

برنامج
الأغذية
العالمي



Programme
Alimentaire
Mondial

World
Food
Programme

Programa
Mundial
de Alimentos

**Executive Board
Second Regular Session**

Rome, 27–28 May 2004

EVALUATION REPORTS

Agenda item 2

*For information**



Distribution: GENERAL
WFP/EB.2/2004/2-B
14 April 2004
ORIGINAL: ENGLISH

REVIEW OF WFP'S EXPERIENCE WITH REAL-TIME EVALUATION

* In accordance with the Executive Board's decisions on governance, approved at the Annual and Third Regular Sessions, 2000, items for information should not be discussed unless a Board member specifically requests it, well in advance of the meeting, and the Chair accepts the request on the grounds that it is a proper use of the Board's time.

This document is printed in a limited number of copies. Executive Board documents are available on WFP's WEB site (<http://www.wfp.org/eb>).

NOTE TO THE EXECUTIVE BOARD

This document is submitted to the Executive Board for information.

The Secretariat invites members of the Board who may have questions of a technical nature with regard to this document to contact the WFP staff focal points indicated below, preferably well in advance of the Board's meeting.

Director, Office of Evaluation (OEDE): Mr K. Tuinenburg tel.: 066513-2252

Task Manager, OEDE: Ms P. Hougesen tel.: 066513-2033

Should you have any questions regarding matters of dispatch of documentation for the Executive Board, please contact the Supervisor, Meeting Servicing and Distribution Unit (tel.: 066513-2328).



EXECUTIVE SUMMARY

This document provides a summary of the review of WFP's first experience with real-time evaluation of emergency operations, as requested by the Office of Evaluation at the Third Regular Session of the Executive Board in 2003.

Real-time evaluation differs from ordinary *ex-post* evaluation in that it is conducted simultaneously with implementation of the emergency operation. WFP used the real-time evaluation approach for the first time with EMOP 10200 for the southern Africa food crisis from July 2002 to June 2003.

A review was commissioned to assess the usefulness to stakeholders of real-time evaluation and to inform any future use of the approach. The review studied the Office of Evaluation and Monitoring documents, correspondence and background materials relating to real-time evaluation in general and the real-time evaluation of EMOP 10200 in particular. Stakeholder groups were interviewed to obtain their perceptions of real-time evaluation and their views on its future use.

Stakeholders see real-time evaluation as potentially useful in improving quality by (i) solving operational problems as they occur, (ii) enabling organizational learning for improvement of future EMOPs and (iii) providing an independent assessment of results. Stakeholders' suggested changes to real-time evaluation have been incorporated in the review's recommendations. WFP has accepted and will implement the recommendations.

DRAFT DECISION*

The Board takes note of the information in document WFP/EB.2/2004/2-B on WFP's review of experience with real-time evaluation, including the Programme's acceptance of the review's recommendations, which it will implement in its continued use of real-time evaluation of emergency operations.

* This is a draft decision. For the final decision adopted by the Board, please refer to the Decisions and Recommendations document issued at the end of the session.



