

Executive summary

Overview

The purpose of this Professional Peer Review is to provide an independent assessment of the functioning and quality of the WFP evaluation function. It is intended for use by WFP's leadership, WFP's Executive Board and the OEDE, as well as the DAC and United Nations Evaluation Group (UNEG). The findings of the Professional Peer Review are to be presented to WFP's Executive Board meeting in February 2008 and will inform further discussions and decisions about the function and administration of OEDE. The assessment report is also to be provided for information to the DAC/UNEG Task Force and will be made publicly available.

The Professional Peer Review takes the central evaluation function, i.e. OEDE, as its starting point but also includes analysis of decentralised evaluation in WFP and the ways in which the organisation as a whole engages in evaluation. Significant attention is paid to the attitudes and perceptions that frame how WFP uses and learns from evaluation. In line with the *Framework for Professional Peer Reviews of Evaluation Functions in Multilateral Organisations*, established by the DAC/UNEG Task Force, the Peer Review of the Evaluation Function of WFP has applied three core criteria that need to be satisfied for evaluation functions and products to be considered of high quality:

- A. *Independence* of evaluations and the evaluation system(s). The evaluation process should be impartial and independent in its function from the process of policy making, and the delivery and the management of assistance. Independence of the evaluation function is a pre-condition for credibility, validity and usefulness.
- B. *Credibility* of evaluations. The credibility of evaluation depends on the expertise and independence of the evaluation managers and the evaluators and the degree of transparency of the evaluation process.

Credibility requires that evaluations should report successes as well as failures. Partner countries and representatives of other partners should, as a rule, fully participate in evaluation in order to promote credibility, ownership and commitment.

- C. *Utility* of evaluations. To have an impact on decision-making, evaluation findings must be perceived as relevant and useful and be presented in a clear and concise way. They should fully reflect the different interests and needs of the many parties involved in humanitarian assistance and/or development co-operation. Ensuring the utility of evaluations is only partly under the control of evaluators. It is also a function of the interest of managers, and member states through their participation on governing bodies, in commissioning, receiving and using evaluations.

The findings of this Review are based on the following:

- Preparatory work included a desk review of relevant documentation and initial interviews at WFP headquarters in Rome.
- Based on these preliminary data collection exercises a preliminary factual report was prepared and circulated to OEDE for verification and comments.
- Field visits were made to WFP Regional Bureaux and Country Offices in Johannesburg, Lilongwe, Jakarta, Bangkok and Vientiane where WFP staff, partners and other stakeholders were interviewed.
- A meta-evaluation of twelve OEDE and decentralised evaluations was conducted. In conjunction with this desktop assessment, interviews were also conducted with evaluation team leaders, evaluation managers, Country Office/Regional Bureau staff and others who were involved with these evaluations.
- This data was compared with a review of ALNAP's assessments of WFP evaluations.
- A web-based survey of the views of WFP staff was conducted which received 87 responses.
- Peer Panel interviews with selected stakeholders were conducted, including: the Executive Board (a meeting held with fourteen members), executive staff, OEDE (director and staff), regional and country directors, division directors, regional monitoring and evaluation focal points, and other WFP staff.
- In total 124 interviews have been carried out.
- The draft assessment report was discussed with WFP in a Review Meeting on October 25, 2007. Based on this discussion the final assessment report has been produced.

The three criteria or dimensions used for the analysis – independence, credibility and utility – relate to both objective and subjective judgements. The UN’s normative framework (the UNEG Standards and Norms, as well as the OECD DAC criteria used in this Review) is based on a mixture of corporate, cultural and perception-based standards. The source information for the Review is therefore drawn from an analysis of WFP’s organisational structure, related financing, corporate managerial practices and the subjective opinions of staff regarding the three dimensions. The latter perceptions relate to WFP’s readiness to utilise evaluation and are hence particularly relevant in an assessment of an organisation’s evaluation function; they have therefore formed a significant component of the source data used by the Peer Panel.

Main findings and conclusions

General

The Panel concludes that the *Independence* of the WFP evaluation function is adequate in comparison to similar organisations; that the *Credibility* of *products* of the WFP evaluation function is uneven, and that the *process* of the function is somewhat more credible but also problematic; and that the criteria of *Utility* of the WFP evaluation function are partially met with regard to contributing to programming but that structures and mechanisms to promote utility are weak in most other respects.

OEDE is a strong unit with committed, well-trained and highly motivated staff. During the past seven years OEDE has invested much efforts on improving WFP’s evaluation function. OEDE is now addressing a number of weaknesses such as the quality of reports, limited attention to strategic evaluation planning, lack of full management engagement and follow-up etc. The Peer Review Panel feels that if these changes are implemented, they will address many of the findings of the Panel.

