portfolio Evaluation; RPE: Regional Portfolio Evaluation; OpEv: Operation Evaluation; IE: Impact Evaluation; Syn: Synthesis Evaluation; StEv: Strategic Evaluation of selected WFP evaluations; PE: Policy Evaluation

8.5. Evaluation Types

86. OEV anticipates that the Evaluation Policy will need to be updated to clarify terminology, and the role of OEV in standards, guidance and quality assurance. Box 1 summarises the current understanding of different evaluation types in WFP. In the majority of discussions during the review, including with regional bureaux, the need to clarify terminology was highlighted. This concerns the typology of evaluation, plus decentralized versus centralized, and other types of 'evaluative' exercise, mid-term reviews, management reviews, surveys, and assessments.

8.6. OEV Evaluation Portfolio

87. The Director OEV has considerable licence in the selection of evaluation subjects. The OEV portfolio provides a balance of defined evaluation types, all of which are well articulated and which appear to be adding value for the different types of stakeholders concerned.

88. OEV's portfolio has changed over the past year, with the addition of the series of operations evaluations and an increase in coverage of humanitarian emergency evaluations. The Panel considers that the OEV's portfolio is now better balanced as a result.

89. Strategic evaluations were influential in developing the new WFP Strategic Plan 2014-2017, while OEV has committed to an 'evaluability assessment' for the Strategic Plan, and will plan for later evaluations of elements of the plan. Overall, the Panel considers that evaluation planning is being conducted in a professional and consultative way.

90. Strategic, policy and impact evaluations are well received by the Board and senior managers, and are influential. The Panel supports their use into the future. OEV evaluation managers have some questions about the future relevance of other types of evaluation to the future shape of the organization. Some expressed doubt that evaluations of single operations will remain relevant.

91. For countries with 3-4 operations that are not on a large scale, it would be more relevant to evaluate them together, at one time, rather than evaluating individual operations, and to include WFP's strategic positioning in that country in the scope, as per country portfolio evaluations. Where there are very large operations, for example in Sudan, it makes sense to evaluate a limited set of operations. Evaluating all operations at once would be possible, albeit a major undertaking, and carry the risk of the evaluation results being too general⁴⁷, a criticism sometimes levelled at Country Portfolio Evaluations (CPEs). In any event, evaluating WFP strategic positioning in the country is relevant to all evaluations. In the Panel's view, regional bureaux should be able to judge whether individual operations evaluations or a CPE are the most appropriate for each country, in discussion with the relevant country office.

92. CPEs are presented to the Board, while the new series of single operations evaluations will not be⁴⁸. OEV reports that this can make country directors less enthusiastic about CPEs than operations evaluations. However, the Panel's view is that CPEs remain an important part of OEV's

⁴⁷ Feedback from one region indicated that two CPEs in the region were out of date by the time they were complete and asked that the scope of such evaluations bring the timescale up to the date of the field mission, not to a historic end date.

⁴⁸ Any future evaluations of Level 3 emergency operations will be presented to the Board.

portfolio. If CPEs can meet field staff expectations that they be more timely and 'field oriented' (see 8.3 below), their perceived relevance to country teams should improve. CPE guidance has recently been reviewed. These evaluations should be monitored to see if they are providing up to the moment results and sufficiently specific recommendations for the elements of the WFP operations and programme.

8.7. Evaluation Criteria

93. Evaluations of different types are selected for inclusion in the OEV Work Plan using certain criteria, as per the Evaluation Policy. The Panel considers that these criteria need further development. Annex 10 includes a possible set for further discussion within OEV. A revised policy should cover evaluation selection based on:

- ❖ Corporate strategic significance (including, for example, Level 3 emergencies);
- ❖ Size and complexity;
- ❖ Risk (e.g. financial, environmental, political, non-performance, reputational, etc.);
- ❖ Opportunity (to demonstrate impact, to provide particularly insights, or to evaluate jointly with other partners);
- ❖ Geographic and thematic coverage.

Box 1: Evaluation Types in WFP

According to the evaluation pages on the WFP web site, OEV currently conducts five types of evaluations, plus the Annual Evaluation Report submitted to the Board, as follows¹:

1. *Strategic Evaluations*: take a global perspective and focus on strategic issues for WFP, especially those highlighted in the strategic plan.
2. *Policy Evaluations*: examine how effective WFP policies have been in achieving their objectives and propose amendments.
3. *Country Portfolio Evaluations (CPE)*: help country offices understand the relevance of their portfolio to the national context and help offices make choices when it comes to CP revision.
4. *Impact Evaluations*: provide an in-depth analysis of the impact that WFP work has on beneficiaries.
5. *Operations Evaluations*: tell the story of relevance, efficiency and effectiveness of an operation.
6. *Annual Evaluation Report*: synthesises evaluation findings during the year.

The evaluation typology in the evaluation policy no longer matches the WFP evaluation web page description, which is in line with current OEV practice and the policy needs to be adapted to match evaluation types listed above.

According to the 2008 evaluation policy, **“Decentralized evaluations are those of single operations that are managed by regional bureaux or country offices. These evaluations follow the same standards as those managed by the Office of Evaluation, including the recruitment of external consultants for the task.”** This definition holds good. In one sense, any evaluation not undertaken by OEV is ‘decentralized’, because it is being undertaken by WFP management, even if the commissioners are in HQ. As per the definition above, Operations Evaluations are evaluations of single operations. Where these are conducted by country offices or regional bureaux, they are ‘decentralized operations evaluations’. In effect, ‘decentralized’ means management-led.

According to the evaluation policy, **“Self-evaluations are undertaken by WFP staff in country offices without involvement of consultants other than for the facilitation of the exercise, if necessary. They report on the implementation of the annual work plan, providing a comparison of planned versus actual results. Given that they are self-evaluations, they do not have to meet the requirements for independence spelled out in this Policy”.** Self-evaluation (also called auto-evaluation) is recognized by the OECD-DAC Glossary of Key Terms in Evaluation and Results Based Management as “An evaluation by those who are entrusted with the design and delivery of a development intervention”. However, the term is causing confusion in WFP and consideration should be given to using ‘management review, or simply ‘review’, instead.

The table below shows some differences between different types of ‘evaluative’ exercise (see Annex 12 for further discussion).

	Led by Management	Independent?	Using External Consultants?	Fit with UNEG Standards?	Submitted to the Board?
Review	Yes	No	If useful/affordable but not required	No	No
Decentralized evaluation	Yes	Partially	Yes	Yes*	No
OEV evaluation (centralized)	No	Yes	Yes	Yes	Yes**

* but not necessarily meeting the full requirements of the WFP Evaluation Quality Assurance System (EQAS):

** single operations evaluations by OEV are not presented to the Board individually but only as a synthesis

94. In addition: It should be feasible to extract sound findings, recommendations and lessons learned; the evaluation should be timely in terms of reporting results or informing programme design; there should be sufficient resources to conduct a high-quality evaluation within the time period available; and there should be demand for the evaluation (from e.g. the Executive Board, senior management, or donors) including mandatory evaluations driven by policy commitments

8.8. Joint Evaluations

95. In the review period, WFP has taken part in the following joint evaluations:

- ❖ A series of impact evaluations with UNHCR on Protracted Refugee crises;
- ❖ Joint Thematic Evaluation of FAO and WFP Support to Information Systems for Food Security;
- ❖ Joint Evaluation of the Global Logistics Cluster, with the Netherlands and UNICEF;
- ❖ Joint Evaluation of the Food Security Cluster (on-going).

96. The demand for joint evaluation with UN agencies or with government was raised surprisingly little during interviews for the Peer Review. As a result of the UN's Quadrennial Comprehensive Policy Review, WFP should anticipate greater internal and external demand for joint evaluation, especially at country level, including for UNDAF evaluations and evaluations of joint programmes. Expectations are high for all UN operational agencies to be working more closely together in partnership, which will then lead to requests for evaluation of these partnerships.

8.9. OEV Staffing

97. As of January 2014, the Office has 10 professional staff (including the Director (D2), 4 P5s, and 5 P4s), plus one temporary P2 position (not shown). OEV employs junior consultants as evaluation analysts on temporary contracts funded from the non-staff budget. In addition, there are 3 general service staff.

Core OEV Work Programme	WFP staff on rotation	Externally recruited evaluation specialists	Locally Recruited	Total
Director (D2)		1		1
Senior Evaluation Officers (P5)	1	2.5		3.5
Evaluation Officers (P4)	3	1		4
General Service Staff (G6 & G5)			3	3
Operations evaluations series (temporary)				
Senior Evaluation Officers (P5)		0.5		0.5
Evaluation Officers (P4)	1			1
Total	5	5	3	13

98. The table does not include the locally-recruited consultant research analysts that provide considerable extra resource in support of evaluations. (They also do not appear on the OEV organigram). Given the pool of talent these researchers represent, OEV could well employ them for new tasks, such as within a 'help desk', should a decision be taken to establish one.