REAL-TIME EVALUATION

1. By definition, “evaluation” means to appraise the value of something, for example the degree of success achieved in an activity, but evaluation of assistance activities has increasingly shifted towards promotion of learning, with less emphasis on accountability. Most development and emergency-assistance organizations use their evaluation units to assess their results for accountability purposes and to improve the quality of their work by learning from experience and building staff skills. There is a continuing debate among evaluators as to whether a single evaluation could address both learning and accountability, given that one requires a confidential and collegial approach and the other is public and independent.
2. Evaluations are traditionally conducted on completion of a project and usually concentrate on (i) assessing relevance, effectiveness, efficiency and sustainability for accountability purposes and (ii) identifying lessons to be learned to improve future performance. They rely heavily on results data, records and discussions with people present during implementation — but these sources of data are often not available after emergency operations (EMOPs), because staff are redeployed, monitoring systems collect incomplete data and records do not provide the context of decision-making. Evaluations can therefore be disappointing in terms of assessment of results and facilitating learning; those that do identify lessons to be learned cannot feed them back to the country office staff in time for them to be useful. The learning benefit is therefore longer-term, not operation-specific.
3. Real-time evaluation (RTE) is an attempt to overcome these impediments and thereby enhance the quality of current and future operations and provide valid performance information to stakeholders. It differs from *ex-post* evaluation principally in that (i) evaluation takes place during implementation of an operation and (ii) it often entails more than one field mission to cover different phases of the operation. These differences have implications for the profiles of evaluators, the number of field missions, methods of work, costs, and the nature and readership of reports.
4. RTE has not been used extensively to evaluate assistance activities and has been employed only recently to evaluate emergency assistance. It was introduced by the World Bank in 1996 when a quality-assurance group (QAG) was created with the primary objective of promoting operational excellence by increasing accountability and enhancing learning. The QAG strategy involves two complementary approaches: (i) providing staff and managers with real-time feedback on operational performance and identifying systemic issues in a transparent fashion, and (ii) identifying and highlighting skills and resources needed for high-quality work, using lessons learned to support training. The QAG reports directly to the managing director of operations; its reports are not made public. Its staff are experienced current and former World Bank personnel and external consultants.
5. RTE was introduced in the Office of the United Nations High Commissioner for Refugees (UNHCR) in 2000 to improve the quality of its emergency work and has been used in at least three operations. UNHCR has carried out RTE for quality improvement through learning about the operation being evaluated and developing staff skills using experienced personnel from its evaluation and policy analysis unit in teams of one or two; enquiries during field missions have been somewhat unstructured. RTE has been reported internally for accountability to senior managers; some reports have also been published on



the UNHCR website. UNHCR appears satisfied that RTE has promoted some immediate learning during operations.

6. A number of non-governmental organizations (NGOs) have employed RTE to examine their operations. The practice appears to be starting amongst bilateral donors and other organizations. The United Nations Children's Fund (UNICEF) reviewed the RTE experiences of other organizations in 2003.
7. OEDE studied the literature and commissioned a consultant evaluator in 2001 to produce an RTE conceptual framework that would enable it to strengthen its work in EMOPS and in particular to promote learning.
8. The framework produced contains most of the elements of a management review and a traditional evaluation, and also takes advantage of simultaneous implementation and evaluation to observe and record operational processes, which enables assessment of the applicability and usefulness of WFP's operational guidelines and generates advice to the EMOP team for immediate action. The framework was fairly closely adhered to in preparing the RTE.

WFP's RTE

9. WFP's experience with RTE stems from evaluation of EMOP 10200 in southern Africa, which ran from July 2002 to June 2003 and delivered nearly 1 million mt of food to six countries at a cost of US\$500 million.
10. RTE was undertaken by a team of two external consultants and a current WFP staff member, augmented by two former WFP staff members, one in each of the second and third field missions. The evaluation manager from OEDE joined as a full mission member for the third and final mission.
11. Preparatory studies were undertaken as soon as the EMOP was approved; the three field missions were timed to coincide with the establishment phase, implementation phase and closing phase of the operation.
12. The initial broad terms of reference (TOR) were narrowed during the operation to focus on important issues. They directed the RTE to promote operation-specific and institutional learning during implementation and to assess results.
13. Brief aides-mémoire were presented to staff at the end of each mission before the team left the field; internal reports were produced after the first two missions. The final report based on the findings from the three missions was presented to the Executive Board in October 2003.
14. The RTE environment was difficult: EMOP 10200 was a complex regional emergency covering six countries; many of the individuals concerned had different perceptions of what RTE was and how it should operate.

Discussion of Review Interview Data

15. The perceptions of people interviewed for this review are summarized below. Each heading combines a number of the questions.