The evaluation function is of more variable quality at the level of Regional Bureaux and Country Offices. Levels of motivation and of invested resources depend on the interests and priorities of the offices concerned. Given the nature of decentralisation within WFP, OEDE has not been in a position to exert significant influence on the quality of these evaluations. Nonetheless, the findings of both OEDE and Regional Bureaux led evaluations are routinely reflected in the design of follow-up programmes and have some influence on policy formation. All OEDE evaluations are made public and submitted to the Executive Board.

Independence

The Panel concludes that the *Independence* of the WFP evaluation function is quite good in comparison to similar organisations. Though generally satisfactory, the Panel has noted a danger that this independence could be eroded in the future due to a lack of recognition of the importance of evaluation independence among many WFP staff.

The Executive Board and Executive Director fulfil responsibilities regarding the appointment of a professional head of evaluation in an appropriate manner. Evaluation resources are currently safeguarded. OEDE is outside of line management while at the same time sufficiently integrated into WFP leadership structures to facilitate impact. This is seen by the Panel as being largely appropriate.

There are, however, some weaknesses in the independence of the evaluation function. Accountability for the implementation of recommendations is unclear. Some OEDE staff are concerned that their careers may be affected by their evaluation role, which could lead to inappropriate risk averse behaviour in their management of sensitive evaluations. There are also insufficient safeguards to prevent partiality and conflicts of interest amongst external evaluators. The role of Regional Bureaux in both oversight and advisory support to Country Offices has problematic implications for the independence of their role in decentralised evaluation in that their evaluation function involves public critique of programming while they must also ultimately maintain collegial day-to-day relationship with Country Offices.

With some notable exceptions, OEDE usually has full access to information and is free to express its findings. Evaluations take the views of all relevant stakeholders into account, but the evaluation process does not provide for sufficient dialogue with stakeholders outside of WFP.

Credibility

The Peer Panel has assessed *Credibility* in terms of both the evaluation products and the processes through which evaluations are managed. The Panel has assessed WFP evaluations with regard to both formal quality standards and also the perceptions of credibility within WFP at different levels of the organisation. In some cases these perceptions are quite different from the Panel's assessment of quality. Some decentralised and self-evaluations are seen as highly credible within the countries where they have been conducted since they relate to the core concerns of the relevant stakeholders. Several of those evaluations have nonetheless been judged as unsatisfactory by the Panel and indeed by most stakeholders at headquarters due to inherent faults in methods and failure to adhere to basic quality standards. The Panel concludes that the overall credibility of *products* of the WFP evaluation function is uneven, and that the *process* of the function is somewhat more credible but also problematic.

An official evaluation policy should provide the foundation for credible evaluation. Such a policy exists, but it consists of a layered series of documents which detracts from clarity and applicability. Evaluation policy is not sufficiently used to guide practice.

Evaluators and Regional Bureaux have been unclear regarding what is expected in terms of quality due to a lack of specification within OEDE itself and concerns that headquarters' expectations do not take into account resource and time constraints in the field. This first deficiency is in the process of being addressed by drafting of clearer standards and procedures. OEDE also intends to take on a more proactive role in supporting decentralised and self-evaluation, but problems will remain due to the shrinking levels of human and financial resources, especially within the Regional Bureaux.

On the whole the OEDE evaluation function is impartial. The views of all stakeholders are often sought, but there appears to be an uneven emphasis on stakeholders who are more accessible and articulate, with beneficiary views in particular under-represented. There are claims that some hired evaluators have preconceived ideas and that they are influenced by conflicts of interest. A more competitive approach to recruiting external evaluators could counter this problem and increase the chances to find well qualified consultants.

The process of preparing for evaluations, management and advising and supporting teams in the field is handled in a highly professional manner by OEDE. Terms of Reference are generally of good quality, but they are at times too standardised and over-ambitious. This is partially related to weaknesses outside of the evaluation function, as programme goals are often not clear or well structured. There is in most cases a dearth of monitoring data and existing data is often accepted without sufficient scrutiny.

The quality of evaluations is mixed. This may improve with the planned establishment of new quality standards by OEDE in the near future. Particular weaknesses have been noted in methodology and in some crosscutting issues such as human rights, where WFP as whole lacks capacity and guidance. A failure to take into account the cost implications of recommendations, together with factors related to the nature of priority setting in WFP has damaged the credibility of evaluations among some WFP staff.

Partners are engaged in evaluation primarily as either hosts or key informants. The Panel views this as inappropriate as a basis for encouraging two-way accountability and learning. The extent to which stakeholders are consulted in the planning, design, conduct and follow-up of evaluations is patchy and overly concentrated on the implementation stage.

