Guidelines for Project and Programme Evaluations

Final Version
July 2009

www.entwicklung.at

Austrian Development Cooperation

Imprint:
Austrian Development Agency
The Operational Unit of the Austrian Development Cooperation
Zelinkagasse 2, 1010 Vienna, Austria
Phone: +43 (0)1 90399-0
Fax: +43 (0)1 90399-1290
office@ada.gv.at
www.entwicklung.at

These Guidelines were written by:
Austrian Development Agency, Evaluation Unit

Vienna, October 2008

English Translation, Vienna, July 2009

Content

1. Introduction

The Guidelines are intended to support project partners which implement projects or programmes supported by the Austrian Development Cooperation (ADC) during the process of planning, commissioning and managing project and programme evaluations. Furthermore, it addresses also those officers of the Austrian Development Agency (ADA) and the Coordination Offices (COs) that are in charge of contracting and coordinating project or programme evaluations[1].

The Guidelines delineate the **administrative processes,** which need to be applied if the costs of project-/programme evaluations are included in the approved budget, or if project or programme evaluations are commissioned by ADA headquarters officers or COs.

ADA uses the internationally approved definitions for Monitoring, Evaluation and Review, which correspond to the OECD/DAC Glossary[2]:

> ***Evaluation:*** *The systematic and objective assessment of an on-going or completed project or programme, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. (OECD DAC Glossary).*
>
> ***Monitoring:*** *A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds. (OECD DAC Glossary).*
>
> ***Review:*** *An assessment of the performance of an intervention, periodically or on an ad hoc basis. Reviews are usually less comprehensive and/or in-depth than evaluations. They tend to emphasize operational aspects. (OECD DAC Glossary).*

Evaluations are generally conducted by independent, external experts. In general, an evaluation analyses complex issues and captures intended and unintended effects. ***Evaluations investigate the reasons why certain aspects of a project or programme have or have not been implemented as planned.***

Evaluations are carried out either during the project cycle (Mid-term Evaluation, Formative Evaluation) or at the end of a project or programme (Ex-post Evaluation, Final Evaluation, Impact Evaluation, Summative Evaluation)[3].

[1] This document is not applicable for evaluations directly commissioned by the ADA Evaluation Unit (particularly “strategic evaluations“) or projects or programmes based on a direct agreement of ADA with a foreign state or a foreign public institution (National Execution). Likewise, it is not to be applied for Impact Assessments or Impact Evaluations.

[2] See also Annex 7.1.

[3] Definitions are illustrated in the Annex.

Evaluations have the following characteristics:

- the five OECD/DAC evaluation criteria of Relevance, Effectiveness, Efficiency, Impact and Sustainability are covered,
- cross-cutting issues, such as poverty, gender and environment, are taken into consideration4,
- and the intervention logic (e.g. Logframe) is analysed.

Additional criteria may also be added such as “Participation“ or “Responsibility“. The five OECD/DAC criteria are defined as follows:

Relevance	Are we doing the right thing? How important is the **relevance or significance** of the intervention regarding local and national requirements and priorities?
Effectiveness	Are the objectives of the development interventions being achieved? How big is the **effectiveness or impact** of the project compared to the objectives planned (Comparison: result – planning)?
Efficiency	Are the objectives being achieved economically by the development intervention? How big is the **efficiency or utilisation ratio** of the resources used (Comparison: resources applied – results)?
Impact	Does the development intervention contribute to reaching higher level development objectives (preferably, overall objective)? What is the **impact or effect** of the intervention in proportion to the overall situation of the target group or those effected?
Sustainability	Are the positive effects or impacts sustainable? How is the **sustainability or permanence** of the intervention and its effects to be assessed[5]?

In Annex 7.2 and 7.3, these five criteria are illustrated in more detail.

If it is not possible to consider all <u>five</u> OECD/DAC criteria, this needs to be justified in the Terms of Reference (ToR).

In <u>all</u> evaluations, the <u>evaluation questions</u> must be formulated in accordance with the OECD/DAC evaluation criteria of Relevance, Effectiveness, Efficiency, Impact and Sustainability.

ADA follows the international trend and defines those evaluations that are managed by the project partners themselves as "internal evaluations" (even if external experts are engaged). Those evaluations that are managed by ADA (Headquarters, Coordination Offices) are considered as "external evaluations", as they are not subordinate to the direct project management.

[4] Further details are to be obtained from the Annex.

[5] Further clarifications are to be obtained from the Annex.

Consequently, ADA distinguishes between three types of evaluations:

External Evaluation Type I: Evaluations of programmes, cross-cutting issues, instruments and projects commissioned and managed by the ADA Evaluation Unit.

External Evaluation Type II: Evaluations of programmes, projects and topics commissioned and managed by the ADA Coordination Office or an ADA Unit Headquarters.

Internal Evaluation: Evaluations commissioned, managed and/or implemented by project partners themselves (with or without external experts).

As far as the budget is concerned external evaluations of Type I are assigned to the ADA Evaluation Unit. External evaluations of Type II are assigned to the budget line, from which the respective project is financed. The costs for internal evaluations need to be considered in the project or programme budget.

Every project or programme must be evaluated once internally within the project or programme cycle. Additionally, projects or programmes can also be evaluated externally. In a particular case and in the event of an external evaluation being planned, an internal evaluation can be abstained from. In any case, the costs of an evaluation must be in an appropriate proportion to the scope of the project. Consequently, an evaluation is rather unlikely in case of small projects (see also Chapter 5.). In general terms, approximately 3–5 % of the entire project or programme budget is provided for an evaluation.

2. Purpose of reviews

In comparison with project and programme evaluations, reviews are less complex analyses. Neither do they necessarily contain an analysis according to the five OECD/DAC criteria nor must they be responsive to cross-cutting issues. However, it is also recommended to analyse the intervention logic (e.g. Logframe) in each review.

Reviews are appropriate if the project or the programme is analysed, in order to, e.g.

- interpret already existing results,
- work out lessons learnt together, or
- develop future strategies, which result from lessons learnt.

ADA supports reviews managed by the project partner and recommends that the latter be carried out together with external experts or moderators (facilitators)[6]. Depending on their focus and method, reviews are also often referred to as “internal evaluations“, “self-evaluations“ or “participatory evaluations“.

The project partner requires ADA (CO/Unit) to approve of the ToR. Beyond that, however, ADA (CO/Unit) is not involved in the planning or coordination process of a review.

[6] According to our experience it needs to be pointed out that numerous planned evaluations have actually turned out to be reviews.

The project partner has to add the review results of a review to the annual or final report, as well as an analysis of the utilisation of these new findings. However, the documents can also be forwarded to ADA immediately upon completion.

Since there are no international standards for reviews available, the principles and standards for evaluations need to be applied accordingly.

Regular progress reports or monitoring discussions and meetings which serve e.g. the preparation of the annual report, are not classified and recognised as reviews.

If required, ADA can also commission reviews.

Given the different scope of reviews and evaluations, it is expected that the complexity and costs of a review will be much lower than those of an evaluation.

3. Purpose of project and programme evaluations

It is a strategic goal of ADA to enshrine project and programme evaluations in a comprehensive manner in the project cycle management. Therefore evaluations need to be included in the project document. Evaluations contribute to secure the optimal quality and impact of development interventions. They also help managers of projects and programmes to manage and improve their implementation.

The purpose of evaluations is:

- Learning from experience: With the assistance of evaluations, successes and failures can be interpreted. Based on those experiences, both current and future projects and programmes can be improved.
- Transparency: Evaluations illustrate the responsible utilization of the resources and justify the results and their effects vis-à-vis the contractor, the partners, the target groups in the recipient country and the tax payers.
- Deepening understanding: Evaluation is a tool for deepening knowledge and understanding of the assumptions, options and limits of development cooperation (DC). Evaluations are intended to contribute to a comprehensive discussion and reflexion about development cooperation.
- Improved communication: An evaluation is intended to foster communication and understanding within and between the groups mentioned above, even if this can only be managed in different ways and with different participations in each case.

4. International evaluation principles and standards

Important principles of every evaluation are:

- **Objectivity**: In its conception and implementation every evaluation needs to achieve a maximum level of objectivity and impartiality. Statement of facts needs to be methodically clearly distinguished from assessments. It is important that different perspectives are taken into account, as well as strengths and weaknesses. Results, conclusions and recommendations need to be supported by evidence and must be comprehensible.
- **Independence of evaluators**: The evaluators must have expert knowledge. Credibility also includes the independence of evaluators from all staff involved operatively.
- **Participation of all parties concerned in the entire process**: An evaluation needs to be as participatory as possible (e.g. developing jointly the Terms of Reference with the project partners, the possibility of all parties involved to comment on the results or the evaluation report).
- **Transparency and Focus**: The evaluation assignment must be clearly defined and focussed: Description of the initial situation (project programme details), objectives of the evaluation, central questions, methodologies, qualifications of the evaluation team, reporting requirements. In most cases, an evaluation cannot cover all OECD/DAC criteria to the same extent, hence a strict definition of priorities is essential.
- **Reliability**: The utilisation and preparation of basic data is necessary in order to prove the assessment and the conclusions in a credible fashion. The evaluation results stated in the evaluation report must be comprehensible.
- **Completeness and clarity of reports**: The report has to be structured according to the OECD/DAC criteria and evaluation questions. All evaluation questions must be answered.
- **Fairness and protection of the interests of the parties involved**: Security, dignity and rights of the persons involved in the evaluation must be protected.
- **Utility**: Evaluation recommendations are used for improving projects or programmes. Feedback to political and operative decision makers must be guaranteed through a clear responsibility for the implementation of the evaluation results.