Expectations and Purposes

16. There was no common view of the purpose of RTE. The TOR allowed the team to determine its main purpose and the issues it could most usefully address once it had visited



the field; the TOR for each field mission narrowed the focus. Team members did not all agree on the priority for learning rather than accountability and hence on methods of work. Field staff, on the other hand, were almost all expecting assistance in implementing the operation and in identifying and resolving difficulties.

17. In these circumstances it is not surprising that most senior and middle-level field staff felt that RTE failed to meet their expectations. The team was not equipped to provide operational solutions, which require detailed knowledge of WFP systems and procedures. The limited field time reduced the team's capacity to go into detail.
18. RTE was more successful with regard to institutional learning, though many field staff regarded the problems identified and brought to the attention of Headquarters units as attributable to their own unrecognized efforts; field staff undoubtedly identified most if not all of the institutional constraints appearing in the RTE reports. It is unclear to what extent the RTE contributed to prompting timely action to resolve them, but its role should not be dismissed. There might have been a more positive perception of the role of the RTE if its reports had acknowledged the part played by staff in identifying issues; the team feared it could have caused difficulties for some staff. An Executive Board member said that the institutional learning was the most important outcome from his point of view.
19. Most interviewees said that RTE should be used in future only for evaluation of large complex EMOPs and those with novel features that could cause problems. Two interviewees suggested that the next use of RTE should be in an EMOP when the regional director concerned invited OEDE to mount the exercise. This could provide a positive environment in which RTE may be further refined.
20. The TOR called for identification of positive attributes of the operation that could inform learning. The reports do mention several very successful aspects, but they are not the subject of any recommendations. This is unfortunate because highlighting success and aspects that should be repeated or institutionalized is part of creating a positive view of evaluation and a willingness to learn from it. A positive mention in the report is good, but showing how it can be used in other operations is more useful feedback.

The Value of RTE

21. Senior and middle-level field managers of EMOP 10200 mostly felt that RTE had failed to produce operation-specific learning and that institutional learning had been the result of their own analysis. Some junior field staff reported that RTE helped in technical learning and increasing morale. Headquarters staff were divided, but most believed that RTE had produced valuable institutional learning, some of which had already been applied to subsequent operations. Several interviewees saw value in RTE as a form of management review and as a means of ensuring compliance with policy and procedures in operations.
22. Executive Board members interviewed said that the use of RTE reflected a commitment to learning in WFP and that they valued the evaluation report, which was substantive and balanced. RTE increased donor confidence in WFP generally, but had no effect on contributions.
23. All interviewees, except two field staff and one Headquarters staff member, thought that RTE had potential, whether or not it satisfied their expectations. Views as to its principal value were divided almost equally between:
 - operation-specific learning — quality improvement and immediate feedback on weaknesses;
 - institutional learning — systems, procedures and policies; and



- accountability for results to senior managers, donors and the Executive Board.
24. All except the two field staff thought that RTE offered advantages over *ex-post* evaluation in terms of benefits such as an outsider's view of an operation as it unfolds, better understanding of result assessment, opportunity for action reviews and the ability to take corrective action immediately.
 25. Many field staff thought that interim RTE reports could be submitted to donors and the Executive Board; others believed that only a final report should be made public so that frank discussions between the evaluation team and field staff were not inhibited. Most interviewees thought that interim reports should be WFP internal documents used to monitor developments and initiate institutional learning.
 26. The modest cost of RTE relative to the total cost of the operation led almost all the interviewees to conclude that RTE that fulfilled its potential would be cost-effective.