Utility

The Panel concludes that the criteria of *Utility* of the WFP evaluation function are partially met in that it provides a major contribution to the formulation of follow-up programmes, but that structures and mechanisms to promote utility are weak in most other respects. There is insufficient corporate strategic thinking regarding the use of evaluation and where evaluation should fit on the learning-accountability spectrum. Evaluation is rarely used to enhance accountability to partners.

Evaluation is insufficiently integrated into many of the processes by which WFP sets, monitors and analyses policies. WFP rarely uses evaluation as a way of encouraging broader accountabilities in respect to its position within the humanitarian and development communities.

Evaluation is primarily focused on outputs, as opposed to outcomes and impact, which reflects the demands of many stakeholders for practical programming advice. In a narrow sense of contributing to an understanding of how to ‘do things right’, evaluation makes a notable contribution to programme design and management. In a wider perspective of learning about ‘doing the right thing’, performance is not so good. Despite some efforts within evaluations to present evidence that can stimulate greater reflection within WFP over the changing role of food aid, for example, the corporate view of evaluation has tended to focus primarily on its utility for making modest adjustments to existing approaches. A notable exception to this is the interests of the Executive Board which seeks to use evaluation as a major input into their deliberations on the future role and corporate priorities of WFP.

Structures for management response to evaluations at headquarters level are fragmented and require a major overhaul. Management responds to evaluation through the process of programme formulation, but in a broader sense this is weak. New plans in OEDE to tie evaluation closely to logical frameworks may enhance utility through a focus on outcomes, but lack of prevailing understanding and use of logical frameworks within WFP will make this difficult. These problems ultimately stem from the fact that respect for the role of evaluation and a readiness to acknowledge mistakes are reportedly weak at WFP.

The intention to use evaluations is evident as programme evaluations are timed to coincide with formulation of new phases. At decentralised levels there is a close link to utility since there is a direct desire to use evaluation to inform and justify new programmes and phases. Local flexible solutions are often found to use evaluation within a broader learning process. A negative aspect of this may be that the accountability aspect of evaluation, whereby an independent assessment may question the overall relevance of country priorities, may be overshadowed by a concentration on feeding into the design of attractive programming based on documented success.

Evaluation makes an inadequate contribution to overall knowledge building within WFP and virtually none among partners. Access to reports and findings through the website, debriefings, etc., is acceptable, but promotion of the use of evaluation products is not sufficiently proactive. Evaluation is not well integrated into other learning initiatives within WFP. Opportunities are missed to distil and disseminate learning systematically. Evaluation does make a major contribution to building knowledge within the Executive Board, particularly through thematic/strategic evaluations. There is a demand for greater efforts to present syntheses of evaluations conducted. At decentralised levels evaluation is more directly related to programme formulation and learning is somewhat more effective.

Not enough effort has been made to find opportunities to build local capacities. Some deficiencies noted regarding utility relate to the need to rethink stakeholder engagement throughout the evaluation process so as to ensure broader buy-in. The quality of utility is directly related to the ability of WFP to apply findings and recommendations together with partners, but this aspect of the evaluation function is generally very weak.

Recommendations

The Peer Panel's recommendations are intended to suggest ways to improve and amend aspects of WFP's evaluation function where the Peer Panel has identified problems or shortcomings in relation to UNEG's norms and standards or to established evaluation practice.

Evaluation policy

OEDE should develop an evaluation policy that encapsulates the previous evaluation policies and fully meets all UNEG norms and standards for evaluation. It should also be designed as a transparent vehicle for promoting greater communication among internal and external stakeholders regarding the aims and intended utility of evaluations. The role and purpose of and the relationship between self-evaluation, decentralised evaluations, and external evaluation should be studied and clearly articulated not only in the evaluation policy, but in the overarching policy and strategy documents of WFP. The mandate for OEDE should form part of WFP's evaluation policy.

Wider accountability

A key challenge for WFP's accountability is to expand its institutional accountability to include intended beneficiaries via host government and/or NGO Cooperating Partners. OEDE should develop an 'accountability map' of key WFP stakeholders, both internal and external, to help in clarifying roles and responsibilities. Based on this 'accountability map', guidance should be provided both for WFP staff in func-

tional units and partners identified in the map as key stakeholders to help them in fulfilling their accountability responsibilities.

Participation in evaluation

OEDE should look for ways of promoting, and providing incentives for staff to adopt more participatory approaches in evaluations. Engagement with partners at country, regional or global levels is primarily a responsibility of other parts of WFP. Nonetheless, OEDE should play a more proactive role in promoting substantive involvement of relevant stakeholders. The responsibility of the OEDE evaluation manager and evaluation team leader is to make optimal use of participatory systems already in place during the planning, implementation and utilisation of the evaluation. OEDE has a role to play in building capacity of WFP field staff to ensure that WFP field staff are provided with appropriate support and guidelines to facilitate participatory approaches during evaluation processes.

Management response to evaluations

The lines of responsibility for management response are currently blurred. WFP should, both in principle and in practice, establish a clear division of responsibility regarding management response between the evaluation function and the organisation's line management. After an evaluation has been submitted to the Executive Director, OEDE should not be involved with drafting or compilation of responses from different parts of the organisation. The management response mechanism should include rules about the timeframe for the response and procedures for follow-up of the management response as well as for reporting to the Executive Board and informing the OEDE about the results of the follow-up.

A similar system for management response should be used for decentralised evaluations. The same kind of division of responsibilities cannot be established when a country director both commissions an evaluation and decides on management response, for example. However, when an evaluation is carried out by external evaluators a country director can still independently formulate a management response and report to the Regional Bureau director or the Executive Director. Follow-up should also be the Country Office's responsibility with reporting on the results upwards within WFP. Management response and follow-up mechanism should be transparent with relevant documents easily accessible for WFP and partners and routinely posted in electronic form.

Quality of evaluation

The increasing emphasis on improving the quality, rigour and harmonisation of OEDE's work, as well as the focus on systematic processes, quality checks and tools such as stakeholder maps and evaluation matrices are highly encouraged. The capacity of OEDE staff should be maintained over time to stimulate interest in the evaluation field and encourage professionalism. We recommend that ample time should be allocated and incentives should be provided for staff to keep up with new developments in the field of evaluation.

Mechanisms should be found to improve the quality, credibility and ownership of evaluation recommendations. Such mechanisms may include developing recommendations in dialogue with primary stakeholders, and/or leaving recommendations up to those responsible for decisions and action in WFP, based upon engagement by primary stakeholders around the findings and conclusions of the evaluation report.

Learning and accountability

The tension and complementarities between evaluation for accountability and for learning seem not to be acknowledged everywhere in WFP. This requires more attention. Ideally the learning element should be linked to a larger organisational knowledge management strategy.

OEDE should continue recent efforts to systematically harvest lessons from existing evaluations as well as external fora such as ALNAP, the IASC and relevant partners. Innovative methods for extracting and sharing of evaluation lessons should be investigated, building on the experiences of other organisations with extensive experience in this field.

Results-based management

WFP should give high priority to address the fragmentation among its various results-focused data collection, reporting and analysis tools. A thorough review of existing field monitoring systems and applications is vital to ensure that evaluations as well as the corporate monitoring system have access to more reliable, relevant and comparable data. Ways should be developed and maintained to ensure that all interventions are linked to proper monitoring mechanisms, both at local and corporate levels. OEDE should make an evaluation of WFP's approach to results-based management a high priority for its future strategic evaluations, building on the recent work done by internal audit on results-based reporting.

Selection of evaluation teams

It is recommended that OEDE develops a more transparent, rigorous and competitive approach to the selection of team leaders. If possible, team leaders should be identified early on and be involved in the identification and selection of the rest of the team. All evaluation teams should include at least one evaluation specialist, preferably the team leader, who has sufficient knowledge about and experience from current evaluation approaches and methods.

Staffing and budgeting of OEDE

The Panel considers the current mix of internal WFP career staff and externally recruited professional evaluators suitable. However, given the technical nature and professional skills profile of these positions, the Panel recommends WFP to allow OEDE to select internal staff based on a professional recruitment process rather than through the standard reassignment exercise. WFP should continue to (a) allow external recruitment of evaluation specialists; (b) base OEDE's staff profile on the profile of evaluators developed by UNEG; and (c) consider how to ensure an appropriate career path for evaluation specialists within WFP and within the UN system.

To address concerns that only a small portion of the overall evaluation budget is within the direct control of OEDE, WFP's senior management should devise ways to safeguard the funding allocated to evaluations for the next biennium. The use of Direct Support Cost of projects appears a budgetary necessity for the time being but it is not an ideal situation. It is critical that the Executive Director and senior management ensure the full use of these funds for evaluation by holding managers accountable for following implementation plans. Furthermore, management should consider 'earmarking' strategic and sensitive decentralised evaluations for OEDE-management to thus reduce the risk that offices seek to bypass OEDE and to ensure full independence where most needed. The establishment of a centrally managed fund for both OEDE evaluations and decentralised evaluations should be investigated.