99. OEV is complying with the evaluation policy by maintaining a balance in its professional staffing between specialist evaluators and programme staff on reassignment from within the organization. The specialist posts are 'non-rotational'. Their contracts are time limited, with a two-year initial period and renewable on an annual basis thereafter, with 2-year renewable contracts after 5 years' service. The contractual rights of staff on rotation are more stable. Assignment to OEV is

for at least 4 years renewable annually thereafter in the corporate Re-Assignment exercise up to 4 times. All OEV staff have the right to apply to move elsewhere in the organization.

100. The 2007 Peer Review found the mix of specialist and programme staff on rotation to be 'appropriate'. OEV continues to see advantages in the current arrangement and the Panel sees no reason for change, given the balance of skills and experience it provides within the Office. All non-specialist staff attend the IPDET⁴⁹ training, or a shortened European version, on joining OEV. The Panel considers that helping generalist staff to learn evaluation management skills as quickly as possible, with on-going peer support and coaching from OEV colleagues, is a priority.

8.10. OEV Team Management

101. Overall the OEV team spirit is generally good and the Director and Deputy Head are respected within the Office. Having received additional funding in the 2013 budget round, which was sustained in 2014, OEV feels somewhat under pressure to perform and the Director and staff bear heavy workloads.

102. Team roles and management of the office were discussed as part of the OEV Internal Review. The team has a flat structure, with most of the professional staff reporting to the Director, while the evaluation analysts report to the Deputy Head⁵⁰. The Director is directly involved in the quality assurance process and in the clearance of individual OEV evaluation reports.

103. In the Panel's view, the Director should be released from detailed attention to individual evaluation reports so that more of her time can be spent on keeping an overview of the evaluation function across the organization and on making use of evaluation to influence performance improvement in the organization, and on internal and external partnerships. The structure of the Office could usefully be revised to become more vertical by giving more responsibility to P5 officers for particular areas of OEV business, for example, evaluation quality assurance, or the utilization of evaluation knowledge, which would be motivating for these officers and reduce the management load on the Director. (This process has already started).

104. The Panel considers that OEV has a strong case for the establishment of a Deputy Director post at D1 level, to be recruited by open competition. IFAD and UNDP have D1 Deputy Directors of evaluation⁵¹. These units are slightly larger than WFP's but the expectations on OEV are high to maintain quality and relevance and the Director needs a deputy who can shoulder a good proportion of the OEV internal management load, especially if the Director is given responsibility for oversight of the evaluation function across WFP.

105. Evaluation managers are heavily involved in all stages of the evaluation process apart from data collection and analysis, even when using LTA companies⁵².

⁴⁹ IPDET: International Program for Development Evaluation Training

⁵⁰ An OEV proposal to have the deputy director role re-classified as D1 was not approved.

⁵¹ UNICEF, UNFPA, FAO, and ILO do not have D1 deputies.

⁵² While the evaluation policy allows that OEV staff may be part of evaluation teams, in practice, for the past few years, OEV staff have not participated as evaluation team members, and the Director does not favour them doing so.

8.11. Vesting Independence in the Office of Evaluation

106. Most UN evaluation offices stake the independence of the evaluation on the assumed independence of the external evaluation team that has undertaken the evaluation. The evaluation units of IFAD, OIOS, and UNDP, by contrast, take full ownership of the evaluation reports. OEV takes partial responsibility by creating the summary evaluation report using the consultant's report as a basis, and seeks the consultants' agreement to the text of the summary. The final wording of the recommendations, however, is left to the consultants.
107. The tendency in the international financial institutions and in the more developed evaluation functions in the OECD-DAC is for evaluation units to take full ownership of their evaluation results. The Panel proposes that WFP adopt this approach. This will put WFP out of step with the majority of UN agencies but represents a logical step in the maturing of the evaluation function. The Director of Evaluation should take full ownership of the evaluation report.
108. This has important implications because it resolves the question of whether evaluation managers should be directly involved in the conduct of evaluation. The independence of the evaluation process and eventual report becomes the responsibility of the manager, acting on behalf of the Director, rather than that of the consultants, who are engaged to assist the evaluation manager.
109. Points in favour of the Director OEV taking full ownership of the report include:
- ❖ Having successfully established a high degree of independence, which is respected in WFP, the idea that contracted consultants are more independent than the Director OEV is unhelpful and confusing;
 - ❖ As noted in Annex 12, consultants are not truly independent once under contract to the organization;
 - ❖ OEV should be able to conclude evaluations more quickly because the Director will be able to make judgement calls on the final content of evaluations with advice from the evaluation manager, still in consultation with the consultants, but perhaps with fewer rounds of consultation and revision;
 - ❖ It will bring WFP in line with more mature independent evaluation offices of the International Financial Institution (IFIs) and bilaterals, and a few UN organizations.
110. However, there is a genuine concern amongst OEV staff on rotation that evaluation managers could be pressurised by management in offices under evaluation to modify evaluation results (for example, pressure on the evaluation managers can be exerted through perceived or actual threats to limit their future job prospects in WFP). The Director of OEV should make it clear to the evaluation managers that if they are pressured to change an evaluation report based on possible unfavourable treatment once they progress from OEV, this should be brought immediately to the attention of the Director, and, if necessary, to the Ethics Office.

8.12. Appointment of the Director Office of Evaluation

111. The procedure for the appointment of the Director of Evaluation specifies that a member of the Board should be part of the selection process, in line with good practice in the UN system. In the appointment process for the current Director, one Board member was included, which is the minimum appropriate level.

112. The Director of OEV is appointed for a maximum of eight years, in two four-year terms, which is in line with comparator UN agencies. IFAD appoints for one term of six years, and this is the pattern in the most international financial institutions. In the Panel's view, a one-term appointment is better because it leaves no room for the independence of the Director to be compromised during their first term by the desire to be reappointed for the second term.

113. The circular on the appointment of the Director of Evaluation⁵³ contains an anomaly in that it allows for the Director's re-entry to WFP⁵⁴. This is out of step with good practice in the UN system, whereby there is no possibility for the Director of Evaluation to return to the organization, so as to protect the Director and the organization from any conflict of interest and to preserve the independence of the function. The Panel recommends that the relevant circular be amended to remove the Director's right of return to WFP, in line with the parallel circular for the appointment of the WFP Inspector General.

8.13. Prioritising OEV Resources

114. The challenge for OEV is how to reconcile the different and growing needs to:

- ❖ Conduct evaluations;
- ❖ Communicate evaluation results and contribute to learning;
- ❖ Oversee and report on the state of the evaluation function across the organization;
- ❖ Strengthen the evaluation function, both centralized and decentralized;
- ❖ Contribute to inter-agency and system-wide evaluation and standard setting, evaluation networks and conferences.

115. To some extent, OEV is at risk of being overtaken by its own success. WFP has an increasing need for evaluation that the organization looks to OEV to meet because of its mandate and its reputation. As a result, the demand for evaluation and evaluation advice outstrips OEV's capacity by a considerable margin and the Office is working at a rate the Director considers unsustainable. The Panel formed the same impression. OEV needs to prioritise its resources clearly but has been limited in its ability to do so because of the organization's lack of clarity about the future of operations and decentralized evaluations.

⁵³ OED 2010/004.

⁵⁴ With regard to the completion of the Director's term, the circular states "There shall be no possibility of re-employment within WFP *to another post or on a consultancy for at least one year*" (italics added).

116. Considerations include:

- ❖ OEV is keen to put greater effort into stimulating learning and use of evaluation. OEV has stepped up its production of syntheses of OEV evaluations and would like to create more policy-oriented products.
- ❖ The Board has an appetite for more evaluations, but some Board members are concerned that OEV does not have enough staff for its current work programme. OEV needs to manage the Board's expectations for evaluation output as it puts more effort into other functions.
- ❖ Survey feedback and interviews confirm that country office staff expect support from OEV. OEV needs to clarify to what extent it will respond to these demands.
- ❖ OEV management, senior management and the Board have yet to address how much of OEV's effort should go into functions other than conducting evaluations, and WFP has choices to make concerning OEV's future role.

117. The Panel proposes that:

- ❖ OEV should take responsibility for standard setting and oversight of the evaluation function across the whole organization, including the monitoring and assessment of the quality of decentralized evaluations;
- ❖ OEV should continue to put the bulk of its effort into the conduct of independent, centralized evaluation while taking on the new oversight role;
- ❖ If the Executive Management Group decides to strengthen the decentralized evaluation function, OEV will need to give more attention to; guidance, training, 'help desk' advisory support, and to backstopping the regional bureaux in their oversight of decentralized evaluation, while RMP provides support and guidance on reviews, assessment and other management learning exercises.

118. If resources remain constant, there will be a need to consider trade-offs between these roles and to re-prioritise.

9. Evaluation Quality

9.1. The Evaluation Quality Assurance System (EQAS)

119. Since the 2007 Peer Review, OEV has placed increasing emphasis on analytical depth and transparency in its evaluations, with more stakeholder consultation and sharper reporting, including better formulated recommendations. In support of this goal, from 2007, WFP developed a detailed Evaluation Quality Assurance System (EQAS), which has been regularly reviewed and reinforced. EQAS provides a comprehensive step-by-step guide to the evaluation process, from conception to utilization. The development of EQAS puts WFP at the forefront of the systematization of evaluation processes in the UN and, in WFP's case, it has the extra benefit in giving generalist WFP staff on rotation coming into OEV a ready-made and detailed roadmap for

conducting evaluations. However, OEV staff expressed concern that EQAS may have become too heavy and complicated, a concern shared by the Panel, which proposes a simplification of EQAS.

9.2. Quality Assessment of WFP Evaluations

120. In 2012, CIDA led a review of WFP effectiveness⁵⁵, which included a quality assessment of 52 OEV evaluations undertaken in the period 2006-2011. The review rated the quality of WFP evaluation reports highly overall. No evaluation scored less than 26/45 in the CIDA-led study.

121. This Peer Review has also conducted a Quality Assessment of WFP evaluations using a sample of 20 evaluations (7 decentralized and 13 OEV evaluations). Evaluation reports were rated against a set of weighted criteria agreed by the Panel and on which OEV was consulted. At the request of OEV, a hybrid approach to assessing the reports was adopted. In addition to a review of all documents in the sample, 5 of the 13 OEV evaluations were further analysed from a longitudinal perspective. A summary assessment is included as Annex 4 and a full report of the quality assessment is available as a separate report.

122. In every category of analysis, centralized evaluations scored higher than decentralized evaluations, although decentralized evaluations also scored well. The overall quality gap between OEV and decentralized evaluations was 17%. OEV evaluations were of consistently high quality (74%-91%, average 81%), while decentralized evaluation scores were more varied (42%-84%, average 64%). This variety is to be expected, given the diversity of country offices managing the decentralized evaluations without any consistent evaluation training programme or quality assessment.

123. The document review showed no significant change in evaluation report quality over time from 2009-2012, which did not marry with the impression gained from interviews and survey results, where the feedback pointed to gradually improving standards. The table below shows a summary of the document review ratings by criteria.

Evaluation Criteria	Weighting 1-10; 1=least important, 10= most important	Average Score for Centralized Evaluations	Average Score for Decentralized Evaluations
Recommendations	10	80%	46%
Findings and Conclusions	8	84%	71%
Analysis	8	83%	62%
Evaluation Methodology	6	82%	69%
Executive Summary	6	76%	55%
Evaluation Context	4	86%	73%
Evaluation Purpose and Scope	4	83%	71%
Accessibility/Clarity	3	86%	79%
General Information	1	88%	83%
TOTAL AVERAGE (Weighted)		81%	64%

124. The longitudinal analysis of five OEV evaluations showed that they played a central role in shaping the priorities and strategic direction of WFP at country or global level. All of the

⁵⁵ Review of the World Food Programme's Humanitarian and Development Effectiveness 2006-2011, April 2012.

evaluations reviewed were mentioned in follow-up strategies and policies and their findings influenced management's decisions. This indicates both the value of OEV products and WFP's ability to use and learn from them, especially at policy/strategy level.

9.3. Feedback from the field

125. Survey respondents agreed that OEV uses skilled evaluators that OEV evaluations include high quality analysis, and that evaluations are conducted independently. Responses were positive overall but more mixed on whether OEV evaluations include recommendations that can be readily implemented, and several respondents disagreed that recommendations are well used in improving performance, and on whether there is effective communication by OEV of lessons learned from evaluations. All respondents agreed that 'the quality of OEV evaluations has improved significantly since 2007', and that 'the OEV evaluation agenda includes the most important evaluation topics'.
126. Country Offices interviewed (4) reported that their engagement with OEV evaluations was productive (variously impact evaluation, a case study in a global level evaluation, or CPE). In all cases, they feel that they were well consulted by OEV and have been able to make good use of the evaluation results in programme redesign and with partners.
127. Responses from the staff survey carried a strong message that country offices want to see OEV evaluations being more field-oriented and practical, especially in their recommendations. Some saw the current approach as too 'corporate', and without a proper consideration of the country context. (Salient quotes are included in a footnote⁵⁶ and Annex 7 includes a summary of the staff survey results).

9.4. External Reviewers and Reference Groups

128. From 2011, OEV expanded its use of external reviewers for those evaluations with especially high levels of stakeholder interest with diverse perspectives. These review panels are separate from the independent consultants who conduct evaluations, and provide an additional dimension for the quality assurance of methodology and/or content.
129. Every OEV evaluation has a reference group made up of internal stakeholders that provides advice and feedback during the evaluation process. These reference groups have an advisory role only. Group members learn about the process of managing evaluation and how to engage

⁵⁶ • "Evaluations need to be more considerate on realities on the ground in country offices rather than coming solely with a corporate approach and measuring country office performance against corporate objectives.

- Evaluations are considered as being "perfect" and in the instance of the private sector evaluation were not completely reflective of the situation or environment. Considerable effort is then required to change views and perceptions that are the result of evaluations. Recommendations are not really viewed as "recommendations" - but rather mandates.

- Less concern about standards/independence and more concern about actionable findings which can be put to practical operational purpose in-country. Less upstream focus, more focus on supporting countries during the conduct of evaluation and ensuring the recommendations are implemented (i.e. much more downstream focus is required).

- ... the views of the country office should be given more weight in the process to ensure its views are better taken on board, particularly when it is felt that the evaluators are driving their own agenda rather than balancing their view with the realities on the ground.

- For country programme and impact evaluation, it needs to be translated into realistic programmatic recommendation and in a way that help the country office to prepare the next programme."

stakeholders in evaluations and are potential future champions of evaluation and the use of evaluation evidence. The use of reference groups is said to have improved evaluation quality, which mirrors experience elsewhere in the UN. The Panel supports the use of reference groups as a means of improving evaluation quality, engagement and ownership.

9.5. Long Term Agreements with Evaluation Service Providers

130. As a key step in managing workload and improving quality, OEV has formed agreements with ten consultancy companies to provide evaluation services under Long Term Agreements, through which it conducts most, but not all, of its evaluations⁵⁷. According to OEV, the LTAs provide access to higher quality and a more diverse range of consultants than when hiring individual consultants⁵⁸, and outsourcing the location, hiring and management of consultancy teams accelerates the process of launching evaluations because the normal two-three month procurement process can be avoided. While there are gains at the start, concerns were expressed by both OEV and the LTA companies about the length of time taken to finish evaluations, especially the phase from zero draft to completion⁵⁹. There is a good deal of frustration in OEV about the low quality of draft reports received from some evaluation teams. LTA companies, for their part, find the quality of OEV evaluation management variable.

131. WFP has gone further than other UN organizations in its outsourcing model and there is much for the UN system to learn from WFP's experience. The Panel supports OEV's intent to evaluate the LTA approach in due course. (See Annex 9 for more detail on LTAs).

9.6. Potential areas for quality improvement

132. The Panel proposes that OEV should consider the following areas for their potential for evaluation quality improvement:

- ❖ **Upstream/Downstream.** The Panel supports OEV's move to extend its evaluation processes 'backward and forward', where 'backward' means upstream work on evaluability, theory of change, logic models and engaging stakeholders more fully in the inception of evaluation, while 'downstream' involves increased use of workshops and other interaction with stakeholders in the draft and final report stages.
- ❖ **Humanitarian Standards.** With regard specifically to emergency programming, there are de facto international standards by WFP's main UN and NGO partners against which emergency operations can be evaluated. The most important of these is the Sphere Standards for Humanitarian Response⁶⁰, which are themselves based on humanitarian principles.

⁵⁷ OEV has recently completed an 'Orientation Guide for Evaluation Companies: Key facts about WFP and its Operations', November 2013.

⁵⁸ In 2011 and 2012, three quarters of the evaluation consultants hired under LTAs had not worked for OEV before.

⁵⁹ Some OEV staff expressed concern about the quality of drafts being submitted by LTA firms and there appears to be heavy engagement with numerous drafts being exchanged between the OEV manager and the team leader following the receipt of the first draft. OEV may select one LTA company for an evaluation or run a competition between a group of LTA companies for the evaluation.

⁶⁰ <http://www.sphereproject.org/handbook/>

- ❖ **Partnership.** WFP will need to place greater emphasis on evaluating its activities within the context of government-led structures and initiatives. EQAS already includes guidance on evaluating the integration of WFP strategy into government-owned structures, and its contribution to national processes and structures, including coordination mechanisms. WFP’s behaviour as a collaborative partner will, like other agencies, come increasingly under the spotlight and more specific criteria for the evaluation of partnership behaviours and outcomes may be required.
- ❖ **Beneficiaries.** One of the new Organizational Design principles is “Beneficiaries are the focus of all our work”⁶¹. Both WFP’s EQAS and the UNEG standards are deficient on the need for, and means of, consultation with beneficiaries, and on beneficiary led evaluation. WFP continues to be a leader in the development of the IASC’s commitments on Accountability to Affected People⁶², which could be used to inform WFP’s evaluation guidance.

10. Management Response to WFP Evaluations

10.1. Management Response arrangements

133. Following the 2007 peer review, WFP made an important innovation by moving the responsibility for management responses to evaluation away from OEV and into line management⁶³. This has been developed to the point where WFP now has a more systematic approach to management responses than most other UN organizations.

134. RMP is responsible for compiling the management response for each OEV evaluation, in consultation with the relevant managers, which is then cleared by the office of the DED, or the ED, depending on the type of evaluation. The management response is then submitted as a separate Board paper along with the evaluation report, and is introduced to the Board by the Director of RMP after the Director of OEV has summarised the evaluation findings⁶⁴. The relevant manager is on hand at the Board meeting, if available, either in person or remotely, to answer any queries from the Board, but does not actually present the response.

135. More than one director raised the question with the Panel as to whether it is appropriate for the Director of RMP to introduce the management response, rather than the manager most responsible for implementation. The Panel considers that it would be better to have a designated manager with the greatest stake in the implementation of the management response (or their Executive Management Group supervisor) lead the presentation to the Board because it is with the relevant manager that the accountability for past and future results lies. It should also lead to

⁶¹ Fit for Purpose, page 2.

⁶² WFP co-chairs the relevant task force in the IASC. See also <http://www.humanitarianinfo.org/iasc/pageloader.aspx?page=content-subsidi-common-default&sb=89>

⁶³ This in line with UNEG Standard 1.4 requires that “Appropriate evaluation follow-up mechanisms should exist within the organization, ensuring that evaluation recommendations are properly utilized and implemented in a timely fashion and that evaluation findings are linked to future activities.”

⁶⁴ There is a procedural shortcoming in the process in that the Board’s comments during the formal Board session on the management response are not taken into account in the management response i.e. it is not revised after the Board discussion. This will not always be necessary but there is no step in the procedure that requires consideration of how or whether the response should be updated.

a more informed debate with the Board. (See Annex 11 for more discussion on the management response system).

10.2. Quality of Management Response

136. There are a number of challenges with the management response process. The quality of responses is uneven, with a majority of clear, definitive response statements to recommendations, mixed with other responses that are either unclear or formulaic in making statements of intent without clarity on what action will be taken when. The quality of the response necessarily suffers when the recommendations themselves are of low quality, for example, where they are unclear or take no account of the resources required to implement them. Management accepts that there are some shortcomings but also commented on the imperative to have a 'good enough' response by the deadline for Board submission.

137. The Panel considers that management responses could be improved if there was more than two weeks⁶⁵ from the finalisation of the report to provide a response, unless all the recommendations are very straightforward and simply dealt with. Four weeks would be more appropriate especially for strategy, policy and thematic evaluations, or where more extended deliberation between various divisions and offices is required.

10.3. Implementation Follow-Up

138. During the review, many staff and Board members commented on the need to improve the accountability for the implementation of evaluation recommendations, but there is something of a contradiction in the review findings on this point. On one hand, RMP dedicates considerable energy to following up how and whether managers have taken the action they committed themselves to in the management response, sometimes years into the past, and to reporting on implementation rates. RMP's follow-up process must be one of the most methodical in the UN system⁶⁶. On the other hand, evaluation follow-up was the most frequent topic in the staff and Board survey feedback. (See feedback comments in footnote)⁶⁷.

139. The challenge is to know what actually happened as a result of the evaluation, rather than just knowing an overall percentage implementation rate which, while useful in its own right, does not provide clarity on what action was taken and whether it led to the changes the management response intended. While this may be frustrating for the Board and other stakeholders, the Panel

⁶⁵ In practice, management has more than two weeks to review the evaluation because of its engagement in reference groups and workshops before the evaluation is finalized. However, the managers responsible are not formally asked to respond to the recommendations until the report has been finalized, as the recommendations can and do change up until the final version of the report is released.

⁶⁶ Over the past year, RMP has been undertaking an intensive exercise to close outstanding recommendations.

⁶⁷ Staff survey: "The area of weakness is how the recommendations are responded to and the accountability systems that need to be put in place to measure, monitor and report on implementation and the effectiveness (or otherwise) of implementation". One respondent wanted this extended to decentralized evaluations; "Utilization of decentralized evaluations needs to be standardized/ institutionalized throughout the country offices, following the successful system that is used for OED-led evaluations i.e. use of response matrix and follow ups".

Board member survey feedback: "it (the implementation report) does not provide any details on impact/results". "Weakness: follow up of the implementation of the recommendations", "WFP management does not always implement/follow-up the recommendations made by the Office of Evaluation in a transparent and comprehensive manner", "Enhance follow up", "More efforts should be made by the EB to follow up on management response to recommendations".

proposes that the organization takes a forward looking approach to focus on ensuring that evaluation findings and recommendations are reflected in future strategy, policy and projects/programmes, rather than looking back to the implementation of management response.

140. The Panel suggests that WFP consider following the example of the JIU by making the follow up tracking system open to internal and external stakeholders, including Board members. It would also be helpful for OEV to undertake a review of the impact of OEV evaluations periodically, selectively targeting one theme or geographical region at a time, with a summary of findings included in the AER.

11. Knowledge Management

11.1. No home in WFP for Knowledge Management

141. Since 2007, WFP has made some progress towards the UNEG Standard to “Promote a culture that values evaluation as a basis for learning”. At the same time, there is a consensus in WFP that organizational learning systems are inadequate⁶⁸ and, while knowledge management is much discussed, there is no dedicated knowledge management (KM) function in the organization. Staff at all levels are concerned that this represents a serious organizational weakness. This limits the uptake of learning, including from evaluations. WFP evaluations would be of still greater value if there was an active KM function into which their results could be fed. For now, KM is placed with Human Resources and WFP has yet to resolve which Division should be driving it in future.

11.2. Collation of evaluation lessons

142. Some organizations have set up databases of lessons learned from evaluations. WFP has not yet done so in recent years. The Emergency Preparedness Division has recently developed a database of lessons learned from emergencies and RMP has a database of details of follow up actions in response to OEV evaluations, but has yet to be ‘mined’ for lessons learned. The Panel does not see great value in OEV pursuing the development of an evaluation lessons database in the medium term⁶⁹. Experience has shown that such databases are little used in the absence of an active knowledge sharing function.

11.3. Publication of Evaluations

143. In addition to the full evaluation report, OEV requires a summary report from the evaluation team with a word limit of 5,000 words, or approximately 15 pages. These reports are then edited, translated and placed on the WFP public web site. Only this summary report and the management response are submitted to the Board, although Board members also have access to the full reports online. OEV also generates a two-page Evaluation Brief. The TOR, Full Report, Summary Report

⁶⁸ A survey comment captures the general sentiment; *“We are missing the boat because we do not use a systematic approach to generate, use and apply lessons learned [from evaluation].”*

⁶⁹ The review TOR called for an assessment of how the results from WFP evaluations, joint evaluations and system wide evaluations results are communicated and used by member countries, donors, cooperating partners and other stakeholders. This aspect has been only partially covered in considering the impact of evaluations via the longitudinal analysis conducted under the quality assessment of evaluation reports.

and Management Response are all made publically available on the WFP web site, while all the relevant documentation for the evaluation is assembled by OEV on the WFP Intranet, WFP-Go⁷⁰.

11.4. Synthesis and dissemination of evaluation results

144. The 2009 OEV report 'Closing the Learning Loop'⁷¹ set OEV on a course of a more concerted effort to improve learning from evaluations. Several of its recommendations have been implemented⁷². To achieve a wider communication of evaluation results, OEV has invested heavily in the development of a range of products including:

- ❖ Synthesis of series of evaluations;
- ❖ Tailor-made inputs to the Strategic Review Committee;
- ❖ Tailor-made inputs to Policy Committee and Programme Review Committee;
- ❖ Top 10 lessons
- ❖ Tailor made briefs or workshops;
- ❖ Evaluation Country Synthesis;
- ❖ Annual Evaluation Report (AER).

145. The online surveys asked how often these various products were referred to. The Annual Evaluation Report and the Top Ten Lessons were reported as the products most often referred to^{73,74}. Within its limited resources, OEV has had success in communicating evaluation results. The AER in particular provides a strong, widely-read platform for the dissemination of evaluation results and freedom for the Director OEV to make recommendations to the organization based on them.

12. Applying the Learning from WFP Evaluations

12.1. OEV Knowledge Management Strategy

146. The latest iteration of the OEV document 'Closing the Learning Loop', sets out OEV's proposed strategy for knowledge management and learning for 2013-2014. The strategy includes two main strands, 'Improved Knowledge Management in Evaluation practice' and 'Enhanced contribution of evaluation to learning in WFP'.

⁷⁰ The peer review did not include any review of web or intranet access to products or their use.

⁷¹ <http://www.wfp.org/content/closing-learning-loop-harvesting-lessons-evaluations-report-phase-1>

⁷² The various learning products are summarized in OEV's 'Evaluation into Use: How OE Stimulates Learning for Programme Improvement'.

⁷³ These results should be interpreted carefully. Some products are targeted narrowly to a specific audience. It may also be that there is a survey bias in that respondents were relatively senior on average, and therefore perhaps less likely to refer to OEV products than WFP staff as whole.

⁷⁴ In the open-ended responses, some survey respondents indicated that OEV products are not reaching them.

147. The Panel has focused its attention on the second theme, the contribution of evaluation, as an element of utilization, because of the potential for WFP to increase the impact of evaluation on the organization's performance. To realise this potential, OEV needs to analyse where evaluation knowledge from its own, and other, evaluations can be applied. The diagram below shows some of the 'contact points' where OEV already applies evaluation learning, or could apply it more fully in the future.

12.2. Application of evaluations at country level

148. Based on interviews with current and former country directors and from the staff survey, it appears that, at country level, OEV evaluations were seen as good quality and the results were used directly in revising programmes, and in presenting evidence to government and other partners. Immediate benefits were reported from engagement of government in presentations of evaluation results. OEV impact evaluations have sometimes provided governments with evidence of a type and quality not available from other sources.

149. The engagement of stakeholders throughout the evaluation process, from TOR design to participation in reference groups, to early debate of the emerging recommendations, is itself a form of learning, and, for those directly involved, may be more powerful in terms of changed thinking and action than any written evaluation report, especially in WFP's action-oriented environment. OEV has been deepening stakeholder engagement in evaluation processes and the Panel encourages it to go still further in this direction.



12.3. Evaluation Training

150. There is potential for WFP training courses to include a module on the basics of evaluation to increase the general understanding and utilization of evaluation, in particular 1) management training and 2) induction of new staff, covering for example; what evaluation is, its benefits, and how evaluation is commissioned, supported and used.

12.4. Application of Evaluation in WFP strategy and policy development

151. There is ample evidence that at policy and strategy levels, evaluation findings are well used. Judicious selection by OEV of relevant evaluation topics and delivery of good quality products has allowed OEV evaluations to be very influential⁷⁵. The series of strategic evaluations on Food Aid to Food Assistance was the first OEV series of inter-related strategic evaluations, focusing on strategic decision-making and systemic issues. The evaluation has been influential in the drawing up of WFP's new strategic plan, in which the synthesis report is quoted several times⁷⁶. According to OEV, the evaluations of food aid to food assistance represented a watershed in senior management interest in OEV evaluations. Until then senior managers had not generally been commenting on evaluations, as they are routinely asked to. With the series on food aid to food assistance, management attention moved to a new level.

152. WFP is obliged to evaluate all corporate policies within eight years of their agreement. WFP is demonstrating good practice in the direct linkages it makes between evaluation and strategy and policy. In the Compendium of WFP Policies relating to the Strategic Plan⁷⁷, evaluation appears as the main tool by which a policy will be or has been assessed. Evaluations and the role of evaluation are referred to consistently throughout the document, which, reportedly, was not the case in the previous edition.

12.5. Application of Evaluation in Country Strategy and Project Review

153. OEV attends the Strategic Review Committee (SRC), which reviews country strategies⁷⁸, but has not participated in the Project Review Committee (PRC) in recent years. The PRC represents the last step in the project approval process, when design issues have already been settled⁷⁹. On occasion, OEV makes inputs to projects at the design stage drawing from specific evaluations, such as the recent evaluations of the Congo PRRO and EMOP for South Sudan.

154. In the Panel's view, there is a major opportunity for OEV to improve both the uptake of evaluation lessons and the demand for evaluation by engaging at the project design stage, before

⁷⁵ Survey respondents were asked to highlight those evaluations they considered most influential (without a list to choose from). The evaluation most frequently cited as influential was the series of four strategic evaluations on Food Aid to Food Assistance, highlighted by Board members, senior management, regional and country based staff, and OEV itself.

⁷⁶ Other evaluations cited multiple times by survey respondents were: Impact evaluations of school feeding; WFP's role in ending long term hunger; Strategic Evaluation of WFP's Role in Social Protection and Safety Nets; the Annual Evaluation Report; Contribution of Food Aid in Protracted refugee situations.

⁷⁷ WFP/EB.2/2012/4-B.

⁷⁸ There is some question as to whether Country Strategies will continue to be used in future in WFP. If the system is discontinued, OEV will lose an important entry point for evaluating programmes, as country strategies provide the starting place for programme logic on which evaluation depends.

⁷⁹ Currently there are 3-4 PRC virtual meetings per week, the majority of which concern budget revisions, which are of little relevance to OEV.

the PRC takes place. To be able to make an impact, OEV will need to be very selective in targeting the largest or strategically most important projects and programmes, with a view to informing their evaluability, to verify that evaluation learning has been used in their design, and to see if there is a plan for evaluation built into the project.

155. Information is not currently available on which projects are being designed and when, so OEV will need intelligence from the regional bureaux concerning the projects under design or re-design, and OEV will need to invest in assembling and targeting specific evaluation lessons. As part of the BPR, WFP is in the process of redesigning its project review and approval process, a process which OEV could influence to ensure that evaluation findings are better utilised in project design, and not only the prior evaluations of that specific project.

13. Operations Evaluations

13.1. Deficit in Operations Evaluations 2008-2013

156. The 2008 Evaluation Policy committed the organization to undertake 30 operations evaluations per year, including 20 decentralized (i.e. conducted by country offices or regional bureaux). In practice, between 2008 and 2013, the annual number of operations evaluations ranged from 4 to 15.

157. In 2009-10, OEV ran an evaluation skills training programme for country and regional staff with the aim of increasing the numbers of decentralized operations evaluations. This programme was stopped after it led to no noticeable increase in operations evaluation activity⁸⁰. The Board subsequently agreed that OEV should focus on strategic, policy and impact evaluations, a decision that has led to the OEV evaluation portfolio becoming of greater strategic relevance. However, the low level of operations evaluations remained a concern⁸¹.

13.2. Progress in the OEV Operations Evaluations Series 2013-15

158. In 2012, OEV proposed a strategy for attaining the operations evaluations policy target over a period of three years, under a new model agreed to fund these evaluations. The intention is to hand over, progressively, the responsibility for operations evaluations to management, in particular to regional bureaux. In the Panel's view, given the current lack of capacity in the regional bureaux, and the absence of a plan to reinforce their capacity, this is unrealistic.

159. OEV noted that, through their engagement in the operations evaluations series, the Regional M&E Advisers are becoming more familiar with the evaluation management process. However, some of the advisers expressed concern that they will not be able to manage evaluations to the same standards as OEV and, in any case, their workload may not allow them to give the commissioning of evaluations sufficient attention. These advisers have already demonstrated their value in working with OEV to plan evaluations in their regions but the Panel is concerned that they will not have the capacity to take the proactive leadership in the planning and

⁸⁰ Although there may be a legacy from this training in the relatively good quality of some decentralized evaluations.

⁸¹ Information on the numbers of operations being evaluated through decentralized evaluation is incomplete. Whatever the level of coverage, decentralized evaluations are not reported to the Board either in individual or synthesis form, so they are not considered either by senior management or governance.

commissioning of evaluations in their region after 2015 when OEV is due to hand over responsibility for operations evaluations.

160. In October 2012, a Special Account was established to “meet accountability requirements for single WFP operations”. The DED wrote to Regional Directors announcing the scheme. Since then, the six Regional M&E officers have become the prime interlocutors with OEV for the new operations evaluations series. The establishment of the Special Account is a positive and innovative development, and OEV has made good progress in recruiting 7 LTA companies and initiating the first 12 evaluations. OEV is aiming to conduct 66 evaluations over three years using US\$10.17 million and has compiled a work programme in consultation with regional bureaux. (See also Annex 13).
161. The question of how resources will be secured for operations evaluations beyond 2015 is yet to be resolved. The scheme is due for a ‘Proof of Concept’ initial review during 2014. Given that this is the first time the organization has succeeded in ensuring that funding for operations evaluation is guaranteed, there may be a case for continuing the special account past 2015, if it proves successful. The Proof of Concept exercise should seek to establish what the on-going central subsidy might be, i.e. the proportion of funds that will not be recharged each year.
162. The Panel considers it a significant achievement for OEV to have launched the operations evaluations series quickly and at scale, and that OEV has managed to outsource a large proportion of the management of the evaluation process to the LTA companies. There has been some limited impact on the Office in terms of capacity to carry out its traditional role. The Panel’s concern is that this has been done in parallel with, and with limited understanding of, current country office initiatives in decentralized operations evaluations. The BPR assessment of WFP’s decentralized evaluation capacity is expected to close the information gap and provide options for making OEV commissioned and country office-commissioned operations evaluations more complementary.
163. The operations evaluations outsourcing arrangement passes most of the evaluation management load to the seven LTA companies, in effect adding to OEV’s evaluation management capacity. This innovative arrangement seems to be working well for OEV so far, even though OEV estimates that 12 days of OEV time are still required for each evaluation, even with the outsourced model. The model has also been designed with a view to regional bureaux being able to take over the management of the LTAs by 2016. However, the Panel is concerned that regional bureaux may not have the capacity to manage the LTA companies. It is also possible that country offices may see the model as too expensive and opt for alternative approaches instead⁸².
164. The evaluation policy statement that “the Office of Evaluation compiles the findings of all single-operations evaluations into an annual report”, was in reference to evaluations conducted by OEV. In future, the Board will be interested in the synthesis reports from the current OEV operations evaluations.

⁸² Currently planned to average \$130k per evaluation. Country offices may push for the regional bureaux to adopt cheaper options, for example, finding 1-2 individual consultants to undertake operations evaluations, with a budget of, say, US\$50-60k, which may then undermine the LTA arrangement. This scale of this risk will be informed by the ongoing BPR capacity assessment of decentralized evaluation.

13.3. Evaluation of Emergency Operations

165. At the outset of the Peer Review, OEV and the DED expressed concern that WFP was not undertaking enough evaluations of emergency operations (EMOPs) compared with the very high proportion of WFP spending that goes to EMOPs, some 80%. By the time of its second visit, the Panel found the situation somewhat changed.
166. OEV is embarking on a series of strategic evaluations of WFP emergency preparedness and response, including PREP, a US\$90 million programme to develop WFP preparedness capacity. OEV is also part of an interagency accountability and learning initiative with ALNAP for the response to the Syria crisis⁸³. In parallel, the Emergency Preparedness Division systematically generates lessons learned from major emergencies and has recently produced a synthesis of lessons-learned document for recent Level 3 responses.
167. These activities show increased attention to emergency-related evaluation by OEV but the current plan for the OEV operations evaluations series includes only two EMOPs. There is an on-going challenge of OEV evaluation processes taking too long to produce findings to be relevant to fast moving emergency situations. A lighter, quicker model is needed for EMOP evaluations.
168. In the past, OEV has been keen to avoid duplicating inter-agency real-time evaluations (IA-RTE), which in practice have gained limited traction⁸⁴, and OEV has not undertaken real time evaluations of major WFP emergency operations in their early stages. It would be useful for the Director OEV and the Emergencies Director to discuss whether there is a need for WFP real time evaluation, or whether both Directors consider OME's Emergency Preparedness and Response Branch lesson learning exercises to be sufficient. WFP should give consideration to making the evaluation of all WFP Level 3 emergency responses mandatory, either as part of an inter-agency exercise or as a WFP-only exercise.

14. Decentralized Evaluation

14.1. Activity levels and Expenditure for Decentralized Evaluation

169. According to country office self-reporting, there has been an overall increase in activity from approximately 12 to 40 decentralized evaluations per annum in the period 2008-2012, although many exercises labelled as evaluations may prove not to be so on inspection by the BPR assessment exercise.
170. Interviews and survey results indicate a strong interest in evaluation at regional and country level in WFP, but country offices also report several constraints: insufficient staff time, lack of financial resources, lack of qualified staff, and lack of standard guidance and tools. (Resources for evaluation may have actually declined over time).

⁸³ <http://www.syrialearning.org/>

⁸⁴ The IASC Transformative Agenda has moved away from RTE for Level 3 emergencies to 'Real Time Operational Review' (RTOR), which, as a management-led review, may not involve UN evaluation offices.

171. In the staff survey, country offices were asked ‘How many evaluations has your office conducted since the start of 2009?’ The average for the 16 offices responding was 2-3, with 3 offices reporting 6 or more. Country offices reported that they intended to continue with roughly the same level of evaluation activity. 2 offices reported spending more than US\$100,000 each year on evaluations, 4 reported that they spent nothing, with the most common expenditure bracket being US\$25-50,000 per annum.

172. As noted earlier, WFP systems cannot show how much has been spent on decentralized evaluation. This is an important weakness. JIU figures collected for its current assessment of evaluation functions across the UN show that UNICEF and UNDP spent close to US\$11 million and US\$10 million respectively on decentralized evaluation in 2011. Because WFP is not coding expenditures against evaluation, it is not possible to benchmark WFP against such figures.

14.2. Responsibilities for Supporting Decentralized Evaluation

173. According to the Evaluation Policy, OEV’s mandate includes “developing the evaluation capacities of WFP at various levels” but also places the responsibility for decentralized evaluation with the regional bureaux and country offices, without defining their respective roles. In line with the evaluation policy, the WFP Programme Guidance Manual states, “Country offices, regional bureaux and other headquarters divisions have the responsibility for managing decentralized and self-evaluations”. The manual continues “Strengthening these ... is a central objective of the M&E strategy approved in February 2012 and will be the focus of significant work by RMP, with support from the Office of Evaluation over 2012-2014”.

174. OEV and RMP have developed a good working relationship over the OEV-led operations evaluations and in relation to Regional M&E Advisers, but there is still insufficient clarity over the division of labour between them with regard to decentralized evaluation.

175. The Panel proposes that OEV is made responsible for standard setting and support to centralized and decentralized evaluation, while RMP retains responsibility for guidance and support to management reviews and all other forms of reporting and lesson learning processes as part of the management functions of country offices and regional bureaux. OEV will need to continue to support the Regional M&E advisers in their evaluation role⁸⁵.

14.3. Evaluation Capacity of regional bureaux

176. So far, Regional Directors have not resourced evaluation adequately for their bureaux to be able to take responsibility for managing decentralized evaluation in their regions. Regional Directors are expected to provide oversight of the evaluation function in their regions and the directors acknowledge that their offices have limited capacity to play this role. Six P4 Monitoring and Evaluation Officers were appointed in 2013, line-managed by the Regional Programme Adviser and with a technical reporting line to RMP.

⁸⁵ The six Regional M&E advisers will attend a ‘mini-IPDET’ (International Program for Development Evaluation Training) course in Kosovo in May 2014.

177. These posts are seen by WFP management to represent a significant increase in technical capacity, responsibility, leadership and oversight of WFP's monitoring and evaluation function but, with regard to evaluation, this is not so. Regional Directors are clear that the main task of these M&E advisers is monitoring, rather than evaluation, and that these posts do not represent any significant increase in evaluation capacity. The advisers themselves are keenly aware that they have limited time for evaluation, though some have initiated useful assessments of evaluation activity in their region. Their ability to initiate evaluations and or deliver on evaluation capacity development for country offices will be limited.

178. Despite this, there is a risk that regional bureaux staff and/or management will silo off evaluation to the M&E adviser only, while regional bureaux also have, for example, Regional Programme Advisers, some of whom have undertaken evaluations successfully in the past, and who, in any event, should be engaging with evaluation activities. If Regional Directors want their bureaux to become effective in supporting evaluation, they should be considering the capacity of the bureau as a whole, not just that of their M&E adviser.

179. In a May 2013 teleconference organized by the DED on the Peer Review, regional bureaux commented that:

- ❖ They are looking for more assistance from HQ;
- ❖ The current attempt to revive operations evaluation 'cannot be allowed to fail';
- ❖ Evaluations commissioned from within the region may not be seen as independent. In practice the independence of evaluations has varied with the quality of the consultants;
- ❖ Resourcing of evaluation is the key issue.

14.4. Roles and Responsibilities

180. In OEV's 2012 Concept Note on its support to the M&E Strategy, OEV committed itself to devising standards for decentralized evaluation and for management response, reporting, publication and follow-up, for developing guidance and training materials, and for providing a periodical post hoc quality assessment of decentralized operations evaluations, while RMP and regional bureaux were to establish an accountability framework for decentralized operations evaluations, together with a Management Information System for such evaluations, and for capacity building. The Panel's view is that all these functions should be carried out by OEV in partnership with regional bureaux.

181. There is no reporting relationship between OEV and the six Regional M&E Advisers, which are under RMP's technical direction. The Panel proposes that this should be changed to provide also for a technical reporting line on evaluation to OEV, which should have oversight of their performance in relation to evaluation, while RMP retains its technical line with regard to monitoring. On evaluation, the advisers should report to the Deputy Head or another senior officer in OEV.

182. In UNICEF and UNDP, the Directors of Evaluation carry some responsibility for the decentralized evaluation function but this is not yet the case in WFP. The Evaluation Policy states that: “The Director of the Office of Evaluation is accountable for ensuring the quality, credibility and utility of evaluations” (para 18) but this appears to refer only to OEV evaluations. Country offices and regional bureaux have few or no staff dedicated to evaluation, in contrast to UNICEF and UNDP, where some country offices do, and the remainder have M&E officers.

183. Given that 12 WFP country offices account for 80% of turnover and cover 65% of all beneficiaries (as at mid-2013), there is a very good case for each of these country offices to appoint a dedicated evaluation officer⁸⁶, given that each has a budget of \$150 million plus. These evaluation officers would be funded from the Direct Support Costs (DSC). This may be easier than adding to regional bureaux capacity, which would mean a further call on PSA. However, the Panel considers that country based evaluation officers should only be appointed if they are to receive technical support and guidance from regional bureaux and OEV, and have a network to join. Country office evaluation officers should not report through the management line that directly oversees projects but should report to, for example, the deputy country director.

15. External Relations

184. While WFP is rightly concerned with maintaining and developing an effective internal evaluation function, it also has a responsibility to play a role in the evaluation community and especially within the UN system. The QCPR has emphasised the increasing need for joint working and evaluation between agencies and the WFP Board has an interest in WFP’s role vis-à-vis evaluation the other Rome-based agencies.

185. Of the evaluation networks, ALNAP (Active Learning Network for Accountability and Performance in Humanitarian Network) and the UN Evaluation Group (UNEG) are the most important forums for WFP. The Director was the UNEG Vice Chair for a period of months in 2013 that included the development of a new UNEG Strategic Plan. In the review period, WFP has been a member of UNEG task forces on impact evaluation, joint evaluation, and gender equality and human rights, plus working groups organized by UNOCHA on real time evaluation and on joint humanitarian impact evaluations. WFP also participates in the Network of Networks for Impact Evaluation (NONIE) and maintains relations with the International Initiative for Impact Evaluation (3IE). ALNAP indicated that WFP has made a consistently strong intellectual contribution to the network. Since ALNAP was founded, WFP has not been a member of its Steering Committee but OEV has now taken up this role.

186. WFP is a major humanitarian player and, in its external engagement, OEV has now decided to focus on the improvement of humanitarian evaluation, including inter-agency evaluation. Accordingly, the Director has stepped down from a role as vice-chair of UNEG in order to devote more time to inter-agency humanitarian evaluation through the IASC, UNEG, ALNAP and other forums.

⁸⁶ This will require a new job profile as there is currently no such post in WFP.

187. According to OEV, evaluation networks and societies have shown particular interest in its impact evaluation methodology, on which OEV members have been invited to make presentations at several evaluation forums. WFP will continue to be asked to share its experience in impact evaluation, and in general engagement in UNEG, ALNAP, IASC, and evaluation associations.

188. It is important that WFP continue to share its evaluation experience with other organizations, networks and associations because it has much to contribute. It can also learn from these entities, and sustained engagement can contribute to the continued professional development of OEV and other M&E staff. In the Panel's view, OEV is making a constructive contribution to external networks and evaluation associations, and is targeting the topics of most relevance to WFP.

16. Selected UN Comparator Evaluation Functions

189. Of the Rome-based agencies, IFAD programme of loans of grants for 2013 totalled US\$1.3 billion, while its Independent Office of Evaluation (IOE) had 12.5 professional and 6 general service staff and a total staff and non-staff budget US\$6.01 million, slightly higher than WFP. IOE retains US\$1.2 million (20% of the annual budget) for promoting learning and knowledge management, including evaluation syntheses and learning themes, and communication and outreach. IFAD was subject to peer review in 2010 by the Evaluation Cooperation Group⁸⁷.

190. FAO's Evaluation Office has dual reporting lines to both the FAO Council and to the Director General. FAO has an internal Evaluation Committee that reviews evaluation policy and vets and follows up on management responses. Every project of more than US\$4 million is independently evaluated. All projects contain a financial component for evaluation, which the Evaluation Office draws on to fund evaluations. Funding for evaluation has been gradually increasing and is now set at 0.7% of the total FAO budget of US\$1 billion. There are mandatory follow up reports by independent consultants, after one year for project evaluations and two years for thematic evaluations.

191. The Rome based agencies do not provide useful comparators for decentralized evaluation, as their evaluation functions are centralized. UNDP and UNICEF both have similar overall funding to WFP and a network of country offices. Both organizations have invested heavily over a long period to develop decentralized evaluation. According to the UNDP annual evaluation report for 2011, 248 decentralized evaluations were commissioned during the year by UNDP's 62 decentralized M&E staff. However, only 20% of these evaluations were rated Satisfactory or Highly Satisfactory, and UNDP is reviewing its strategy with regard to decentralized evaluation.

192. In 2012, 95 per cent of UNICEF-supported evaluations were managed by field offices. Each of the 7 regional offices has a P5 Regional M&E officer and there were 114 UNICEF staff at P3-equivalent level or higher with 'evaluation in their job title', spending on average 14% of their time on evaluation. In 2010, UNICEF country and regional offices spent US\$11.2m on evaluation and UNICEF conducted 140 evaluations. In 2011, 43% of evaluation reports were rated satisfactory or

⁸⁷<https://wpqr1.adb.org/LotusQuickr/ecg/Main.nsf/7e6e83714d63fac348257731002a960f/3f296435ce2c24c648257731002a9639/?OpenDocument>

better compared with 20% in 2002, and 93% of evaluations included a management response, up from 10% in 2009. (See Annex 14 for more details).

17. Summary assessment against the Peer Review criteria

56. The terms of reference for the peer review asked the panel to make an independent assessment of the WFP evaluation function against the three principal criteria of independence, credibility and utility.

Independence

57. Overall, the functional independence of centralized evaluation in WFP is high, without full structural independence. Intentionality⁸⁸ is well established for centralized evaluation and WFP is making well-planned and rational decisions about the selection of OEV evaluations. Functional independence of the centralized evaluation function is expressed as follows:
- ❖ The Director of OEV consults WFP managers when drawing up the OEV work plan but retains the final choice over evaluation subjects in the work programme submitted to the Board.
 - ❖ The Director has full discretion in the selection of subjects for evaluation.
 - ❖ The Director has full authority over the management of OEV's human and financial resources for evaluation, but not over the level of those resources.
 - ❖ OEV has not generally sought funding for its activities directly from donors.
 - ❖ OEV is independent in supervising evaluators.
 - ❖ The Director of OEV submits evaluation reports to the Board and publishes evaluations on WFP's public website without interference from management.
58. The Director of OEV issues an annual evaluation report on progress against the OEV work programme, summarizing the main OEV evaluations for the year and making new recommendations based on a synthesis of OEV evaluation findings, which require a management response that is presented to the Board.
59. The panel considers that independence and utility could be strengthened by OEV assuming full ownership of its evaluation reports, as OEV is better placed than evaluation consultants to determine the final shape of each report and its recommendations.
60. There is a risk that WFP may overemphasize the need for independence in its decentralized evaluations, where the primary aims are impartiality and quality. Ensuring a degree of separation between the project manager and those organizing the evaluation will increase the level of confidence in being able to achieve a balanced, impartial result.

Credibility

61. The evaluations managed by OEV are highly credible and are respected by the Board, senior management and external parties. The panel finds that OEV evaluation reports are of high quality, backed by a highly structured set of standards and procedures in the EQAS.

⁸⁸ Intentionality is an evaluation norm referring to a clear organizational intention to plan and use evaluation to inform decision-making and improve performance.

62. Compared with other United Nations evaluation functions, WFP's central evaluation function performs well in terms of spending, numbers of evaluations, transparency and quality assurance. The Director and staff of OEV are seen as credible in inter-agency fora and evaluation networks and are often asked to participate because they make valuable contributions. OEV has competent staff but insufficient capacity for some of the functions it is expected to perform, especially in helping to ensure that evaluation results are utilized and in developing evaluation capacity across WFP. OEV's reputation remains strong, based on its generation of high-quality reports.
63. The good use that WFP makes of evaluation reports is further evidence of their perceived reliability and relevance. The low coverage of operation evaluations since 2010, compared with WFP's policy commitment, is being addressed. Future credibility depends on OEV ensuring that its portfolio of evaluations remains relevant to WFP as it evolves.
64. The credibility of the evaluation function is somewhat reduced by shortcomings in the management response system and by field staff's perception that evaluations are not always sufficiently realistic and rooted in the country context.

Utility

65. Over the review period, evaluation has become more accepted; the past tendency for it to be marginalized has been reduced. All types of OEV evaluation are seen to be relevant and useful and are used across WFP. OEV evaluations are a useful accountability tool as their reports are discussed by the Board, with the exception of those for single operation evaluations, which are presented in synthesis reports.
66. The review found several examples of WFP staff using evaluation to benefit country programmes and global policy and strategy development. In general, the learning from evaluation is limited by the lack of a corporate knowledge management function. However, OEV-managed evaluations are used well, not only for accountability to the Board, but also in influencing policy and strategy. They are also well used at the field level for immediate programme improvements.
67. OEV intends to make greater use of evaluability assessments, especially for WFP strategies. Evaluability assessments provide not only a check on whether the proposed intervention can be evaluated, but also a systematic analysis of its coherence, feasibility and metrics. This should help to increase the quality of assessed strategies. The panel considers that evaluations can be made even more influential by selectively targeting evaluation lessons towards WFP project design, evaluability and evaluation planning.

18. Recommendations of the peer review

68. The panel presents the following recommendations for consideration by WFP management and the Board.
69. The panel's recommendations are divided into two sets. Recommendations 1 to 10 apply regardless of the evaluation model that management selects under Recommendation 1. Recommendations 11 to 15 are relevant only if WFP selects an evaluation model that includes the intentional strengthening of decentralized evaluation, i.e. model 2 or 3.

70. **Recommendation 1: Evaluation model for WFP.** WFP management should take decisions concerning the most appropriate model for the evaluation function in WFP, using the three models set out in paragraph 15 and making sure that the necessary human and financial resources are made available to implement the preferred model. These decisions should be made when the results of the BPR assessment of decentralized evaluation capacity become available and before the management response to the peer review is submitted to EB.2/2014.
71. **Recommendation 2: Evaluation policy.** WFP should revise the 2008 evaluation policy in line with the selected evaluation model to ensure that financial resources for evaluation are protected; criteria for evaluation selection and coverage are clearly specified; the roles of the Board, OEV and management at different levels of the organization are clarified; evaluation terminology is revised; and the typology of evaluations is brought into line with current WFP practice. The policy should also specify WFP's role in developing national evaluation capacity, and the involvement and role of partners in country-level evaluation.
72. **Recommendation 3: Oversight of the evaluation function.** The Board should request the development of a set of key performance indicators to support its oversight of evaluation across WFP, giving OEV adequate time to establish the necessary systems. The panel suggests the following as a possible set of indicators:
- 1) the numbers, types and coverage of evaluations taking place across WFP;
 - 2) the human and financial resources used for evaluation;
 - 3) progress in the development of WFP's capacity and competence in evaluation; and
 - 4) ratings of the quality of evaluations, including decentralized evaluations if model 2 or 3 is selected.
73. The Board should ensure that WFP management has systems and processes in place to maximize the use of evaluation results in policy and strategy development, as well as in project and programme design. The strategic use of evaluation results should be the Board's main focus in its deliberations with WFP management, over and above discussion of individual evaluations.
74. **Recommendation 4: Management response.** WFP management should improve the quality and effectiveness of management responses to evaluations, in particular by giving due attention to the ownership of follow-up. This requires the active engagement of relevant senior managers and other stakeholders during the evaluation process and beyond. Specifically:
- ❖ OEV should continue to improve the quality of evaluation recommendations by ensuring that they are clear, realistic and relevant to WFP's context and structures. OEV should also strengthen its dialogue with evaluation consultants and management at the draft report stage, including through its current good practice of holding workshops with major stakeholders so that, as far as possible, recommendations are agreed before the evaluation is complete – although OEV may retain recommendations with which management does not agree.
 - ❖ WFP management should nominate a member of the Executive Management Group to be responsible for engaging management in each OEV evaluation from its outset through to its presentation to the Board, with RMP in a supporting and advisory role.

- ❖ WFP should provide staff with online access to the RMP database for tracking progress in implementing the management responses to recommendations, so that the staff responsible can enter completed actions online, as they already do for audit recommendations.
 - ❖ WFP should replace the current management response template with a more detailed format for setting out action plans in which management can specify more clearly how it intends to act in response to each evaluation.
 - ❖ WFP management may request OEV's informal comments on how well a draft management response corresponds to the findings, conclusions and recommendations of an evaluation. The management response will still be fully owned by management, and OEV will in no sense be approving or taking responsibility for its content.
75. **Recommendation 5: Vesting evaluation independence in OEV.** For each evaluation it manages, OEV should take ownership of all aspects of the evaluation report, including the recommendations. Independence is vested in OEV, not in evaluation consultants.
 76. **Recommendation 6: Evaluation quality assurance.** OEV should commission an external consultant to review the EQAS and identify evaluation steps and elements that can be simplified, reducing the time inputs of staff and consultants, the number of steps and the duration of evaluations. OEV should also take account of the panel's proposals on improving evaluation quality.
 77. **Recommendation 7: Utilization of evaluation.** In redesigning its project and programme planning and approval process, WFP management should ensure that evaluation evidence is taken into account, not only from any evaluations of the project under consideration, but also from other relevant evaluations; management should request OEV's assistance in identifying such evaluations if necessary.
 78. OEV should strengthen its inputs to WFP's revision of project and programme planning, design and approval processes to encourage the use of evaluation evidence and improve arrangements for evaluation within projects, in particular by promoting the design of projects that can be evaluated effectively (evaluability); the use of prior evaluations of a project, and other relevant evaluations; and planning to facilitate evaluation from the outset.
 79. **Recommendation 8: Evaluation training.** OEV should develop evaluation training in partnership with the Human Resources Division, including modules for evaluation planning and management in WFP management training; and WFP induction courses covering the essential elements of evaluation concepts and purposes, and the evaluation policy.
 80. **Recommendation 9: Roles and responsibilities.** WFP management should ensure that regional monitoring and evaluation advisers have a reporting line to OEV for technical oversight and support on evaluation, and review of their performance with regard to evaluation.
 81. WFP management should clearly delineate the roles of OEV and RMP, giving OEV the lead responsibility for evaluation standards and guidance at all levels of WFP, with RMP taking the lead on monitoring and all forms of project/programme review by management.
 82. WFP management should request Regional Directors to take responsibility for receiving and reviewing management responses to the single operation evaluations and decentralized evaluations in their regions.

83. Internal Audit and OEV should agree how to identify audit risks regarding compliance with the WFP evaluation policy, and should develop standard questions for testing the compliance of individual country offices, regional bureaux or Headquarters divisions where risks of non-compliance are identified.
84. **Recommendation 10: Monitoring and evaluation guidance.** OEV should work with RMP to make sure that WFP guidance on evaluations and, separately, on monitoring and review are distinct but mutually compatible and understood by country offices; that any regional and global training of M&E officers is coordinated; and that management reviews are used as key inputs to operation evaluations, country portfolio evaluations and other types of evaluation.

Further recommendations

85. Should WFP select an evaluation model that includes further enhancement of decentralized evaluation – evaluations managed by regional bureaux and country offices – the following recommendations would also apply.
86. **Recommendation 11: Evaluation strategy.** OEV should develop an evaluation strategy in line with the selected model for evaluation. The evaluation strategy, separate from the monitoring and review strategy, should set out how WFP will develop evaluation capacity, resourcing, selection, coverage, and utilization across the Organization.
87. **Recommendation 12: Role and designation of the Director of Evaluation.** To support the Board’s governance of the function, the Board should request the Director of OEV to oversee and report on the evaluation function across WFP. In its annual report, OEV should include an assessment of the quality of decentralized evaluations and the salient issues emerging from these evaluations. In line with this change in role, the Board should request WFP management to redesignate the Director of OEV as the Director of Evaluation.
88. To avoid any possible conflicts of interest, the Board should limit the Director’s term to a single period of six years, non-renewable, and without the right of return to WFP. This would replace the current arrangements – a term of four years renewable once, with the possibility of re-employment in WFP.
89. The Executive Director should redesignate the Director of OEV as the Director of Evaluation. The Director should retain direct responsibility for OEV-managed evaluation and should also be responsible for standard-setting, oversight and support of evaluation across WFP.
90. **Recommendation 13: Decentralized evaluation standards and guidelines.** OEV should develop appropriate and realistic standards for decentralized evaluations based on a simplified version of its EQAS and reflecting the norms and standards of the United Nations Evaluation Group. Once these standards are agreed, OEV should issue guidelines to country offices on the management of decentralized evaluations.
91. **Recommendation 14: Evaluation expertise.** WFP management should ensure that the WFP People Strategy includes the development of a staff cadre for assessment, monitoring and evaluation in WFP, so that regional bureaux and country offices have the human resource capacity and expertise to implement the evaluation strategy.
92. **Recommendation 15: Evaluation database.** OEV should develop an online database for all centralized and decentralized evaluations into which country offices and regional bureaux can

upload their evaluation reports. OEV can use this database to make an annual assessment of the quality of evaluation reports, with a summary included in the Annual Evaluation Report. WFP may be able learn from the database solutions developed for this purpose by the United Nations Development Programme and the United Nations Children's Fund.

Response to all Recommendations

93. In responding to and implementing these recommendations, WFP will need to adopt a sequenced approach to the development and revision of documents relating to evaluation, as follows:

- ❖ Develop the management response to the peer review. This will require inputs from management and the Board, with advice from OEV (OEV should not prepare the response). The response will need to take into account the BPR assessment of decentralized evaluation capacity, due by May 2014.
- ❖ OEV adjusts its work programme for 2015 taking into account the results of the peer review and the BPR assessment of decentralized evaluation capacity.
- ❖ Revise the evaluation policy.
- ❖ Develop an evaluation strategy, either in parallel with or after revision of the evaluation policy, in line with the selected model for evaluation.
- ❖ Update the references to evaluation in programme guidance.
- ❖ Make sure that the evaluation strategy and the monitoring and review strategy are consistent.

Abbreviations

AED	Assistant Executive Director
AER	Annual Evaluation Report
CEB	Chief Executives Board for Coordination
CIDA	Canadian International Development Agency
COMET	Corporate M&E Tool
DAC	Development Assistance Committee
DED	Deputy Executive Director
DSC	Direct Support Costs
ED	Executive Director
EMOP	Emergency Operation
FAO	UN Food and Agriculture Organization
GHD	Good Humanitarian Donorship
IASC	Inter-Agency Standing Committee
IFAD	International Fund for Agricultural Development
JIU	United Nations Joint Inspection Unit
M&E	Monitoring and Evaluation
MRE	Monitoring, Reporting and Evaluation
M&SE	Monitoring and Self-Evaluation
OECD-DAC	Organization for Economic Co-operation and Development /Development Assistance Committee
OEV	WFP Office of Evaluation
PRRO	Protracted Relief and Recovery Operation
PSA	Programme Support and Administration
QCPR	UN Quadrennial Comprehensive Policy Review
RFP	Request for Proposal
RMP	Performance Management and Monitoring Division
TOR	Terms of Reference
UNDP	UN Development Programme
UNEG	UN Evaluation Group
UNICEF	UN Children's Fund