Team Composition

27. An RTE team must clearly have the knowledge, skills and experience to achieve its purpose. If a unit selects the team to serve one purpose and other units have a different purpose in mind, there is unlikely to be agreement on the suitability of the team, as was the case in RTE of EMOP 10200. All stakeholder groups agreed, however, on the essential characteristics of future teams.
28. Most stakeholder interviewees believed that future teams should be led by and include external consultants. External consultants provide an outside view, enhance credibility for accountability assessments and are more likely to introduce new ideas; they must be familiar with current WFP operational practice and systems if the team is to provide informed advice for operational and institutional learning.
29. WFP staff members of RTE teams who come from regions other than that of the EMOP and from Headquarters units are essential to provide institutional understanding and facilitate relations between the RTE team and field staff.
30. Teams that are less familiar with WFP may be able to identify problem areas, but may be unable to define a particular problem and generate practical solutions.
31. Some field staff believed that RTE teams need sound evaluation experience as well as technical experience.
32. Field staff were particularly concerned that all RTE team members should have good interpersonal skills that enable them to interact positively with field staff at all levels. Opinions vary as to whether this requires WFP team members to be peers or more senior staff. Seniority usually implies long experience and credibility, which are positive attributes, but it may inhibit establishment of close working relationships with junior staff.
33. There may be occasions when a team should adopt a provocative and demanding approach, but identifying them requires considerable skill and judgment.
34. All WFP interviewees believed it was necessary to restrict RTE teams to three or four members, even though it would limit the range of areas of expertise that could be included and hence the scope of an evaluation.

Methods of Work

35. Most interviewees thought that three field missions were necessary for RTE, but some senior managers, OEDE staff and field staff considered that two would be sufficient; the timing of the missions was their main concern.



36. Most interviewees, particularly field staff, believed that the first field mission should take place only after one to six months, after the establishment phase; they thought that a minimum of six weeks was required to establish an operation and for staff to be available for RTE. Some suggested that this would remove an opportunity to assist or observe the establishment phase, when RTE involvement might be of value.
37. Some field staff were concerned about what they saw as the unsystematic way in which the RTE team had approached selection of issues and collection of data. An informal approach in which teams and field staff work together to identify issues and organize systematic data collection is probably appropriate for initial missions to determine the skills required for the subsequent mission. Systematic investigation of focus issues would make some field staff more confident of RTE.
38. A team member and an OEDE interviewee proposed that the first field mission should be conducted by generalists — possibly the team leader and the OEDE task manager — who would design subsequent missions and identify the skills and experience needed.
39. Some field staff suggested that RTE teams should rely on data generated by the EMOP, thus reducing the burden on implementing staff.
40. Many interviewees thought that the burden of checking reports and aides-mémoire was too great and that the process delayed feedback for action. Most Executive Board members interviewed thought the final summary report was useful; one saw a disadvantage in the lack of interim reports and of general lessons for WFP as a whole.

CONCLUSION AND RECOMMENDATIONS

Conclusion

41. The first RTE was in many respects successful and confirmed that the approach should be of substantial value as a tool for learning and accountability in EMOPs: (i) it assisted with some operational learning, at least for some junior staff, (ii) it made a useful contribution to institutional learning and informed WFP at the policy level and (iii) it provided balanced and substantive accountability information, which helped to increase Executive Board confidence in WFP. Various steps could be taken, however, to adjust the RTE model to make it more useful to WFP in the future.

Recommendations

- a) The scope and purpose of RTEs need to be clearly defined and manageable. WFP will need to decide which purpose has priority, because no RTE can be expected to achieve everything that the tool can offer. This may be done on a case-by-case basis, or by establishing as RTE policy a priority, for example, for specific operational learning, with improved accountability as a by-product. It is doubtful that the same team could achieve effective learning and accountability evaluation. Using two RTE teams should be considered for EMOPs, in which both learning and accountability are important: team A would undertake the first and second field missions and emphasise learning; team B would undertake the third mission, using the information collected by team A as an input to a summative evaluation for accountability. Evaluation of a single operation will not form the basis of major WFP policy change; a number of evaluations indicating similar needs would be required.