The utility of an evaluation also depends on the most appropriate moment of the implementation within the project or programme cycle. Furthermore, the expenditure of time and financial means of an evaluation need to be in a reasonable proportion to its benefits and the scope of the project or programme (Internationally about 3–5 % of the total project or programme budget is normally spent on evaluations).

During the **evaluation** the evaluation team must consider the OECD/DAC evaluation quality standards and/or the ones of the Gesellschaft für Evaluation (DeGeval). The familiarity with these standards is absolutely necessary to ensure that evaluations meet the international standards and requirements[7].

[7] See Annex and DAC Evaluation Network. DAC Evaluation Quality Standards. OECD. March 2007, http://www.oecd.org/site/0,3407,en_21571361_34047972_1_1_1_1_1,00.html
Gesellschaft für Evaluation (DeGEval). Standards für Evaluierung. Köln. Juni 2004, http://www.degeval.de/

What does the mandate of an evaluation clearly exceed?

- An evaluation itself cannot decide whether or not a project or programme should continue. Evaluations are solely commissioned to make recommendations and to present observations.
- An evaluation is not to be used for the justification of decisions (e.g. to depict a project or programme as being better than it is, or to terminate it immediately).
- An evaluation is not an instrument of crisis management.

For all project or programme evaluations it applies that the **coordination process** must meet the international evaluation standards and the international quality criteria.

Generally, managing an evaluation the following tasks need to be carried out:

- Preparation of the Terms of Reference (participatory process with partners)
- Preparation of a budget plan
- Selection of the evaluation team
- Support of the evaluation team (compilation of documents, possible logistical support, etc.)
- Organisation of an introductory workshop
- Organisation of a workshop for the presentation of the inception report
- Quality control approving the inception report
- Organisation of a workshop for the presentation of the final draft evaluation report
- Quality control approving the final draft evaluation report
- Quality control approving the final report
- Approving the final report
- Implementation plan of the evaluation recommendations (Management Response)

The main tasks of the coordination are to be fulfilled entirely by the contracting entity of the evaluation. According to the evaluation definitions of ADA, responsibility lies with the project partner during the internal evaluations; in case of external evaluations, responsibility lies with ADA.

Details on the individual managerial tasks, the international quality standards of an evaluation, cross-cutting issues (poverty, gender, environment), on the checklist for gender, the formats for ToR, the inception report, the data collection planning worksheet, the evaluation report, the management response as well as on further literature and relevant internet addresses can be found in the Annex.

5. The management of project and programme evaluations by project partners (internal evaluation)

Depending on the financial volume of a project or programme, ADA is involved in project or programme evaluations managed by the project partner to different extents.

The following overview refers to the projects or programmes under the administrative support of the ADA Unit of Countries and Regions.

Project Programme size (Unit for Countries and Regions)	**Inclusion of ADA**
Under EUR 500,000.-	Approval of the ToR The final evaluation report and the management response need to be sent to the CO and/or the ADA Headquarters for information. *<u>Information:</u> Quality control and approval of the evaluation report are done by the project partner.*
EUR 500,000.- to EUR 1 million	Approval of the ToR The final draft evaluation report is sent to the CO and/or the ADA Headquarters for comments. Equally, the final evaluation report as well as the implementation plan of the evaluation recommendations are sent to the CO and/or the ADA Headquarters for information. *<u>Information:</u> Quality control and approval of the evaluation report are done by the project partner.*

For individual projects and programmes administered by ADA Units NRO Cooperation (Microprojects) and Humanitarian Aid, Private Sector Development and the Department for Development Communication and Education in Austria, no compulsory evaluations on the part of ADA are required given the relatively low financial volume or their kind respectively. Within the framework of the budget request, however, an evaluation can be approved on an individual basis. Other instruments of the NRO cooperation are evaluated in accordance with their guidelines and directives.

Referring to an evaluation coordinated by the project partner, ADA has to approve the ToR. Without this approval, no evaluation can be financed by ADA. The ToRs are generally examined by the CO, the responsible unit and the Evaluation Unit. The assessment of the ToR is considered as a significant measure for quality assurance.

Project or programme evaluations managed by the project partner do not require a participation of ADA in the recruiting process (e.g. submission and approval of the CVs of the individual evaluators)[8]. However, it is clearly stated that the selection of an evaluation teams or an individual person must be documented in writing and submitted to a competition procedure. See also Annex 7.6.

If serious quality defects (e.g. in the evaluation report) are determined, ADA can partly or fully deny the costs of an evaluation.

[8] The ADA Departments and the ADA Evaluation Unit upon request provide help with the identification of qualified evaluators (e.g. at the propagation of the ToR in relevant international forums.)

6. The management of project and programme evaluations by coordination offices or headquarters of the Austrian Development Agency (external evaluation)

External project or programme evaluations are either coordinated by the relevant CO or ADA departments or units at headquarters respectively. The decision whether a project or programme evaluation will be conducted is taken by ADA in consultation with project partners.

Programmes and projects with a volume of more than EUR 1 million are evaluated externally and coordinated by ADA. It is decided on a case-by-case basis whether such evaluations are managed by a CO, a department, a unit or by the Evaluation Unit.

If necessary, programmes or projects (Unit for Countries and Regions) the volume of which is below EUR 1 million may in addition to internal evaluation also be evaluated externally.

A larger number of projects and programmes with the same topical or geographical focus may also be analysed within one evaluation only. Such an evaluation will be managed by ADA (CO, Unit or Evaluation Unit) upon agreement.

As far as more comprehensive projects, programmes or other interventions are concerned, the intention is to carry out more joint evaluations, impact evaluations and budget support evaluations in the future.

The ADA Evaluation Unit supports, upon request, other ADA departments, units and partners in their evaluation projects, or manages evaluations by itself. Also see Guidelines for Strategic Evaluations.

7. Annexes

7.1 Definitions

7.2 OECD/DAC evaluation criteria

7.3 OECD/DAC evaluation criteria for humanitarian aid

7.4 Format for Terms of Reference

7.5 Cross-cutting issues (poverty, gender, environment)

7.6 Coordination details of an evaluation

7.7 International evaluation quality standards (DAC Evaluation Standards)

7.8 Gender checklist

7.9 Format for an inception report

7.10 Format for a data collection planning worksheet

7.11 Format for an evaluation report

7.12 Format for a management response

7.13 Challenges for an evaluation or the evaluation team

7.14 Further literature and internet addresses

Annex 7.1
Definitions

If not specified differently, all definitions are taken from the OECD/DAC Glossary of Key Terms in Evaluation and Results Based Management, Paris, 2002.[9]

Audit An independent, objective assurance activity designed to add value and improve an organization's operations. A distinction is made between regularity (financial) auditing, which focuses on compliance with applicable statues and regulations, and performance auditing, which is concerned with relevance, economy efficiency and effectiveness. Internal auditing provides an assessment on internal controls undertaken by a unit reporting to management while external auditing is conducted by an independent organization.
Cluster Evaluation An evaluation of a set of related activities, projects and/or programs.
Evaluability Extent to which an activity or a program can be evaluated in a reliable and credible fashion. Note: Evaluability assessment calls for the early review of a proposed activity in order to ascertain whether its objectives are adequately defined and its results verifiable.
Evaluation The systematic and objective assessment of an on-going or completed project or programme, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability.
Ex-Ante Evaluation An evaluation that is performed before implementation of a development intervention
Ex-Post Evaluation Evaluation of a development intervention after it has been completed
External Evaluation The evaluation of a development intervention conducted by entities and/or individuals outside the donor and implementing agencies.
Formative Evaluation Evaluation intended to improve performance, most often conducted during the implementation phase of projects or programs.
Impact Evaluation/Assessment a) Impact: Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended. b) An assessment of impact using before/after and/or with/without comparison. (DFID. Guidance on Evaluation and Review for DFID Staff, DFID, London, 2005.)
Internal Evaluation Evaluation of a development intervention conducted by a unit/and or individuals reporting to the management of the donor, partner, or implementing organization.
Joint Evaluation An evaluation to which different donor agencies and/or partners participate.

[9] http://www.oecd.org/site/0,3407,en_21571361_34047972_1_1_1_1_1,00.html

Meta-Evaluation The term is used for evaluations designed to aggregate findings from a series of evaluations. It can also be used to denote the evaluation of an evaluation to judge its quality and/or assess the performance of the evaluators.
Mid-Term Evaluation Evaluation performed towards the middle of the period of implementation of the intervention.
Monitoring A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds.
Ongoing Evaluation An evaluation designed to help the development of a project as it proceeds. It is more wide-ranging and thorough and objective than an internal review. (DFID. Guidance on Evaluation and Review for DFID Staff. London. 2005.)
Output to Purpose Review (OPR) The periodic review of ongoing projects, focusing on progress at the Purpose level, and whether the performance of the project implementation in terms of achievements under each Output are sufficient to reach the Purpose at the end of the project. Additional outputs may be added as a result of the OPR. (DFID. Guidance on Evaluation and Review for DFID Staff. London. 2005.)
Participatory Evaluation Evaluation method in which representatives of agencies and stakeholders (including beneficiaries) work together in designing, carrying out and interpreting an evaluation.
Peer Review is basically an examination of one state's performance or practices in a particular area by other states. The point of the exercise is to help the state under review improve its policymaking, adopt best practices and comply with established standards and principles. (OECD Peer Review at a Glance)
Process Evaluation An evaluation of the internal dynamics of the implementing organizations, their policy, instruments, their service delivery mechanism, their management practices and the linkages among these.
Review An assessment of the performance of an intervention, periodically or on an ad hoc basis. Reviews are usually less comprehensive and/or in-depth than evaluations. They tend to emphasize operational aspects.
Self-Evaluation An evaluation by those who are entrusted with the design and delivery of a development intervention.
Study/Research The systematic investigation into and study of materials and sources in order to establish facts and reach new conclusions.
Summative Evaluation A study conducted at the end of an intervention (or a phase of that intervention) to determine the extent to which anticipated outcomes were produced. Summative evaluation is intended to provide information about the worth of the program.
Triangulation The use of three or more theories, sources or types of information to verify and substantiate an assessment. Note: by combining multiple data sources, methods, analyses or theories, evaluators seek to overcome the bias that comes from single informants, single methods, single observer or single theory studies. (DFID. Guidance on Evaluation and Review for DFID Staff. London. 2005.)

Annex 7.2
OECD/DAC evaluation criteria

Definition of evaluation

"The systematic and objective assessment of an on-going or completed project, programme or policy its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability."

An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision making process of both recipients and donors. (OECD. Development Assistance Committee Working Party on Aid Evaluation. Glossary of Key terms in Evaluation and Results Based Management. Paris. 2002
(http://www.oecd.org/dataoecd/29/21/2754804.pdf).

7.2.1 Relevance

OECD/DAC: ***The extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, global priorities and partner' and donor's policies.***

Are we doing the right things? What is the **relevance or significance** of the intervention regarding local and national requirements and priorities?

- To what extent does the intervention comply with development policy and planning of the recipient country or the partner government?
- How important is the intervention for the target group and subgroups (e.g. women), and to what extent does it address their needs and interests?
- To what extent do the basic principles of Austrian development policy – poverty reduction, promotion of democracy and human rights, gender equality, respect for the cultural background and ecology – correspond with the respective strategic goals and programmatic targets?

Examination questions of German development cooperation

To what extent does the development intervention aim at the solution of a **core problem of the target group(s)?** Is the most recent perspective taken into account? Does it play a role in terms of development policy (according to gender, ethnic groups, conflict parties, etc.)?

To what extent does the development intervention correspond with the most recent objective of the partner country's development policy (Government: **Poverty Reduction Strategy Paper** (PRSP) or similar other relevant groups in case of conflict of interests, if applicable)? Is the solution of a core problem that is important in terms of development policy or a decisive development shortage of the partner country being tackled by the development intervention?

To what extent does the objective of the development intervention in terms of development policy correspond with the objectives and directives of the Federal Ministry for Economic Cooperation and Development (BMZ) (poverty reduction, **Millennium Development Goals** (MDGs), cross-cutting issues, gender equality, participatory development, good governance, protection of environment and resources, as well as crisis prevention, objectives of the country concept and the focus strategy paper(s), targets of relevant sectoral concepts)?

To what extent does the basic orientation and conception regarding development policy of the development intervention correspond with the most recent requirements, standard of knowledge and framework conditions (For example, is the **cause-effect hypothesis** plausible?)?

Source: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, Referat 120, Evaluierung der Entwicklungszusammenarbeit. Bonn. Juli 2006.

Examples for questions of relevance:

a) What is the relevance of the instruments and the projects/programmes selected by the Austrian Development Cooperation (ADC) for the partner countries of ADC? To what extent do they correspond with the priorities, the needs and the practical requirements of the partner countries? (ADA Education Evaluation, 2005)

b) What development and other effects are the development and business partnerships supported by ADC aspiring to? (ADA-Evaluation of the Instruments of Partnerships and Business Partnerships, 2008)

c) Was the humanitarian assistance provided in line with the humanitarian policy and procedures of the Netherlands, as well as the needs, priorities and rights of the affected populations? (Evaluation of the Dutch Humanitarian Assistance 2000–2004, 2005)

7.2.2 Effectiveness

OECD/DAC: *The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance. Note: Also used as an aggregate measure of (or judgement about) the merit or worth of an activity, i.e. the extent to which an intervention has attained, or is expected to attain, its major relevant objectives efficiently in a sustainable fashion and with a positive institutional developmental impact.*

Have the objectives of the development intervention been achieved? How big is the **effectiveness or impact** of the project compared to the objectives planned (Comparison: result – planning)?

- To what extent will the objectives of the intervention be (most likely) achieved?
- To what extent is the target group reached?

Examination questions of German development cooperation

To what extent were the originally defined objectives of the development intervention realistic? To what extent do they still meet the most recent requirements and the most recent standard of knowledge? To what extent have the (direct) objectives of the development intervention been achieved in accordance with the (adjusted, if applicable) target system? What are the (concrete) contributions of interventions of the German DC for achieving the objectives of the development intervention? What factors were crucial for the achievement or failure to achieve the project objectives so far (indication of strengths and weaknesses, e.g. the monitoring and evaluation system)? What is the quality of development-policy and technical planning and coordination (e.g. BMZ, project partner, partner)? How can it be judged? What other effects – also negative ones – can be established regarding services and (project/programme) objectives? *Source: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, Referat 120, Evaluierung der Entwicklungszusammenarbeit. Bonn. Juli 2006.*

Examples for questions of effectiveness:

a) Has the Austrian Education Cooperation contributed to sustainable capacity building, knowledge transfer in partner countries as well as to the improvement of the scientific-technical potential? (ADA Education Evaluation, 2005)

b) How effective is the support in the initiation of business partnerships (advise and information of interested companies by ADA, travel allowances, feasibility studies)? (ADA-Evaluation of the Instruments of Partnerhsips and Business Partnerships, 2008)

c) To what extent did the humanitarian assistance provided achieve its purpose? (Evaluation of the Dutch Humanitarian Assistance 2000–2004, 2005).

7.2.3 Efficiency

OECD/DAC: *A measure of how economically resources/ inputs (funds, expertise, time, etc.) are converted to results.*

Are the objectives achieved in a cost-efficient manner by the development intervention? How big is the **efficiency or utilisation ratio** of the utilised resources? (comparison: provided means – results):

- Is the relationship between input of resources and results achieved appropriate and justifiable? What is the cost-benefit ratio?
- To what extent have individual resources been used economically?
- Are there any alternatives for achieving the same results with less inputs/funds?

Examination questions of German development cooperation

How high were the costs? (e.g. by instruments, sectors, interventions)? To what extent were the costs and benefits of the development interventions in a reasonable proportion to each other from a business and economic point of view?

Would there have been cheaper solutions/alternatives concerning the utilisation of instruments and the conceptualization of the development intervention achieving the objectives on a sustainable basis?

Are the services, capacities created and potentials used appropriately?

Were services provided in time and impacts achieved within an appropriate time period?

Source: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, Referat 120, Evaluierung der Entwicklungszusammenarbeit. Bonn. Juli 2006.

Examples for a question of efficiency:

a) Were the financial resources and other inputs efficiently used to achieve results? Issues to be addressed:

- Aid management (programme and project cycle, staffing, tasks and responsibility of ministry departments and embassies, inter-ministerial co-operation include civil-military co-operation)
- Criteria used in the selection of implementing partners (comparative advantage or other)
- Use of monitoring of progress and achievements for programming, learning and accountability (Evaluation of the Dutch Humanitarian Assistance 2000–2004, 2005).

7.2.4 Impact

OECD/DAC: *The positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended.*

- What has happened as a result of the programme or project?
- What real difference hast the activity made to the beneficiaries?
- How many people have been affected?

Does the development intervention contribute to the achievement of overall development objectives (tendentially, overall goal)? What is or are the **impact(s)/effects** of the intervention compared to the total situation of the target group or those affected:

- positive and negative, intended and unintended effects
- technical, economic, social, cultural, political, ecological effects – disaggregated by sex or other relevant social groups, such as minorities

Examination questions of German development cooperation

To what extent were the originally intended, overriding objectives in terms of development policy (goals) realistic? To what extent do they still correspond with the most recent requirements and the most recent standard of knowledge?

To what extent have (according to the most recent requirements and the most recent standard of knowledge) appropriate overriding effects regarding development been achieved so far? What has the development intervention contributed to so far and what is it still contributing to?

To what extent was the development intervention exemplary, created structures and/or had a broad effect/impact in terms of leverage (e.g. adaptation among target groups and organisations)?

What other effects – also negative ones – can be determined on a goal level?

What would the development have been like without the development intervention?

Source: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, Referat 120, Evaluierung der Entwicklungszusammenarbeit. Bonn. Juli 2006.

7.2.5 Sustainability

OECD/DAC: ***The continuation of benefits from a development intervention after major development assistance has been completed. The probability of continued long-term benefits. The resilience to risk of the net benefit flows over time***

Are the positive effects sustainable? How is the **sustainability or the continuity** of the intervention and its effects to be assessed?

- To what extent will activities, results and effects be expected to continue after donor intervention has ended?
- To what extent does the intervention reflect on and take into account factors which, by experience, have a major influence on sustainability like e.g. economic, ecological, social and cultural aspects?
- How self-supporting in particular is the assisted local counterpart?

Examination questions of German development cooperation

To what extent are the positive changes and effects of the development intervention (summarily) sustainable compared to the objectives regarding development policy?

Particularly: How stable is the situation in the surrounding field of the development intervention regarding social justice, economic efficiency, political stability and ecological balance?

What risks and potentials are visible regarding the sustainable effectiveness of the development interventions and how likely is their occurrence? Will the effectiveness of the development intervention most likely improve or worsen in future?

To what extent is/are the target group(s) capable and prepared to receive the positive effects of the development intervention without support in the long term?

To what extent are the (counterpart) organisations (financially, personnel-wise and in terms of organisation) capable and prepared to maintain the positive effects of the development interventions without support in the long term?

To what extent are the target groups and counterparts able to adapt sufficiently to external changes and shocks?

Source: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, Referat 120, Evaluierung der Entwicklungszusammenarbeit. Bonn. Juli 2006.

Example for a question of sustainability:

To what extent did the projects/programmes strengthen local ownership and leadership? (ADA-Education Evaluation, 2005)

Annex 7.3
OECD/DAC evaluation criteria for humanitarian aid

During the revision of the five OECD/DAC evaluation criteria of relevance, effectiveness, efficiency, sustainability and impact for humanitarian aid, the request was made in pertinent circles of experts to concretise or adapt them so that they can better meet the requirements of humanitarian aid.

Four of the five criteria have remained the same in their basic definition. The criterion of sustainability is not specifically examined, because interventions of humanitarian aid generally are of short-term duration. The question relating to sustainability is therefore defined differently and is presented as '*Connectedness*'. New criteria are: connectedness, coverage, coherence.

The following criteria and evaluation questions are suggested by the OECD/DAC and the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP) in humanitarian aid:

7.3.1 Relevance/Appropriateness

ALNAP: *Relevance is concerned with assessing whether the project is in line with local needs and priorities (as well as donor policy).*

ALNAP: Appropriateness is the tailoring of humanitarian activities to local needs, increasing ownership, accountability, and cost-effectiveness accordingly.

The criteria of relevance and appropriateness are complementary, relevance refers to wider goals and priorities, and appropriateness refers to activities and inputs.

Examples for possible questions:

- Was a needs analysis carried out, in which the needs of men, women, boys and girls were identified?
- Did the intervention take into account the livelihood and capacities of the target group?
- Were interventions in some cases more relevant and more appropriate than in other cases?

7.3.2 Coherence

ALNAP: *The need to assess security, development, trade and military policies, as well as humanitarian policies, to ensure that there is consistency and, in particular, that all policies take into account humanitarian and human rights considerations.*

Coherence refers to the consistency/complementarity and freedom of contradiction of guiding general principles of different topics, such as development, trade, military, humanitarian aid, analysing whether human rights have been taken into consideration in all principles or not.

Examples for possible questions:

- How was coordination (coherence) achieved, and/or why was there a lack in coherence?
- What political factors were specifically responsible for the coordination of assistances or relief items or what made the latter more difficult?
- Is coherence necessary or feasible in the present situation at all?

7.3.3 Effectiveness

ALNAP: *Effectiveness measures the extent to which an activity achieves its purpose, or whether this can be expected to happen on the basis of the outputs. Implicit within this criterion of effectiveness is timeliness.*

Examples for possible questions:

- Effectiveness is measured on the basis of the defined outputs and outcomes. How was the decision taken regarding these results or objectives? Was a needs analysis conducted? Were the objectives clearly defined? Who participated in the decision-making process? Was the main target group involved in the project planning phase?
- To what extent have the objectives been achieved or not? What are key paramters/determining factors that have influenced the achievement of the objectives ? Lessons learnt have to be taken into account in future interventions.
- Did the interventions reach the target population? In cases where monitoring data (structured according to sex, socio-economic categories, ethnicity) are not available/have not been collected, they have to be collected in interviews with the main target groups.
- Are the statements of the main target group on the attainment of goals identical with the opinions of the actors having provided humanitarian assistance (e.g. employees of the respective organisation)?
- Have goods, services or other subsidies been delivered or offered at the right time according to the main target group?
- Have the interventions contributed to strengthening the core potentials of the target groups with regard to new risks?

7.3.4 Coordination

ALNAP: *While not a 'formal' DAC criterion, coordination is an important consideration in the evaluation of humanitarian action. Coordination cuts across several criteria, but is included here under the heading of effectiveness, as the two are closely related.*

Contrary to coherence, which deals with the question whether the policies of different participants are consistent, coordination refers to the practical activities of governments and organisations – whether the latter have Joint Common Cluster Groups, have discussed geographic target areas and how the information was shared.

Capturing and assessing coordination requires a discussion with mostly a large number of actors, an analysis whether the responsible government was really involved in the decision-making processes, as well as a description of the role of non-traditional partners, such as the military, for example.

Examples for possible questions:

- Were there any local coordination structures? Were there plans for these local coordination structures? How did the organisations harmonise and coordinate their interventions with other partners? How actively were organisations involved in the coordination?
- What partners were involved in the coordination and how? Why were they included? Were any organisations not involved?
- Were there any reasons for not participating or participating only to a small extent in the coordination?
- Were there any incentives for coordination? Was e.g. the UN coordination supported by donors with any financial means?

- Was a lead agency appointed? Which organisation was appointed lead agency by which procedure? Which tasks has the organisation accomplished for promoting coordination? How effectively is the latter perceived?
- Which factors have restricted the coordination, and which factors have supported it? How was good coordination achieved? What is transferable to other situations in this respect?
- What effects did the coordination have on the interventions of humanitarian aid? Did the coordination lead to better effectiveness and impact of the interventions?
- Was the allocation of financial resources coordinated in an organised manner or were funds provided by donors individually in order to support their own strategic objectives?

7.3.5 Efficiency

ALNAP: Efficiency *measures the outputs – qualitative and quantitative – achieved as a result of inputs. This generally requires comparing alternative approaches to achieving an output, to see whether the most efficient approach has been used.*

During humanitarian aid often a large quantity of material goods is provided, therefore conducting a comprehensive cost-benefit analysis is important. In connection with efficiency, political considerations should also be taken into account, e.g. if an intervention takes place in a conflict area but the government does not approve the support in the long run.

Examples for possible questions:

- Was a needs assessment carried out, in which the needs were clearly assessed and the services required mentioned in accordance with the situation?
- Were the commodities (inputs) utilised as planned?
- Is there a potential to make better use of the resources than previously? Generally, is there a potential for optimisation concerning planning, procurement and logistics? Would it have been possible to obtain certain goods in a better way and, perhaps, cheaper somewhere else?

7.3.6 Impact

ALNAP: *Impact looks at the wider effects of the project – social, economic, technical, environmental – on individuals, gender- and age groups, communities and institutions. Impacts can be intended and unintended, positive and negative, macro (sector) and micro (household).*

Impact refers to the long-term changes and is not equally relevant for all evaluations. Consequently, evaluations carried out during or shortly after an intervention can only partly provide information on long-term effects. Classic impact evaluations are characterised by a very complex methodology.

7.3.7 Connectedness/also Sustainability

ALNAP: *Connectedness refers to the need to ensure that activities of a short-term emergency nature are carried out in a context that takes longer-term and interconnected problems into account.*

Connectedness derives from the criterion of sustainability. Although operations of humanitarian aid are generally planned as short-term interventions, they should nevertheless contribute to interventions planned in the longer term, such as recovery or development.

Examples for possible questions:

- Does a sensible exit strategy exist including schedule and guidelines for the transfer of responsibility and activities to government departments and/or development organisations? Is there a budget scenario for the time after the assistance?
- Were financial means used for relief or recovery?
- What influence did already existing networks have (e.g. national and international non-governmental organisations) on the implemented interventions? Which lessones learnt could be relevant for others?
- To what extent were local capacities developed or strengthened through the humanitarian interventions?

7.3.8 Coverage

ALNAP: *The need to reach major population groups facing life-threatening risk wherever they are.*

Coverage is to be viewed in connection with effectiveness. ALNAP summarises detailed questions concerning the target group in a criterion of its own. Evaluation questions referring to that, however, can also be included in the criterion of effectiveness (OECD/DAC 1999).

Examples for possible questions:

- Who was supported by the humanitarian interventions? Which groups were taken into account and which not?
- What were the main reasons for certain parts of the target groups having received support and protection and others having been excluded?
- Was the support aligned to regionally different needs? What decisions were taken or not in this regard?
- Who has really received support on a local level (village, slum, community and/or refugee camp)? Data should be analysed and interpreted in a sex-disaggregated manner, socioeconomic categories and ethnicity.
- Have all of those in need of protection received protection during the interventions?

ALNAP covers the topic of protection within the criteria of coverage. The OECD/ DAC considers protection as an additional topic:
Beside the criteria already mentioned, the OECD/DAC also mentions protection. If protection is inadequately and there is the risk of members of the target group losing their lives, this must be considered in the analysis of the evaluation. The topic of security or protection should also be included in every evaluation referring to humanitarian aid (OECD/DAC 1999).

Literature

Beck, Tony, Evaluating Humanitarian Action Using the OECD-DAC Criteria. An ALNAP Guide for Humanitarian Agencies, London, 2006
http://www.alnap.org/

OECD DAC, Encourage Effective Evaluation of Conflict Prevention and Peacebuilding Activities, Towards DAC Guidance, 2007

OECD, DAC Evaluation Network, DAC Evaluation Quality Standards, March 2006
http://www.oecd.org/site/0,3407,en_21571361_34047972_1_1_1_1_1,00.html

OECD, DAC, Guidance for Evaluating Humanitarian Assistance in Complex Emergencies, 1999
http://www.oecd.org/site/0,3407,en_21571361_34047972_1_1_1_1_1,00.html

VENRO. Humanitäre Hilfe auf dem Prüfstand: Prinzipien, Kriterien und Indikatoren zur Sicherstellung und Überprüfung der Qualität in der Humanitären Hilfe. VENRO Arbeitspapier Nr. 14, Verband Entwicklungspolitischer Deutscher Nichtregierungs-organisationen e.V., Bonn, August 2005.
http://www.venro.org

Annex 7.4
Format for Terms of Reference

Background
… contains a short description of the project to be evaluated.

Purpose
… explains what learning aspects and accountability functions are referred to. E.g.: a) The evaluation is intended to contribute to an improvement of policies, processes and methods. b) The evaluation makes recommendations on whether a project or programme should be continued. c) The evaluation contributes to the accountability towards the stakeholders and taxpayers (priority: control aspect).

Objectives
… explain why and for whom the evaluation is carried out.
… illustrate why the evaluation takes place at a certain point of time.
… explain what exactly the evaluation wants to achieve. E.g. a) it revises results (output, outcome, impact) and assesses the effectiveness, efficiency and relevance of an intervention. b) represents results, conclusions and recommendations with regard to policies, programmes, etc.

Subject and focus (scope)
The scope defines which topics/themes relating to development interventions are addressed to or taken into consideration (duration, kinds of interventions, geographic scope, target groups, funds of interventions and other aspects). Deliberate restrictions of the evaluation are substantiated, e.g. if one or more of the five evaluation criteria (relevance, efficiency, effectiveness, impact and sustainability) are not applied. A substantiation is also required if additional criteria (e.g. coordination issues, participation) are applied. It also needs to be mentioned whether cross-cutting issues (such as e.g. poverty, gender and environment) are taken into consideration or if the intervention logic (e.g. logframe) will be analysed.

Main evaluation questions
The questions of evaluation should be formulated as concretely as possible and in accordance with the basic evaluation criteria (relevance, efficiency, effectiveness, impact, sustainability), as well as in accordance with other aspects (e.g. coordination issues, participation).

Evaluation approach and methods
… comprise a content-related description of what is expected in the respective phases of an evaluation (desk study, inception phase, field studies, preparation of inception report and final report, presentations). Number of the total working days planned, as well as number and duration of field trips. Brief description of methods. Indication that data will e.g. be collected and interpreted in a sex-disaggregated manner. Furthermore, it should be indicated that the OECD DAC Evaluation Quality Standards are to be applied and that the compliance of the latter needs to be comprehensible in the evaluation.

Timetable
… is a chronological description of the respective phases of an evaluation (tender, acceptance, desk study, inception phase including submission of inception report, field studies, preparation of final draft report and final report, presentations).

Evaluation team
Number and key qualifications of evaluators (geographic and technical expertise, experience in evaluation, cross-sectoral and gender expertise), requirements of the team leader, composition of the team (national/international), qualifications of national experts. Indications on how qualifications can be proven (e.g. CVs, reference evaluations).

Reports
Inception report, final draft evaluation report, final report. Language, scope and maximal length of reports. Indication that the final draft evaluation report and final report need to be structured according to the OECD/DAC criteria and the evaluation questions. Notes which criteria are used for the assessment of the quality of the evaluation report (Beyond that, a reference can be made to the evaluation quality criteria of the OECD/DAC in the ToR):

- Were the terms of reference fulfilled and is this reflected in the report?
- Does the report contain a comprehensive and clear summary?
- Is the report structured according to the OECD/DAC criteria and the evaluation questions?
- Are cross-cutting issues (e.g. poverty, gender, environment) indicated in the report separately?
- Does the report describe and assess the intervention logic (e.g. logframe)?
- Are the conclusions and recommendations based on findings clearly stated in the report, and are they derivable from the latter?
- Does the report clearly differentiate between conclusions, recommendations and lessons learnt?
- Is it comprehensible how the evaluators have achieved their findings?
- Are the recommendations and lessons learnt realistic and is it clearly expressed to whom the recommendations are addressed to?
- Are the methods and processes of the evaluation sufficiently documented in the evaluation report?
- Were the most significant stakeholders involved consulted?
- Were the most important documents taken into consideration, and is the content of the latter reflected in the report?
- Does the report present the information contained in a presentable and clearly arranged form?
- Is the report free from spelling mistakes and unclear linguistic formulations?
- Can the report be distributed in the delivered form?

Coordination/Responsibility
Responsibility and competence for the evaluation. Clarification whether and what logistical support is offered to the evaluation team.

Annexes
E.g. literature list, project and/or programme details

Annex 7.5
Cross-cutting issues (poverty, gender, environment)

Questions regarding cross-cutting issues are in many cases already included implicitly or explicitly in the main questions of the evaluation criteria. In order to facilitate assessments of cross-cutting issues, the cross-cutting issues also need to be summarised separately in the evaluation report.

Examples for questions to the evaluation of cross-cutting issues:

a) Assessment of the project with regard to poverty reduction and Millennium Development Goals (MDGs)

In the assessment of a project it should also be assessed to what extent this project contributes to poverty reduction as well as to the achievement of the MDGs.

Central questions to this issue are:
- Was the conception of the project differentiated by target groups and was a poverty analysis available?
- Was the project embedded in a poverty-oriented environment (national poverty reduction strategy or Poverty Reduction Strategy Paper)?
- Did the project include the participation of poor population groups in economic and political processes?
- Did the project achieve poverty-reducing effects? Did the project improve the living conditions of the target group (access to education, health, nutrition, water, environment)?

b) Assessment of the project with regard to the promotion of gender equality

The assessment of the project should also contain an assessment of the extent to which it contributes to the promotion of gender equality.

Central questions to this issue are:
- Was the conception of the project gender-differentiated and was a gender analysis conducted?
- Did women and men make an equal contribution to the design of the project?
- Do women and men equally benefit from the project?

c) Assessment of the project with regard to the promotion of the environment

The assessment of the project should also contain an assessment of the extent to which it has positive effects on the environment.

Central questions to this issue are:
- What actual or expectable positive and negative effects are foreseeable on the environment?
- What protection and monitoring measures were applied and how effective were they?
- What are the possible risks following the completion of the project?

Literature

Federal Ministry for Economic Cooperation and Development, Evaluation Grid, Bonn, August 2006.

Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), Terms of Reference, Fremdevaluierung eines laufenden Vorhabens/Schlussevaluierung/Ex-Post Evaluierung, 19 December 2006

Annex 7.6
Coordination details of an evaluation

Preparation of the Terms of Reference

The draft of the Terms of Reference needs to be **coordinated with all parties concerned** as early as possible (particularly with local partners, possibly with representatives of the target group or the local government). The amount of time required to prepare the ToR is often underestimated. This phase is generally the longest and most difficult one in the entire evaluation process.

Clearly structured evaluation questions are the core of ToR. They need to be derived for the purpose of the evaluation. The evaluation questions are to be formulated following the OECD/DAC evaluation criteria.

Apart from the decision on when an evaluation should be carried out within the project cycle, climatic, seasonal and cultural-religious (e.g. Ramadan) events also need to be considered.

Selection of the evaluation team

The **evaluation team** should generally consist of at least **two persons** having different professional qualifications. Furthermore, a woman and a man should be represented in the team**.** This is how gender-related issues can be dealt with more easily. At least one member of the evaluation team must have a profound gender expertise[10]. Several gender analysis frameworks are available[11]. Also the ADA Guidelines for Gender Equality and Empowerment of Women, as well as the Toolkit on Mainstreaming Gender of the European Commission are helpful basic documents[12]. The team leader has to guarantee the integration of the gender perspective.

Taking into consideration the funds provided in the budget, the composition of the team of evaluators also remains to be clarified in the Terms of Reference.

Beside **international evaluators** also **national experts or experts from the respective region of the destination country** should be integrated into the evaluation team, as the latter generally have a better knowledge of the local situation.

The **independence** of the evaluator team is of fundamental significance: In terms of an evaluator´s credibility, the latter has to be independent from the organisation implementing the project/programme as well as possible local partners. **On no account** may evaluators have been involved in the planning process of the respective project or programme or in the monitoring of the latter. The appearance of a lack of independence can already jeopardise an evaluation.

During the recruitment of the evaluators, in any case the provisions of the **General Conditions of Contract** established in the contract with ADA have to be considered. The latter specify the **obtainment of several offers**.

---

[10] If there is only one evaluator with no gender expertise, the expertise needs to be obtained from another place.

[11] Oxfam. A Guide to Gender-Analysis Frameworks. Oxfam 1999.

[12] Austrian Development Agency (ADA). Geschlechtergleichstellung und Empowerment von Frauen. Leitlinien der Österreichischen Entwicklungs- und Ostzusammenarbeit. April 2006; European Commission. Toolkit on Mainstreaming Gender Equality in EC Development Cooperation. Brussels. 2004.

Concerning the identification of evaluators the following procedures are considered successful in the past from the perspective of ADA:

- Propagation of the ToR (without information on the evaluators´ fee) in relevant circles with the request for forwarding them
- Additional request to the responsible CO (if existent), the responsible unit and/or other organisations active in the sector/country to communicate the ToR
- Distribution of the ToR on relevant websites (e.g. Reliefweb[13])
- Forwarding of the ToR to relevant consulting firms, consultants´ networks and professional associations of evaluators[14]
- Advertisement in print media.

A mix of the procedures mentioned above together with a long application deadline generally leads to good results.

Given that the evaluation team requires a very specific content-related knowledge (knowledge of the target country, knowledge of the sectors concerned, gender-specific analysis skills, evaluation methods) and at the same time must be independent from the implementing organisation, it is recommended to look, in **procedures as open and transparent as possible,** in order to search for the best suitable experts.

In most cases evaluation teams (consulting firms, agencies, etc.) will make a technical and financial offer on their own separately. These offers are assessed independently of each other. In order for offers to be assessed not only under the aspect of costs it is important to assess the technical offer first. In most cases, a committee is set up for the purpose. It must also be clearly visible from the offers which of the evaluators proposed will act as team leader.

The selection of the evaluation team must be made in accordance with the procedure mentioned in the ToR.

In case of project or programme evaluations coordinated by the project partners, an involvement ofADA in the recruiting process is not necessary (e.g. submission and approval of the CVs of the individual evaluators).

In case of the **conclusion of the contract with the evaluation team** by the project partner it must be particularly taken into account that the organisation awarding the contract includes all regulations into the contract necessary for the project partner to be able to fulfil its obligations vis-à-vis ADA. (This concerns in particular the calculation of travel allowances and the submission of original receipts[15]).

---

[13] http://www.reliefweb.org

[14] On the website of OECD/DAC, evaluations commissioned by the most different donors can be accessed according to target countries (“DeRec“).

http://www.oecd.org/pages/0,3417,en_35038640_35039563_1_1_1_1_1,00.html

In the evaluation reports, companies or experts having carried out evaluations can be identified. Furthermore, there is a German-speaking society for Evaluation, European Evaluation Society, African Evaluation Association and others.

[15] If no appropriate tuning is made, ADA may deny costs that have arisen within the framework of the evaluation during the settlement in case the carrier cannot fulfil its obligations.

Support of the evaluation team

The teamleader needs to be provided with **one copy of all documents** concerning the project or programme to be evaluated already at the **signing of the contract** (desk study). The latter include in any case:

- Project documents (including budget)
- All reports compiled to the project or programme
- Information on other projects or programmes closely connected with the project or programme to be evaluated
- List of contacts containing all persons involved in the implementation of the project or programme (with function, task, contact data and information on language skills) as well as all local partners (including representatives of the target group, as far as feasible).

All documents presented need to be dated and structured in an understandable and clear manner.

Furthermore, the evaluation team needs to be provided with the names of local contact persons. The latter have to provide every necessary support to the evaluation team.

Organisation of an introduction workshop

Following the signing of the contract, the contractor has to organise an introduction workshop for the evaluation team. During this workshop, subject specific information as well the ToR need to be jointly discussed. Generally a whole day should be reserved for this workshop in order to clarify content-related facts right at the beginning of the evaluation.

Following this workshop, the evaluation team will intensively study all documents, papers and evaluation questions (desk study).

Quality control approving the inception report

The inception report **must** contain detailed questions, hypotheses and indicators to the individual evaluation questions. Evaluation questions are generally structured in further sub questions. It is recommended to use a data planning worksheet for the purpose[16].

Furthermore, concrete evaluation methods and instruments should be presented and adapted to the evaluation questions: Which evaluation question is captured with which methods and instruments? Organisational details contain the exact time schedule as well as an interview and/or visitors' list.

Also methodological details on the formulation of cross-cutting issues (particularly gender) and the kind of analysis of the intervention logic need to be mentioned in the inception report.

For the presentation of the inception report, generally a workshop is organised so that methodological details and the questions which have occurred so far can be discussed. Following the workshop, suggestions are included in the inception report by the evaluation team, or a protocol is prepared, which is presented to the contractor for approval.

[16] A format can be found in the Annex.

Before an inception report was not officially approved by the contractor, the evaluation team is not permitted to start the field mission.

Quality control approving the final draft evaluation report

The contractors of the evaluation and ADA (the latter according to need) are involved in the quality control of the report.

The report criteria should already be mentioned in the ToR. The report needs to be critically assessed according to these criteria. ADA recommends the use of the format for evaluation reports.

Report criteria:

- Were the ToR fulfilled accordingly and is this reflected in the report?
- Does the report contain a comprehensive and clear summary?
- Is the report structured according to the OECD/DAC criteria and the evaluation questions?
- Are cross-cutting issues mentioned in the report separately?
- Does the report describe and assess the intervention logic (e.g. logframe)?
- Are the conclusions and recommendations based on clearly defined statements and can they be derived from the latter?
- Does the report clearly distinguish between conclusions, recommendations and lessons learnt?
- Is it comprehensible how the evaluators have achieved their findings?
- Can recommendations and lessons learnt be implemented and is it clearly recognisable whom they are directed to?
- Are the methods and processes of the evaluation sufficiently documented in the evaluation report?
- Were the most important stakeholders consulted?
- Were the most important documents taken into account and is their content reflected in the report?
- Does the report present the information in a presentable and clearly arranged form?
- Is the report free from spelling mistakes and unclear linguistic formulations?
- Can the report be distributed in the delivered form?

A workshop needs to be organised by the contractor also for the presentation of the final draft evaluation report to make it possible that the report can be discussed in detail. Comments made during the workshop are captured by the evaluation team. Additional comments before or after the workshop are generally collected by the contractor and forwarded to the evaluation team.

Quality control approving the final evaluation report

After comments were incorporated by the evaluation team, the final evaluation report will be sent to the contractor and ADA.

Finally, the contractor checks whether all comments have been included in the final report. If the report was approved positively, the contractor will approve it.

Following its completion, the report needs to be made accessible to **all employees involved in the project/programme** of the project partner or partner organisation. Also other partners, such as local government agencies, should receive the report. Furthermore, a **publication** on the Internet should follow as well. Only this way can insights from the evaluation be properly utilized (“learning effects“).

According to the contract with ADA, an electronic copy of the evaluation report will be sent.

Implementation plan of the evaluation recommendations (management response)

The decisive factor for the success of every evaluation is the **implementation of the recommendations**, as otherwise evaluations are useless[17]. Already in the ToR it needs to be clarified by whom and how the evaluation results will be utilised.

The implementation of the evaluation recommendations will be guided by a management response:

A **management response**[18] is a matrix listing all recommendations. The project partner comments on every recommendation and establishes whether they are fully, partially or not accepted at all. For recommendations having been accepted, steps for implementing them need to be noted.

It is the task of the project partner to prepare the management response and to implement the evaluation recommendations. The management response needs to be revised at least once a year in order to see to what extent the recommendations have already been implemented.

The management response needs to be forwarded to ADA (CO, ADA Unit).

Pursuing and implementing evaluation results is essential for good development cooperation.

---

[17] This is why mid-term evaluations are to be preferred to evaluations being implemented at the end of a project/programme.
[18] A format of a management response is outlined in the Annex.

Annex 7.7
International evaluation quality standards (DAC Evaluation Standards)

How can an evaluation be judged that it was carried out professionally? How can an evaluation be assessed formally? The Development Assistance Committee of the OECD has developed quality standards for that purpose[19].

1. Rationale, purpose and objectives of an evaluation

	ToR	Report
1.1 The rationale of the evaluation Describes why and for whom the evaluation is undertaken and why it is undertaken at a particular point in time.	X	X
1.2 The purpose of the evaluation The evaluation purpose is in line with the learning and accountability function of evaluations. For example the evaluation's purpose may be to: - Contribute to improving an aid policy, procedure or technique - Consider a continuation or discontinuation of a project/programme - Account for aid expenditures to stakeholders and tax payers	X	X
1.3 The objectives of the evaluation The objectives of the evaluation, specify what the evaluation aims to achieve. For example: - To ascertain results (output, outcome, impact) and assess the effectiveness, efficiency and relevance of a specific development intervention; - To provide findings, conclusions and recommendations with respect to a specific policy, programme etc.	X	X

2. Evaluation scope

	ToR	Bericht
2.1 Scope of the evaluation The scope of the evaluation is clearly defined by specifying the issues covered, funds actually spent, the time period, types of interventions, geographical coverage, target groups, as well as other elements of the development intervention addressed in the evaluation.	X	X
2.2 Intervention logic and findings The evaluation report briefly describes and assesses the intervention logic and distinguishes between findings at the different levels: inputs, activities, outcomes and impacts. The report also provides a brief overall assessment of the intervention logic.		X
2.3 Evaluation criteria The evaluation report applies the five DAC criteria for evaluating development assistance: relevance, efficiency, effectiveness, impact and sustainability. The criteria applied for the given evaluation are defined in unambiguous terms. If a particular criterion is not applied this is explained in the evaluation report, as are any additional criteria applied.		X

[19] OECD DAC Evaluation Network. DAC Evaluation Quality Standards. March 2006.

	ToR	Report
2.4 Evaluation questions The questions asked, as well as any revisions to the original questions, are documented in the report for readers to be able to assess whether the evaluation team has sufficiently assessed them.		X

3. Context

	ToR	Report
3.1 The development and policy context The evaluation report provides a description of the policy context relevant to the development intervention, the development agency's and partners' policy documents, objectives and strategies. The development context may refer to: regional and national economy and levels of development. The policy context may refer to: Poverty reduction strategies, gender equality, environmental protection and human rights.		X
3.2 The institutional context The evaluation report provides a description of the institutional environment and stakeholder involvement relevant to the development intervention, so that their influence can be identified and assessed.		X
3.3 The socio-political context The evaluation report describes the socio-political context within which the intervention takes place, and its influence on the outcome and impact of the development intervention.		X
3.4 Implementation arrangements The evaluation report describes the organisational arrangements established for implementation of the development intervention, including the roles of donors and partners.		X

4. Evaluation methodology

	ToR	Report
4.1 Explanation of the methodology used The evaluation report describes and explains the evaluation method and process and discusses validity and reliability. It acknowledges any constraints encountered and their impact on the evaluation, including their impact on the independence of the evaluation. It details the methods and techniques used for data and information collection and processing. The choices are justified and limitations and shortcomings are explained.		X

	ToR	Report
4.2 Assessment of results Methods for assessment of results are specified. Attribution and contributing/confounding factors should be addressed. If indicators are used as a basis for results assessment these should be SMART (specific, measurable, attainable, relevant and time bound).		X
4.3 Relevant stakeholders consulted Relevant stakeholders are involved in the evaluation process to identify issues and provide input for the evaluation. Both donors and partners are consulted. The evaluation report indicates the stakeholders consulted, the criteria for their selection and describes stakeholders' participation. If less than the full range of stakeholders was consulted, the methods and reasons for selection of particular stakeholders are described.		X
4.4 Sampling The evaluation report explains the selection of any sample. Limitations regarding the representativeness of the evaluation sample are identified.		X
4.5 Evaluation team The composition of evaluation teams should posses a mix of evaluative skills and thematic knowledge, be gender balanced, and include professionals from the countries or regions concerned.	X	X

5. Information sources

	ToR	Report
5.1 Transparency of information sources The evaluation report describes the sources of information used (documentation, respondents, literature etc.) in sufficient detail, so that the adequacy of the information can be assessed. Complete lists of interviewees and documents consulted are included, to the extent that this does not conflict with the privacy and confidentiality of participants.		X
5.2 Reliability and accuracy of information sources The evaluation cross-validates and critically assesses the information sources used and the validity of the data using a variety of methods and sources of information.		X

6. Independence

	ToR	Report
6.1 Independence of evaluators vis-à-vis stakeholders The evaluation report indicates the degree of independence of the evaluators from the policy, operations and management function of the commissioning agent, implementers and beneficiaries. Possible conflicts of interest are addressed openly and honestly.		X
6.2 Free and open evaluation process The evaluation team is able to work freely and without interference. It is assured of cooperation and access to all relevant information. The evaluation report indicates any obstruction which may have impacted on the process of evaluation.		X

7. Evaluation ethics

	ToR	Report
7.1 Evaluation conducted in a professional and ethical manner The evaluation process shows sensitivity to gender, beliefs, manners and customs of all stakeholders and is undertaken with integrity and honesty. The rights and welfare of participants in the evaluation are protected. Anonymity and confidentiality of individual informants should be protected when requested and/or as required by law.	-	X
7.2 Acknowledgement of disagreements within the evaluation team Evaluation team members should have the opportunity to dissociate themselves from particular judgements and recommendations. Any unresolved differences of opinion within the team should be acknowledged in the report.		X

8. Quality assurance

	ToR	Report
8.1 Incorporation of stakeholders' comments Stakeholders are given the opportunity to comment on findings, conclusions, recommendations and lessons learned. The evaluation report reflects these comments and acknowledges any substantive disagreements. In disputes about facts that can be verified, the evaluators should investigate and change the draft where necessary. In the case of opinion or interpretation, stakeholders' comments should be reproduced verbatim, such as in an annex, to the extent that this does not conflict with the rights and welfare of participants.	-	X
8.2 Quality control Quality control is exercised throughout the evaluation process. Depending on the evaluation's scope and complexity, quality control is carried out either internally or through an external body, peer review, or reference group. Quality controls adhere to the principle of independence of the evaluator.	-	-

9. Relevance of the evaluation results

	ToR	Report
9.1 Formulation of evaluation findings The evaluation findings are relevant to the object being evaluated and the purpose of the evaluation. The results should follow clearly from the evaluation questions and analysis of data, showing a clear line of evidence to support the conclusions. Any discrepancies between the planned and actual implementation of the object being evaluated are explained.		X
9.2 Evaluation implemented within the allotted time and budget The evaluation is conducted and results are made available in a timely manner in relation to the purpose of the evaluation. Un-envisaged changes to timeframe and budget are explained in the report. Any discrepancies between the planned and actual implementation and products of the evaluation are explained.		X
	ToR	**Report**
9.3 Recommendations and lessons learned Recommendations and lessons learned are relevant, targeted to the intended users and actionable within the responsibilities of the users. Recommendations are actionable proposals and lessons		X

learned are generalizations of conclusions applicable for wider use.		
9.4 Use of evaluation Evaluation requires an explicit acknowledgement and response from management regarding intended follow-up to the evaluation results. Management will ensure the systematic dissemination, storage and management of the output from the evaluation to ensure easy accessibility and to maximise the benefits of the evaluation's findings.	-	-

10. Completeness

	ToR	**Report**
10.1 Evaluation questions answered by conclusions The evaluation report answers all the questions and information needs detailed in the scope of the evaluation. Where this is not possible, reasons and explanations are provided.		X
10.2 Clarity of analysis The analysis is structured with a logical flow. Data and information are presented, analysed and interpreted systematically. Findings and conclusions are clearly identified and flow logically from the analysis of the data and information. Underlying assumptions are made explicit and taken into account.		X
10.3 Distinction between conclusions, recommendations and lessons learned Evaluation reports must distinguish clearly between findings, conclusions and recommendations. The evaluation presents conclusions, recommendations and lessons learned separately and with a clear logical distinction between them. Conclusions are substantiated by findings and analysis. Recommendations and lessons learned follow logically from the conclusions.		X
10.5 Clarity and representativeness of the summary The evaluation report contains an executive summary. The summary provides an overview of the report, highlighting the main conclusions, recommendations and lessons learned.		X

Annex 7.8
Gender checklist

Several levels

Management and/or implementation of the evaluation

- Men and women are represented in the evaluation team?
- Gender expertise in the team?
- Are gender issues or topics clarified and assigned within the team?
- Do the indicators and methods of the evaluation reflect the differentiation between men and women (Data for men and women are prepared and interpreted separately. Methodological details can be taken from the inception report)?

Evaluation results/analysis

- Were gender issues taken into consideration in the project and programme planning phase and the implementation?
  a) Is the design of the intervention gender-differentiated and is a gender analysis available?
  b) Do women and men contribute equally to the realisation of the intervention?
  c) Do women and men equally benefit from the intervention?
- What are the intended and unintended effects on men and women?
- Do the results, conclusions and recommendations in the evaluation report refer to gender issues?

Annex 7.9
Format for an inception report

The inception report should contain no more than 20–25 pages.

Introduction
… contains a short description of the background, purpose and scope of the evaluation according to the ToR.

Schedule
The schedule is described and possible deviations and adaptations are explained, as formulated in the Terms of Reference.

Activities
This section contains an overview of the activities already carried out, as listed in the ToR.

Preliminary hypotheses
Presentation of preliminary results on the basis of the five evaluation criteria and evaluation questions respectively, as listed in the ToR.

Methods
It is recommended to prepare an overview /matrix of the main evaluation questions with all the corresponding sub-evaluation questions, indicators, required data, data source, survey methods and the person in charge. Presentation of the data triangulation, data processing and quality assurance. Methodological details on the formulation of cross-cutting issues (particularly gender) and the extent of which the intervention logic will be analysed in the evaluation also need to be included in the inception report.

Further procedure
In this section, details on further activities, including field trips, interviews, discussions, surveys, reporting etc. are mentioned. The internal division of labour in the evaluation team should be clearly mentioned.

Annexes
Terms of reference, overview of documents used.

Annex 7.10
Format for a data collection planning worksheet

1. Major evaluation question	**Statement of data required**	**Data source**	**Survey method**	**Name of person who will collect the data**	**Beginning**	**End**
Subquestion 1						
Subquestion 2						
Subquestion 3						
etc.						
2. Major evaluation question						
Subquestion 1						
Subquestion 2						
Subquestion 3						
etc.						

Annex 7.11
Format for an evaluation report

The evaluation report should contain at most 50-60 pages without annexes.

Title page
Title of evaluation, date of completion of report, name of evaluators (of the institute), name of contractor.

Index, list of abbreviations, map

Executive summary
The evaluation report starts with an executive summary of three to five pages. The summary contains a brief overview of the purpose, objectives, scope, methods of the evaluation and refers to the most important recommendations, results and lessons learnt. If the evaluation report was prepared in German, an English executive summary also needs to be added to the German version. The executive summary must be written as an independent document so that it can be forwarded to third parties.

Background
In this chapter, the fundamental information on the project being evaluated are summarised, i.e. project and programme context (national, political, economic, social, cultural background), project and programme title, project and programme number, duration, name of project partner, location, costs, objectives, expected results and planned changes with regard to the target group (outcome), intervention logic or logframe respectively (the latter must be added to the annex), details on the target groups (number according to sex, ethnicity, etc.).

Introduction
… contains a brief description of the purpose, objectives and scope of the evaluation and briefly explains whether there have been any restrictions during the evaluation.

Methods
This section offers an overview of the quantitative and qualitative methods applied (including an overview and explanation on the number of the persons included per method, as well as criteria for selecting the project locations etc.). Techniques used during collection and processing of data and information (e.g. data triangulation) should be mentioned as well. The evaluation report also mentions possible restrictions (e.g. the non-availability of key informants) by using the methods as well as possible resulting effects on the evaluation, particularly its independence.

Evaluation findings
In this chapter, the evaluation findings are presented in detail. The evaluation report is structured according to the OECD/DAC criteria of relevance, effectiveness, efficiency, sustainability and impact as they are listed in the ToR. The evaluation questions and the corresponding results also need to be attributed to the OECD/DAC criteria. Results referring to the cross-cutting issues (poverty, gender, environment) also need to be considered under the OECD/DAC criteria or the evaluation questions, but additionally need to be described separately. Statements and conclusions must be comprehensible and be supported by data. Wherever it seems relevant, data must be presented and interpreted in a sex-disaggregated manner. Hypotheses must be verified and falsified.

Conclusions
… contain a summary of the results of all evaluation questions and, furthermore, inlcude all information issues (e.g. assessment of the intervention logic) which were mentioned under the scope of the evaluation. The conclusions are based on the results and the analysis, and are comprehensible on this basis. In case information is only presented partially, the reasons should be stated in the evaluation report.

Lessons learnt
Lessons learnt result from the conclusions and can be subdivided e.g. in strategic, policy, sector, management, implementation relevant lessons learnt and others.

Recommendations
In this chapter, recommendations are listed on the basis of the individual evaluation questions. It is important that the recommendations are feasible.,It must also be clearly identifiable to who the recommendations are addressed to. It is recommended to present the recommendations in a matrix.

Annexes
Logframe, terms of reference and schedule of the evaluation, list of key informants, list of documents used, questionnaires or other instruments used in the evaluation; Reports prepared for the field study; Information regarding the evaluators.

Annex 7.12

Format for a management response

In the action plan all recommendations of the evaluation are listed, discussed and documented. In case one recommendation will only be partially accepted or not accepted, then it is necessary to briefly write an explanation.

The status of the action plan should be regularly monitored (at least once a year) and documented.

The project partner is responsible for the preparation and implementation of the action plan.

The target is to have an instrument which documents the implementation of the evaluation recommendations.

Recommendation of the evaluation	Recommendation fully accepted	Recommendation partially accepted	Recommendation not accepted	Recommendation partially accepted or not accepted, please, explain reasons	Recommendation fully accepted, please, explain next step, activity, decision etc.	Date of Implementation	Name of the organization, department, or person responsible for the implementation of the recommendation	Current status of the implementation
1.	Yes/ No	Yes/ No	Yes/ No					
2.								
3.								
etc.								

Annex 7.13
Challenges for an evaluation or the evaluation team

- Different persons have different expectations from an evaluation. As a result, some evaluations are too complex and overloaded with too many questions.
- Evaluation questions are not clearly defined and leave too much room for interpretation.
- Evaluation questions are mixed with research questions so that the evaluation has more the character of a study.
- A lack of interest or resistance against the evaluation, e.g. because the latter was commanded “from the top“ and was not carried out in a participatory manner.
- Project or programme employees project their unsolved problems on the evaluation.
- Evaluation as "crisis intervention" – expectation that an evaluation will solve all problems.
- The decision of the project´s future has already been made prior to the evaluation. The evaluation is used to legitimise the end of a project or the continuation of a project.
- A lack of participation of the project stakeholders in the planning and implementation of the evaluation.
- The number of project documents, “grey“ literature and correspondence is too large or small.
- Uncertainties about the responsibilities (e.g. support, logistics on the spot and accountability at the end of the evaluation process).
- Evaluators have to write too many reports, e.g. monthly activity reports.
- The contact between the evaluators and the employees of the respective project or programme is too close or too little.
- Evaluation results are not presented or not sufficiently discussed with relevant partners.
- No management response was prepared, therefore it is not comprehensible how the evaluation results were used.
- Evaluation results are forgotten in the daily routine of the project.
- Evaluation results are not taken into consideration in the next project planning cycle.

Annex 7.14
Further literature and internet addresses

Austrian Development Agency (ADA). Geschlechtergleichstellung und Empowerment von Frauen. Leitlinien der Österreichischen Entwicklungs- und Ostzusammenarbeit. Wien. April 2006.
http://www.entwicklung.at

Beck, Tony, Evaluation Humanitarian Action using the OECD-DAC Criteria. An ALNAP Guide for Humanitarian Agencies.
http://www.alnap.org/

Canadian International Development Agency (CIDA), CIDA Evaluation Guide. Ottawa. January 2004.(Guide in Englisch, Französisch, Spanisch)
http://www.acdi-cida.gc.ca/index-e.htm

Department for International Development (DFID). Guidance on Evaluation and Review for DFID Staff. London. July 2005.
http://www.dfid.gov.uk/

European Commission. Evaluation Methodology for European Commission's External Assistance. Luxenburg. 2006.
http://ec.europa.eu/europeaid/how/index_de.htm

European Commission. Toolkit on Mainstreaming Gender Equality in EC Development Cooperation. Brussels. 2004.

Gesellschaft für Evaluation (DeGEval). Standards für Evaluation. Köln. Juni 2004. http://www.degeval.de/

Hughes, Jenny und Baumgartl, Bernd: Project Manager's Guide to Evaluation. Navreme. Wien 2005.

OECD. DAC Evaluation Network. Encouraging Effective Evaluation of Conflict Prevention and Peacebuilding Activities. OECD. 2007.
http://www.oecd.org/site/0,3407,en_21571361_34047972_1_1_1_1_1,00.html

OECD. DAC Evaluation Network. DAC Evaluation Quality Standards. OECD. March 2007
http://www.oecd.org/site/0,3407,en_21571361_34047972_1_1_1_1_1,00.html

OECD. DAC Evaluation Network. Guidance for Managing Joint Evaluations. OECD. 2006. (Guide in Englisch und Französisch)
http://www.oecd.org/site/0,3407,en_21571361_34047972_1_1_1_1_1,00.html

OECD. DAC Evaluation Network. Glossary of Key Terms in Evaluation and Results Based Management. OECD. June 2002.
http://www.oecd.org/site/0,3407,en_21571361_34047972_1_1_1_1_1,00.html

Oxfam. A Guide to Gender-Analysis Frameworks. Oxford.1999.
http://www.oxfam.org.uk/

Swedish International Development Cooperation Agency (SIDA). Looking Back, Moving Forward. SIDA Evaluation Manual. Stockholm. 2004.
http://www.sida.se/

United Nations Population Fund (UNFPA)
The Programme Manager's Planning, Monitoring and Evaluation Toolkit. New York. 2004 (Guide in English, French, Spanish, Arabic)
http://www.unfpa.org/

United Nations Development Programme (UNDP). Handbook on Monitoring and Evaluation for Results. New York. 2002. (Guide in English, French, Spanish)
http://www.undp.org/

United Nations Development Programme (UNDP). Managing for Results: Monitoring and Evaluation in UNDP. A Results-Oriented Framework. New York. November 2001

Internet addresses of some evaluation societies

African Evaluation Association: http://www.afrea.org/home/index.cfm

American Evaluation Association: http://www.eval.org/

European Evaluation Society: http://www.europeanevaluation.org/

Gesellschaft für Evaluation: http://www.degeval.de

Other internet sites

The Active Learning Network on Accountability and Performance in Humanitarian Assistance:
http://www.alnap.org/

Direktion für Entwicklung und Zusammenarbeit (DEZA):
http://www.deza.ch/de

ELDIS: http://www.eldis.org/

The Evaluation Center. Western Michigan University:
http://www.wmich.edu/evalctr/

European Commission (EC):
http://ec.europa.eu/europeaid/evaluation/index_en.htm

The International Development Research Centre (IDRC):
http://www.idrc.ca/en/ev-26266-201-1-DO_TOPIC.html

MANDE News: http://www.mande.co.uk/

OECD/DAC:
http://www.oecd.org/site/0,3407,en_21571361_34047972_1_1_1_1_1,00.html

Performance Assessment Resource Centre (PARC):
http://www.parcinfo.org/

The Worldwide Evaluation Information Gateway:
http://www.policy-evaluation.org/

The World Bank Group. Independent Evaluation Group (IEG):
http://www.worldbank.org/oed/