- b) RTE teams must have the right composition in terms of size, skills and experience to achieve their purpose. Operational learning requires a team that has sound understanding of current WFP systems, procedures and practice and excellent participatory evaluation skills. All teams should be familiar with evaluation methods of data collection and analysis and have good interpersonal skills. Team members should be experienced people with strong management skills. It will take time to identify teams with all the necessary attributes; it may be helpful if OEDE created a list of screened candidates to facilitate team selection when the need arises.
- c) It would be helpful to try the next RTE in an EMOP where the regional director or country director invites its involvement. To this end, all regional and country directors could be provided with the full report of the RTE and the review report and asked to consider inviting RTE of a future EMOP in their area.
- d) OEDE should decide case by case in consultation with OD which EMOPs should be evaluated by RTE. The main criterion should be the extent to which WFP would benefit in terms of learning or accountability. It may be appropriate to use RTE only in large complex EMOPs, or in those with some novel feature. The decision should be made as soon as the EMOP design process begins, to allow the RTE team to observe and start gathering information as early as possible. The three-mission model is appropriate for most EMOPs to be evaluated using RTE; the first field mission should not be undertaken in the first six weeks of implementation unless OEDE and OD conclude that there is particular value in RTE during the initial phase of an EMOP.
- e) Regional and country directors responsible for EMOPs should be kept informed about the preparation, purpose and scope of RTE; their views should be taken into consideration. OEDE should be the final decision-maker, however. Regional and country directors should ensure that field staff are kept fully informed about RTE and their role in supporting the team.
- f) RTE methods should include participatory approaches such as action reviews and SWOT analyses and should involve as many stakeholder groups as possible. The team should acknowledge verbally and in written reports any problems or solutions identified by field staff.
- g) Teams and OEDE should continue to use what they perceive as the best methods of collecting and analysing data to achieve their TOR objectives, but they must be selected with the context in mind; RTE team members and staff must be kept informed about purposes and methods as they evolve.
- h) The first field mission may need to be an exploratory mission by the OEDE task manager and the team leader, or a senior member of OD, or an experienced ex-WFP staff member. This will enable data to be gathered and issues identified to guide the composition and approaches of subsequent teams.
- i) Interim reports should be WFP internal documents unless the purpose of a particular RTE is to provide public interim accountability reports. Interim reports should be in the form of aides-mémoire prepared towards the conclusion of each field mission, discussed with the responsible offices and finalized to incorporate their views before the team leaves the field. These reports should be circulated to OEDE, senior management in OD and to WFP units affected by their recommendations. A final report should be written at the conclusion of RTE and circulated as for *ex-post* evaluation reports. Interim reports to the Executive Board, donors and the public, if required, could be prepared by WFP staff on the basis of information from RTE reports.



- j) OEDE should explore means to improve dissemination of RTE findings so that learning from them can be maximized. Two possibilities are better links with staff training programmes and more creative use of information media.
- k) OEDE should investigate ways in which a positive evaluation culture can be established in WFP. RTE is a significant tool in this: it can provide practical and immediately relevant feedback and ideas to country office staff, which will bring about recognition of evaluation as a helpful service; this will be enhanced if RTE focuses on looking for positive features of EMOPs. This recognition needs to extend to RTE recommendations for institutional learning wherever possible: field staff and others will be more likely to learn from RTE if their successes are considered as worthy of replication as their failings are worthy of note.



WFP RESPONSE TO THE RECOMMENDATIONS

OD and OEDE agree that RTE is a promising new instrument for improving real-time performance and for operational and institutional learning. They accept the findings and recommendations of the RTE review, including those concerning the necessary prior consultations between OEDE and OD to clarify expectations among those involved. They would like to add that evaluation work is not limited to field missions, and that fact-finding could start at the onset of a new emergency, well before an EMOP becomes operational.

Regarding improved dissemination of results, it is the intention of OEDE to devise a mechanism to communicate lessons to be learned to a wider audience and include them in staff training. This is not restricted to RTE, but will apply to all evaluation work carried out by OEDE, including lessons learned from decentralized evaluations.



ACRONYMS USED IN THE DOCUMENT

EMOP	emergency operation
NGO	non-governmental organization
QAG	quality-assurance group
RTE	real-time evaluation
SWOT	strengths, weaknesses, opportunities and threats
TOR	terms of reference
UNHCR	Office of the United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund

