



Ministry of Foreign Affairs

Evaluation policy and guidelines for evaluations

IOB | October 2009 | Policy and Operations Evaluation Department | IOB | October 2009 | Policy and Operations Evaluation Department | IOB | October 2009 | Policy and Operations Evaluation Department

Evaluation policy and guidelines for evaluations

Contents

1	Introduction	4	
2	Short history	6	
3	Basic principles of IOB's evaluation policy	10	
4	Definitions and criteria for evaluations	14	
4.1	What is evaluation?	15	
4.2	Types of evaluation	16	
4.3	Evaluation criteria	17	
4.3.1	Efficiency	17	
4.3.2	Effectiveness	18	
4.3.3	Impact	19	
4.3.4	Relevance	19	
4.3.5	Sustainability	20	3
4.3.6	Policy consistency	21	
5	Guiding principles and ethical standards	22	
6	The evaluation process	26	
6.1	Programming	27	
6.2	Preliminary study	29	
6.3	Terms of Reference	30	
6.4	Deploying researchers	32	
6.5	Cooperating with researchers in partner countries	32	
6.6	Methodology	33	
6.7	Reporting	37	
7	Quality control	40	
7.1	Internal peer review	41	
7.2	Reference groups	42	
7.3	Departmental meeting	42	
7.4	Advisory panel	43	
7.5	Self-evaluation	43	
8	Communication and feedback	44	

1 Introduction

This document describes the evaluation policy and guidelines for evaluations of the Policy and Operations Evaluation Department (IOB) of the Dutch Ministry of Foreign Affairs. It places the evaluation policy in the context of the *Order on Periodic Evaluation and Policy Information* (known as RPE 2006)¹, the government-wide framework for the evaluation of government policy in general. It also reflects recent decisions regarding the function of evaluation at the Ministry of Foreign Affairs, and the role IOB plays within it.

IOB's evaluation policy and the guidelines must be seen in the context of a changed environment. An increasingly critical outside world demands a clearer insight into the results of the Netherlands' foreign policy. There is growing pressure for ministers to account not only for the use of public funds, but even more so for the results achieved. This also applies to the various areas of foreign policy. If evaluations are to contribute to this, they must be independent, public, transparent and accessible.

In publishing these guidelines, IOB wishes to provide transparency to third parties: colleagues at the Ministry of Foreign Affairs, the Dutch parliament, organisations and institutions affected by IOB evaluations, users of the evaluation results, and interested citizens. In each evaluation report, IOB accounts for the specific evaluation; this document clearly presents the general principles and guidelines IOB applies internally in conducting evaluations.

5

The guidelines outlined in this document apply in principle to all evaluations conducted by IOB, unless explicitly stated otherwise. It should be borne in mind that the Netherlands has a long tradition of evaluation in the field of development cooperation, which is expressed through international agreements, fora, cooperation, etc. and which impacts on IOB's work. That is not the case for other areas of foreign policy. Because its mandate also extends to these other areas of policy, IOB is in an exceptional position, especially internationally.

The evaluation policy and guidelines reflect the opinions and perspectives that underpin IOB's work. In that sense, this document describes the ideal situation for IOB. Reality will sometimes deviate from this ideal in some ways. There will always be a certain tension between theory and practice. However, IOB aims to work as much as possible according to these guidelines. It is important to remember that neither policy nor guidelines are carved in stone. This must be seen as a growth document, in the most literal sense. New insights may lead to it being adapted in the future. IOB will thus periodically review its policy and the guidelines and, if necessary, revise them. The latest updated version of this document will be available at <http://www.minbuza.nl>.

1 *Order on Periodic Evaluation and Policy Information 2006*, Government Gazette, 26 April 2006, no. 83.

2

Short history

In the first half of the 1970s the budget for development cooperation increased rapidly. So rapidly that concerns were voiced from various quarters, and especially in the media, that this was having an adverse effect on the quality of aid. As a consequence of this negative publicity, the Minister for Development Cooperation felt the need for an independent opinion on development policy, autonomous of the policy departments involved. On 1 July 1977 the Operations Review Unit (IOV) was set up by ministerial order. It was made clear from the beginning that IOV would focus on bilateral cooperation, both governmental and non-governmental (in so far as the latter was government-funded), but would in principle restrict its activities to the Netherlands' contribution to joint interventions. On the basis of this restriction, the minister took the position that IOV could conduct inspections without the permission of the recipient country or organisation. IOV's work programme was drawn up together with the policy departments and submitted to the minister for approval.

Outside the work programme the minister and the ministry management were also authorised to commission IOV to conduct inspections on an ad hoc basis. However, this occurred only sporadically. From the start IOV focused on individual projects, submitting confidential reports to the minister. The reports were confidential because they referred to identifiable events and individuals.

7

From 1979 the House of Representatives of the Dutch parliament called each year for IOV's reports to be made as public as possible. Eventually, it was agreed that general reports would be sent to parliament, together with summaries of project inspections. The first summarised report that IOV sent to the House was the 1984 General Evaluation of Dutch Bilateral Development Cooperation.

In the 1980s, IOV's area of work shifted from the inspection of individual projects to evaluations of sectors, policy themes and country programmes. During verbal consultations with the House of Representatives in 1987 the minister agreed that all IOV reports would be fully public and would be sent to parliament, together with a policy letter from the minister containing his or her response to the report. The reports were used actively in debates by members of both the House of Representatives and the Senate. Universities and other institutions in the public and private sectors also showed a lively interest in the reports once they became accessible.

In 1996 a far-reaching reassessment of Dutch foreign policy took place. It was widely acknowledged that development cooperation, foreign policy and international trade could no longer be seen as independent policy areas. Foreign policy was 'decompartmentalised' and the various areas of international activity were integrated. For the purpose of the review, IOV published a new synthesis study, summarising the results of the evaluation reports of the preceding decade. During the reorganisation that accompanied the policy reassessment, it was logical to modify IOV's mandate. Whereas, before the review, IOV was only responsible for evaluations in the field of development cooperation, its area of operations was now extended to include the

whole of foreign policy. This implied that, in the future, the Unit could address itself to all subjects that fell under the Homogenous Budget for International Cooperation (HGIS). The change in IOV's mandate led to a revision of its position in the organisational structure of the ministry. It would no longer be accountable to the Director-General for International Cooperation (DGIS) but directly to the Secretary-General. The Unit was also given a new name: in the future it would be known as the Policy and Operations Evaluation Department (IOB).

In recent years IOB's position and role has once again been the subject of political debate and discussion within the ministry. Many of the issues discussed in the 1980s were once again on the agenda, including the independence of evaluation, the public nature of the reports, the relationship between external accountability and internal learning. This led in 2008 to the Minister for Development Cooperation reconfirming to parliament the independence of IOB and the public nature of its reports. However, IOB's programming authority was restricted to development cooperation.² In the other fields of foreign policy, primary responsibility now lies with the Director-General concerned, who can request IOB to conduct evaluations.

2 Letter from the Ministers of Foreign Affairs and for Development Cooperation of 16 October 2007 on implementation of the Boekestijn/Gill'ard motion regarding the Policy and Operations Evaluation Department (IOB) (31200V, no. 9). Report from the Permanent Committee on Foreign Affairs on IOB of 29 May 2008, approved on 25 June 2008.

3

Basic principles of IOB's evaluation policy

The Policy and Operations Evaluation Department is the evaluation department of the Ministry of Foreign Affairs. Institutionally, IOB is directly accountable to the Secretary-General. It forms part of the ministry's Audit Committee. Together with the Audit Department and the Inspection and Evaluation Unit (ISB), IOB forms the cluster 'In Control'.

Internationally, IOB is a member of the OECD/DAC network of evaluation services. It also takes part in consultations between the evaluation services of the European Union, and of the countries united in the Nordic+ group. IOB contributes to the activities of a number of other forms of international cooperation, including the consultation committee of German-speaking evaluation services (DACH), the Network of Networks for Impact Evaluation (NONIE) and the International Initiative on Impact Evaluation (3IE).

IOB's mandate is as follows:

The Policy and Operations Evaluation Department meets the need for evaluation of policy and its implementation with regard to all policy areas that are financed by the homogenous budget for international cooperation (HGIS). IOB evaluates the implementation of the Netherlands' foreign policy, giving particular attention to its relevance, efficiency and effectiveness, the sustainability of the results achieved and the consistency and coherence of the policy.

11

In the light of this mandate, IOB has the following tasks:

- 1) to conduct high quality evaluations for learning and accountability purposes;
- 2) to ensure balanced evaluation programming in the field of development cooperation and to advise on programming in other areas of foreign policy;
- 3) to contribute to improvement of the quality of evaluations conducted by, or on behalf of, policy departments, institutions subsidised by the ministry, and multilateral institutions by drawing up generally applicable quality requirements and guidelines for conducting evaluations, advising on the design and implementation of individual evaluations, and conducting systematic assessments of these evaluations or peer reviews of other evaluation services;
- 4) to strengthen capacity to conduct evaluations in partner countries by supporting education and training and participating in joint evaluations;
- 5) to promote the distribution and communication of evaluation results.

Re 1:

IOB's main task is to conduct high-quality evaluations within the framework of the Ministry of Foreign Affairs' evaluation programme. These evaluations can relate to all areas of foreign policy that fall under the HGIS. In this context, evaluation comprises critical and thorough investigation of a policy or programme and explaining the results. IOB evaluations do not stand in isolation but are in principle part of the policy cycle, in which planning, monitoring and evaluation are seen as an integrated whole.

Evaluation is therefore an essential component of the 'knowledge function' within the ministry. Knowledge on the implementation and impact of policy supports two concrete functions:

- it provides the basis for accounting to parliament and society at large for policy and expenditure (accountability function);
- it generates information for the purposes of improving policy formulation and implementation (learning function).

Re 2:

As stated in the RPE 2006 systematic evaluation must be a fixed component of every policy process. Each policy area must be periodically evaluated at the level of operational objectives. The RPE 2006 thus forms the basis of evaluation programming. The programming is also determined by specific requests from parliament, ministers and state secretaries, and budget holders. Within the Ministry of Foreign Affairs, IOB has final responsibility for programming evaluations in the field of development cooperation and advises on programming in other areas of foreign policy.

12

Re 3:

IOB is not the only unit at the Ministry of Foreign Affairs concerned with evaluation. Policy advisors also have to deal with evaluations, as the contact for evaluation programming within their policy department, as a 'client' commissioning an evaluation, sometimes as (joint) implementer, or as a user of the results. In addition to conducting evaluations, IOB therefore has an advisory function. This entails formulating quality criteria and guidelines for implementing evaluations and advising on the Terms of Reference for evaluations that IOB does not conduct itself. During the annual settlement of the programme, agreements are made regarding the extent to which IOB will play an advisory role in evaluations conducted by or on behalf of ministry departments. IOB also advises on the design of the ministry's evaluation function, the evaluation system and instruments, and coordination with planning and monitoring. In addition IOB performs quality tests on programme evaluations conducted by cofinanced non-governmental organisations (MFS), with the aim of strengthening their evaluation capacity.

For the multilateral channel IOB is involved in international reviews of the evaluation services of relevant institutions.

Re 4:

Strengthening evaluation capacity in partner countries is an important theme in development cooperation. Independent evaluation capacity in partner countries is indispensable if the aim of establishing a system based on mutual accountability is to be achieved in reality. IOB sees it as its task to contribute to this by supporting international initiatives aimed at strengthening evaluation capacity. In addition, where possible and safeguarding its own criteria for independence and quality, IOB works together with researchers from partner countries in conducting evaluations.

Ad 5:

Evaluations generate knowledge on the basis of experience from the past regarding what works and what does not, and especially why. IOB considers it important that the results of evaluations are distributed as widely as possible. To achieve this, it discusses the findings in staff meetings, with ministers and state secretaries, and with other national and international parties involved in the evaluations. In addition, where possible and where it does not conflict with its own responsibilities, IOB will offer support in preparing policy responses and descriptions of evaluation findings in the Ministry of Foreign Affairs' Annual Reports.

4

Definitions and criteria for evaluations

4.1 What is evaluation?

There is a wide variety of definitions of the concept of evaluation. In the 1980s and 1990s, the OECD/DAC worked on a shared framework of concepts for evaluation practices in the field of development cooperation. The resulting definition of evaluation serves as a basis for IOB's work.

'An evaluation is an assessment, as systematic and objective as possible, of an ongoing or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, developmental 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.'³

The OECD/DAC definition reflects an understanding of evaluation that is broadly supported in the international community of evaluators. The definition is intended for evaluations in the field of development cooperation. Since IOB's work extends further than development cooperation, concepts like 'developmental' and 'recipients and donors' will not apply to all IOB evaluations.

15

Evaluation is different to reviews and monitoring. The OECD/DAC Glossary describes a review as an assessment of the performance of an intervention, periodically or on an ad hoc basis. For IOB the main difference between a review and an evaluation is that a review is considered a management instrument for operational monitoring of the implementation of an activity, and therefore focuses on operational aspects and recording results, while evaluation is a more in-depth assessment of the aims, implementation and results of an intervention, with the question of attribution as its main focus.

Monitoring is an instrument of internal management. It aims, on the basis of a fixed set of indicators, to provide a regular insight into the progress made in implementing activities. Monitoring allows a budget holder or the implementer of an intervention to stay up to date on inputs supplied and activities performed (outputs). Monitoring shows whether intended activities have actually been carried out. Good, systematic monitoring is of great importance for evaluations, as it provides a significant part of the data on which evaluation is based.

3 DAC principles for Aid Evaluation, OECD/DAC, 1991 (the definition was also included in the DAC Glossary of Key Terms in Evaluation and Results-based Management in 2002).

4.2 Types of evaluation

The Order on Periodic Evaluation and Policy Information 2006 (RPE 2006) prescribes a system of evaluation based on the so-called operational objectives, OD. The system entails periodical evaluation of these operational objectives.

The RPE 2006 distinguishes between policy reviews and ex-post impact studies. Both concepts deviate from the evaluation terminology used internationally, but are the equivalent of the more common terms ‘policy evaluation’ and ‘impact evaluation’. The starting point is that policy reviews, which are linked to the operational objectives, are as far as possible fed by impact studies. In principle they resemble synthesis studies. IOB gives priority to these types of evaluation, where – especially in the field of development cooperation – policy reviews are preceded by a series of ex-post impact studies (impact evaluations).

16

On the basis of the system prescribed in the RPE 2006, IOB impact evaluations will generally focus on interventions (ranging from programmes to various forms of budget support). Policy evaluations (or reviews) address themes or sectors. In addition, integrated foreign policy can also give rise to ‘decompartmentalised’ evaluations of country or regional policies.

Besides these – usually extensive – evaluations, IOB can also conduct short internal evaluations that are limited in nature and necessary to monitor the progress of programmes.

Part of the evaluations in IOB’s programming focuses solely on Dutch programmes and activities. It is logical for these evaluations to be conducted by IOB. However, it is becoming increasingly common to work together with international partners in the field of development cooperation. Projects by individual donors are gradually being replaced by joint financing of government policy in partner countries. It is then also logical that donors give greater preference to the joint programming and implementation of evaluations. In cases where donors work together, for example in providing joint support for sector policy or other far-reaching forms of joint financing (such as basket-funding), IOB evaluations will in principle be conducted in cooperation with other donors, as long as IOB’s criteria for independence and quality are met. The nature of the cooperation can vary from actual participation by several donors in an evaluation, agreeing on a division of tasks within a shared Terms of Reference, to agreement on the division of labour between different themes, sectors or countries.

4.3 Evaluation criteria

Evaluation calls for a tailor-made approach. That does not mean, however, ignoring quality criteria or abandoning consistency in approach. In all evaluations, the evaluation criteria must be applied in an unambiguous and, above all, transparent way. Here, in particular, standardisation is paramount to enable evaluations and their results to be compared. The evaluation criteria applied by IOB are largely identical to those drawn up by the OECD/DAC. The OECD/DAC definition of evaluation specifies five criteria for evaluation of development interventions: efficiency, effectiveness, impact, relevance and sustainability. The evaluation criteria refer to different result levels founded on the logical framework or result chains based on it. Together, these criteria provide a framework within which researchers can present their findings in an ordered and coherent way. The extent to which equal attention can or will be paid to all criteria will vary per evaluation. Furthermore, additional criteria can be applied: evaluation of ‘decompartmentalised’ policy, for example, may require that IOB devotes explicit attention to policy consistency.

4.3.1 Efficiency

Efficiency measures how economically resources (inputs) and the way they are applied are converted to direct results. The concept therefore illustrates the relationship between input and output and refers to a level of results that the implementer can check.

17

In evaluating efficiency, both the quantity and quality of inputs and outputs are assessed. The most economical or cost-effective input is not always the most appropriate and there can be significant tradeoffs between the quantity and quality of outputs. Assessing efficiency also calls for comparison with alternative approaches, with which the same outputs can be achieved. This is by no means always easy because there are often no previously formulated benchmarks and they might have to be developed on the basis of practical experience. That is more likely in the case of, for example, interventions involving construction activities than of those aiming to bring about social or institutional change.

The following elements should be taken into account when evaluating efficiency:

- Have appropriate inputs been deployed at the lowest possible cost?
- Have activities been conducted in an uncomplicated manner?
- Have decisions been made at the right level and has bureaucracy been avoided as far as possible?
- Have overhead costs been kept to a minimum?
- Has duplication been avoided?
- Have conflicts during implementation been prevented or solved?
- Have outputs been achieved within the planned period and budget?

4.3.2 Effectiveness

Effectiveness relates to the extent to which the direct results of interventions (output) contribute to the sustainable achievement of policy objectives (outcome). Policy objectives are the objectives that interventions are intended to achieve and should, in principle, be formulated in the project document. In some cases policy objectives can be identical to the operational objectives specified in the budget or specific components of them.

An intervention is considered effective if its outputs have made a demonstrable contribution to achievement of the intervention's intended objectives. Efficiency and effectiveness refer to two successive levels in the results chain. A positive assessment at one level does not automatically imply a positive assessment at the other. An intervention can be efficient without having an effect, and an effective intervention can be inefficiently implemented.

The effects of an intervention are beyond the direct control of the implementer. Interventions generate outputs which are assumed to generate the desired effects. Evaluating effectiveness aims to establish this causality. The core problem is that it would be ideal to know what would have happened if the intervention had not occurred (the counterfactual). Only then can observed changes be attributed to the intervention and is it possible to speak of the net effects. In other words, there is a problem of attribution. As long as other factors have a visible impact on the observed changes, the problem is relatively simple to solve. It is more difficult when influential factors are not immediately visible (the 'unobservables'). Researchers cannot know what they do not know. Evaluations try to find a solution to this problem (see 6.6).

18

In practice – especially in the case of large and complex interventions, but also with interventions that aim to bring about effects that are difficult or impossible to quantify – evaluations often make use of plausibility argumentation. It is then important that the argument is transparent and credible.

In essence, three steps can be distinguished in the evaluation of effectiveness:

- The measurement of changes in the effect variables in comparison to the situation at the start (baseline);
- Attribution of the observed changes to the intervention;
- Assessment of changes observed and attributed to the intervention in terms of the objectives.

4.3.3 Impact

The broad consensus in the community of evaluators about the concepts of efficiency and effectiveness is less evident when it comes to impact and relevance.

The OECD/DAC defines impact as all significant effects produced by a development intervention, directly or indirectly, intended or unintended, on the ultimate stakeholders and third parties. While effectiveness refers only to the intended effects of an activity, in this terminology impact refers to all possible economic, social, political, technical and ecological effects at local, regional or national level that have a direct or indirect effect on the target group or other parties.

In the stricter context of the logical framework, the concept of impact usually refers to the highest level of results. The question is then to what extent, with the aid of outputs, the effects achieved (outcomes) have contributed to broader, more far-reaching objectives. These can be operational objectives or objectives at a higher level, such as more general poverty reduction.

In the international literature on impact evaluations, the concept of impact is used for the net effects attributed to the intervention, i.e. at the level of effectiveness.

19

One interpretation of impact is not by definition better than another. It is important that concepts are made explicit in evaluation reports. In the interest of mutual comparison it is desirable in ex-post impact studies to use the interpretation of impact evaluations most commonly applied in the international literature. In this respect IOB follows the instructions in the RPE 2006.

4.3.4 Relevance

The concept of relevance is also subject to different interpretations. The OECD/DAC defines relevance as the extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, global priorities and partners' and donors' policies. That makes relevance a container concept, within which no choices are made. This definition is unsatisfactory for several reasons. First, the assessment of whether an intervention is a logical consequence of or is compatible with policy is in essence an ex ante question that should be part of the approval procedure. Second, donor policy, the partner's policy and the needs of the target group do not necessarily all perfectly coincide. That would imply that an intervention can be relevant and irrelevant at the same time. Third, policy priorities can change over time, such that a specific intervention is relevant at one moment and not the next, or vice versa. Lastly, this definition implies that the relevance of the policy itself is not up for discussion. In other words, it excludes problems with policy itself.

In the logical framework, relevance refers to the extent to which the effects of interventions make a sustainable contribution to achieving the ultimate objective (the impact). An intervention has been valuable, or relevant, to the extent that it has

generated effects that bring achievement of the ultimate development objective closer. IOB will apply the concept of relevance in its evaluations in this respect.

4.3.5 Sustainability

Since an activity can hardly be considered effective if the effect it has achieved is not lasting, sustainability is in fact an aspect of effectiveness.⁴ At the same time the concept of sustainability comprises a variety of dimensions that are relevant for assessing an intervention. In the donor community the sustainability of interventions is linked to seven factors, which can be used to analyse interventions. Which factors or combinations of factors are useful for assessing sustainability will depend on the nature of the intervention.

- The involvement of the recipient is usually recognised as a factor that affects the success of an intervention. Involvement is expressed through agreement on objectives and the provision of human and material resources;
- When deploying technology it is important to take account of the financial and institutional capacities of the recipient and the degree to which the conditions for maintenance and renewal are guaranteed;
- Interventions can have an effect on the environment in the short or long term. While short-term effects are usually envisaged and are prevented or mitigated, it is especially important to incorporate long-term environmental effects in the assessment, as they can neutralise positive changes or have other harmful consequences;
- The degree to which changes are sustainable depends very strongly on social and cultural factors, which are often not directly visible. Studying perceptions can provide better insight into this dimension of sustainability;
- Institutional and capacity development play an important role in determining sustainability. When assessing interventions the attention devoted to these aspects will help determine the extent to which sustainability is achieved;
- Financial and economic sustainability refers to the degree to which measures have been taken to guarantee that activities can continue and completed works can be maintained in the future;
- Interventions always take place in a wider natural, political, economic and socio-cultural environment, which cannot be influenced in any way at all but which itself has a strong influence on effectiveness and sustainability. Changes in political stability, economic crises, or natural disasters can have a decisive effect on the sustainability of interventions.

20

⁴ In the context of humanitarian aid this is, of course, different. It usually entails short-term, punctual interventions, the effects of which it is hoped will only be necessary for a brief period.

4.3.6 Policy consistency

This concept refers to the absence of inherent contradictions between different policy areas. Consistency is particularly important for IOB in the evaluation of decompartmentalised policy. Development objectives can conflict with international trade interests or foreign policy considerations. That is why it is important, particularly in cases where objectives have to be achieved in the context of decompartmentalised policy, to determine the extent to which interventions performed from different perspectives are coherent with and do not obstruct each other. The aim is to help achieve intended objectives without the effects of one intervention neutralising those of another.

5

Guiding principles and ethical standards

IOB's guiding principles reflect the core values on which its evaluations are founded. They are interrelated and mutually reinforcing and, as such, they must be seen as providing an overall framework, with which all aspects of IOB's evaluation work must be consistent. The principles determine what can be expected from an inspector and other researchers involved in an evaluation, from evaluation processes and products, and from the methodology. They are consistent with the OECD/DAC's Principles for Evaluation of Development Assistance (1991) and Evaluation Standards (draft 2009).

All IOB evaluations must be consistent with the following guiding principles:

Independence

The judgements of the researcher must not be influenced by external pressure or a conflict of interests. Organisations or institutions that are commissioned to perform an evaluation and members of evaluation teams may not have been previously involved in the interventions that are to be evaluated or in formulation of the policy on which they are based.

Objectivity

The evaluation must be based on verifiable facts. The researcher must make every effort to ensure that the data on which the evaluation is based does not contain any inaccuracies. The presentation of facts should be clearly and recognisably distinguished from judgments.

23

Transparency

The evaluation must be conducted in a way that can be followed clearly by stakeholders and third parties. The questions to be addressed, the data base, the research methods, findings and conclusions must be presented in the report in a clear and accessible way.

Validity

The evaluation must measure what the Terms of Reference specifies as requiring measurement, and must do this in a way that the reader can understand.

Reliability

The results of the evaluation should be stable, i.e. that a repeat of the evaluation will produce the same results and conclusions.

Partnership

In the case of evaluations in the field of development cooperation, where possible and in so far as it does not conflict with other principles, partners from the recipient countries will be involved in the implementation of evaluations. Every effort will be made to cooperate with independent researchers from the partner countries.

Feasibility

Evaluations are only conducted if the data base is accessible and the associated methodology is available.

Correctness

The evaluation must not cause harm to individuals.

Cost effectiveness

The evaluation must be conducted at the lowest possible price, taking account of the other guiding principles.

Usability

The evaluation must result in conclusions and points for attention that are valuable for the user of the evaluation by suggesting actions that lie within the user's capacity.

Evaluation is a form of research that entails a judgment. It differs from pure academic research in that it always results in an opinion on the results of concrete human actions. Researchers therefore have a great responsibility towards stakeholders involved in the evaluation. This applies equally to IOB inspectors and to external researchers contracted by them. Researchers can therefore be expected to adhere to internationally accepted ethical standards for evaluations.

24

The following ethical standards apply to IOB's work⁵:

Respect and understanding for other cultures

Many evaluations are conducted in a different cultural environment to that of the researchers. They can therefore be expected to comply with local cultural codes during interviews and other forms of data collection.

Universal values

Some cultural customs and practices may be incompatible with fundamental rights and freedoms. Although researchers should in principle show respect for other cultures, they must not ignore the effect of certain cultural values on gender relations or minorities and other specific groups. In such cases they should act in accordance with the Universal Declaration of Human Rights and the associated international agreements.

5 The standards are largely based on the evaluation policies of the Danish International Development Agency (DANIDA) and the Canadian International Development Agency (CIDA).

Gender

Evaluations should be carried out taking account of the different roles performed by men and women and the aim of achieving gender equality. In policy areas where gender differences are a dimension or effect of the policy pursued, they should be taken into account in the design and implementation of the evaluation.

Anonymity and confidentiality

Researchers are expected to respect requests from informers to treat the information they provide with confidentiality and to take steps to ensure that confidential information cannot be traced back to the source.

Responsibility

In many countries a critical stance regarding the government can have severe consequences for researchers from the country itself. In evaluations involving local researchers it is important that such risks be recognised in time and steps taken to avoid them, for example, by making clear agreements about who is responsible for the content of the report.

Omissions

Researchers may obtain information that is not directly important in answering the questions addressed in the evaluation, but is nevertheless relevant in the light of the accountability and learning functions of the evaluation. In such cases, the client (in the last instance, the IOB Director) will be consulted to determine whether this information shall be included in the report and in what form.

25

Assessment of individuals

Evaluations are not assessments of individuals. Yet evaluation reports can include findings on leadership and the management or other qualities of implementers that come close to personal assessments. Researchers should then make every effort to present the facts of easily traceable cases in an abstract form. If in doubt, they should consult with the IOB Director.

Evidence of malpractice

Evaluations are not investigations into legitimacy. Nevertheless it is possible that malpractices come to light during evaluations. In that case the researcher should submit the case to the IOB Director. If there is sufficient reason to do so, the Audit Department (ACD) will be informed.

6

The evaluation process

6.1 Programming

The quality of evaluation benefits from a careful and transparent decision-making process. This process starts with the programming of evaluations. IOB has final responsibility for programming evaluations in the field of development cooperation. That means that the IOB Director approves the programming after it has been discussed with the Financial and Economic Affairs Department (FEZ) and the other policy departments involved. With respect to the other areas of foreign policy, IOB advises on the programming, which is approved in the meeting of the Secretary-General and the Directors-General. IOB applies the basic principles laid down in the RPE 2006 for systematisation of the programming.

The evaluation programming distinguishes between three types of evaluation: policy reviews (policy evaluations), ex-post impact studies (impact evaluations) and other evaluations. Policy reviews and ex-post impact studies fall under the RPE 2006. The RPE prescribes that each policy area must be periodically subjected to a policy review or policy evaluation at the level of operational objectives. A policy review should make as much use as possible of ex-post impact studies. Policy reviews and impact studies largely determine the IOB's research agenda.

27

In addition to the RPE 2006 there can be other reasons for including evaluations in the programming.

First, there are political considerations. At certain times policy issues can be the subject of political or social debate. It is then important to feed the debate with insights obtained through thorough and independent evaluation.

Policy relevance can also play a role. This is determined by the priority given to an issue and the moment in which the evaluation is carried out. If it is clear when results are required for the formulation of new programmes or policies and an evaluation can be conducted within that period without loss of quality, IOB ensures that the evaluation is programmed and implemented in time, irrespective of the scale or complexity of the subject.

In the field of development cooperation there is international pressure to improve coordination between the evaluation programming of individual donors and to conduct joint evaluations as far as possible. Joint evaluations include those carried out together with evaluation services in the recipient countries, multidonor evaluations, and various combinations of the two. Working together with evaluation services in recipient countries helps strengthen their evaluation capacity and improve mutual accountability. The choice of which joint evaluations IOB participates in will depend on initiatives developed during international consultations, opportunities to enter into strategic alliances with like-minded evaluation services, the extent to which IOB's criteria for independence and quality are guaranteed, and whether they can be fitted into IOB's existing programming.

IOB also considers it important to allow space in its programming for ad hoc requests. These may be requests from within the ministry for IOB to conduct short evaluations of a limited nature to determine the progress of programmes or to support the decision-making process. In addition, ministers, state secretaries or parliament can request IOB to conduct specific evaluations. IOB makes every effort to comply with these requests.

The evaluation programming in the field of development cooperation relates not only to bilateral cooperation but also to the non-governmental and multilateral channels. In the context of the cofinancing system (MFS) IOB conducts systematic assessments of programme evaluations carried out under the responsibility of the co-financed non-governmental organisations themselves. In addition, IOB selects a number of themes or countries to conduct its own evaluations. To date, IOB has been somewhat reserved in programming evaluations of multilateral cooperation. Together with other donors, IOB has encouraged multilateral institutions to improve the quality of their own evaluations through peer reviews. In the future, where possible, IOB will participate in joint donor initiatives aimed at a more systematic evaluation of multilateral development cooperation.

28

The decompartmentalisation of foreign policy and the subsequent broadening of IOB's mandate mean that IOB evaluates not only interventions in the field of development cooperation, but also activities in other areas of foreign policy. It is not unusual for other foreign policy instruments to be deployed to address a specific theme, alongside development cooperation. This can be good reason to conduct a 'decompartmentalised' evaluation. In many cases, only the Ministry of Foreign Affairs is involved in such an evaluation, but not always. There are also issues of foreign policy in which other ministries play a prominent role. For IOB that is reason to explore the possibilities for conducting evaluations jointly with other ministries. In such cases, and if IOB's conditions regarding independence and quality are met, these evaluations will be included in the programming.

6.2 Preliminary study

Once a subject for evaluation has been included in the programming a preliminary study can be initiated. Especially in this initial phase, a number of important decisions have to be made that will have far-reaching significance for the further progress of the evaluation. The inspector will have to become familiar with the policy area to be evaluated and the available sources of information, and form some idea of the questions to be addressed in the evaluation and the most adequate approach. With large, complex evaluations the inspector can draw up a start-up report. The report describes the reason for the evaluation and the objective. It can also give a brief description of the policy area to be evaluated and formulate core questions for the preliminary study. It can suggest which type of evaluation would be most appropriate and how it might be organised (the expected duration, whether field missions are considered necessary, and when and how external consultants might be needed). Lastly, the start-up report can offer an insight into the expected products of the evaluation and the budget.

Start-up reports are discussed with the IOB management and the internal peer reviewers designated for the evaluation (see 7.1). After decisions have been made on the abovementioned main points, the preliminary study will move on to drafting definitive Terms of Reference for the actual evaluation.

29

The preliminary study comprises a number of activities:

Reconstruction of policy theory and intervention logic

The extent to which a subject can be evaluated is improved by clearly formulated policy principles and programme objectives, a description of the activities undertaken, and a logical relationship between policy, objectives and activities. Practice shows that these criteria are by no means always met. Generally speaking, this applies to both development cooperation and other areas of foreign policy. Especially at the level of operational objectives, policy is often broad-ranging and not very specifically formulated, aims and effects are frequently unclear and the logic of interventions is not sufficiently explicit. In such cases, on the basis of the preliminary study, inspectors themselves will develop a policy theory and a reconstructed intervention logic based on it. They will then present the theory and the logic to the policymaker to determine whether they are in line with what the latter had in mind.

Defining the evaluation object

Policy themes and sectors are often broadly formulated and cover a wide variety of activities. This applies in particular at the level of the operational objectives specified in the ministry's budget document. The preliminary study should therefore establish the limits of the evaluation object – also in the case of wide-ranging evaluations – such that it has a clear focus and thus remains feasible. This offers a better guarantee that the focus of the evaluation remains clear and that only those questions are formulated that

can reasonably be expected to be given a valid answer. Establishing the limits of an evaluation relates to the time dimension, the geographical dimension and the nature of the activities. It is based on conversations with relevant informants (stakeholders, external experts) and on literature and file studies.

Identifying result indicators

An important component of reconstructing policy theory is formulating a results chain, distinguishing various levels of results. Measurable result indicators are identified for each of the result levels, to enable performance to be assessed at that level.

Identifying data sources

Identifying the scale and quality of existing sources of information (files, literature, informants, statistical data bases) and the opportunities for gathering supplementary information determine the extent to which the subject can be evaluated and how the evaluation questions are formulated.

Formulating the research questions

30

The research questions included in the Terms of Reference are determined by the ultimate aim of the evaluation. They relate to the self-evident need for information on results achieved at the level of outputs and effects. This information is of course valuable for both the accountability and learning function of an evaluation. In addition, supplementary questions relating to a specific learning need can be formulated.

Determining the method of evaluation

The choice of research methods depends on a large number of factors. It will be influenced by the nature of the subject, the time and budget available for the evaluation, data that is already available and needs to be collected, and the nature of the research questions.

6.3 Terms of Reference

The Terms of Reference establish the conditions under which the evaluation will be conducted and form the result of the preliminary study. The ToR must contain at least the following components:

- the reason and justification for the evaluation: a short explanation of the relevance of the proposed evaluation;
- the (external) aim of the evaluation;
- the core questions: formulation of the central research question and the secondary questions derived from it;
- the demarcation and scope of the evaluation: explanation of how the field of

- research will be demarcated in terms of theme, geographical location and time;
- approach and choice of methodology, including focus on subsidiary studies: description of the design of the evaluation, the indicators, information sources and research techniques;
- representativeness: explanation of how and the extent to which representativeness plays a role in the evaluation;
- organisation: explanation of the organisation of the evaluation; deployment of experts, organisation of workshops, missions, etc.;
- products: indication of the expected subsidiary products, including subsidiary studies or a data base. Also an indication of whether the report will be published or whether the evaluation is internal;
- time path: the time path for the evaluation;
- budget: a specified outline of the costs.

If the Terms of Reference are for a policy review the questions prescribed by the RPE 2006 must also be taken into account. Policy reviews must address the following questions:

- What problem gave rise to the policy? Does the problem still exist? What is the cause of the problem?
- Why does the government consider tackling the problem as its responsibility? Why is the responsibility at central government level (and not at local or EU level)? What form does that responsibility take and why?
- What objective has the government formulated to solve the problem?
- What instruments are being deployed? How coherent are they? Is there overlap? What is known about the implementation of the policy and the efficiency of operational management?
- What is the effect of the instruments on the formulated objectives (solution to the problem)? Do instruments in other policy areas also have significant effect on the formulated objectives? What are the main positive and negative side-effects?
- How are the budgets determined? On what basis?

31

The draft Terms of Reference are not only put on the agenda of the IOB departmental meeting but should also be discussed in advance with direct stakeholders and external experts (preferably in the form of a reference group, see 7.2), and with the internal peer reviewers. The degree to which and the way in which comments are dealt with must be clear to all parties. The Terms of Reference for an evaluation are approved by the IOB Director; if the subject of the evaluation lies outside development cooperation, the ToR will be approved by the Director-General concerned.

The Terms of Reference form the basis on which a budget will be reserved for the evaluation. It also provides a framework for contracting external researchers. By approving the Terms of Reference, the IOB Director also approves the budget included in the ToR.

6.4 Deploying researchers

The IOB inspector bears full responsibility for setting up, conducting and completing the evaluation. In most cases, the inspector will deploy researchers to conduct the evaluation. The moment at which they will be deployed will vary, as will the extent to which they operate autonomously.

Researchers can be research assistants or consultants from research institutions, universities or consultant agencies.

Research assistants are usually recent graduates who support the inspector in conducting the evaluation. The activities they perform vary from file studies, assistance with field missions or interviewing informants to drafting parts of the final report. The way in which they are deployed will depend on their experience and can change over time. Research assistants are full members of the evaluation team.

32 With the majority of evaluations IOB contracts external consultants.⁶ Two criteria are decisive: the expertise of external experts must be complementary to that of the inspector, and they must have ample research experience, preferably in the context of evaluations. Various modalities are possible for deploying experienced experts. One or more experts can be deployed as a direct counterpart for the inspector and a division of tasks agreed between them. Another alternative is for whole or part of the evaluation to be contracted out to a researcher or a research bureau. The way in which external expertise is deployed depends on the nature of the services required and associated costs. In the case of joint evaluations with other donors, the whole evaluation is generally contracted out.

6.5 Cooperation with researchers in partner countries

In conducting evaluations IOB aims, where possible, to work together with researchers from partner countries. That helps to build independent research capacity in these countries. Furthermore, researchers from partner countries can be invaluable as they are often well informed on local developments and relations. However, not all evaluations are of course suited to such cooperation. The possibility of making use of researchers from partner countries depends in the first place on the subject of the evaluation. Some subjects do not warrant such a choice, for example, because the evaluation will address only Dutch activities or interventions of a global nature. Where the deployment of experts is a logical option (for example, evaluation of sector policy in a specific country) IOB actively explores the possibilities.

6 Contracting out and entering into contractual relationships are determined by the procedures laid down in the Operational Procedures Manual (HBBZ). They will not be considered further in this document.

6.6 Methodology

Evaluation is a form of research that entails a judgment. The customary methodology is largely derived from the social and economic sciences. The Terms of Reference give a general outline of the methodology to be applied in the evaluation. When the evaluation is to be contracted out to external researchers, they are usually asked to elaborate the methodology they will use in an inception report.

In accordance with the OECD/DAC Evaluation Quality Standards, IOB requires that a sound methodology is applied and that it is accounted for in the final report. The methodological choices made in an evaluation are determined by the extent to which they lead to the collection of reliable data and provide a basis for reaching valid and reliable judgments. The methods and techniques used can vary per evaluation.

The validity of an evaluation is determined by the extent to which the design of the evaluation, the method of data collection and the analysis of the data lead to replicable answers to the questions posed in the Terms of Reference. Validity relates to:

- the statement of the problem (the clarity with which the problem is stated, the formulation of the research questions and the definition of the evaluation criteria);
- the evaluation object (the definition, operation and limitation of the evaluation object and situating it in the policy and institutional context);
- the policy theory (the presentation or reconstruction of the intervention logic and result levels, and the operationalisation of result measurement on the basis of indicators);
- the analysis (data sources, data collection and the processing and argumentation of conclusions with findings).

33

The reliability of an evaluation is determined by the extent to which the results are stable, i.e. that repeating the evaluation will produce the same results. Reliability is related to:

- the research methods (specification of and accounting for the research methods used, verification of data and triangulation);
- the scope (representativeness of the sample and selection of case studies and specification of the limits of the evaluation);
- independence (of the source material and of the researchers in relation to stakeholders);
- the progress of the evaluation and quality control (reporting on the progress of the evaluation and organised quality control through internal and external monitoring).

To guarantee the validity and reliability of evaluations as much as possible IOB chooses a combination of research methods ('mixed methods'). This combination relates to the design of the evaluation and to qualitative and quantitative methods and techniques of data collection and analysis.

The design of the evaluation

The majority of IOB evaluations are designed according to the guidelines for 'theory-based' evaluation. That is logical because the aim is to assess the efficiency and effectiveness of policy implementation. That requires taking policy theory as a starting point. The model of theory-based evaluation is very suited to policy reviews. Theory-based evaluations focus on a detailed description of what has happened in the past and why, and analysing the demonstrable causes of these developments. To do that in a responsible manner, it is important in the first instance to reconstruct the policy theory. The evaluation then assesses the validity of the policy theory. The most tried and tested method of identifying aims and intervention strategies is through the logical framework or the often more detailed result chain. This model offers a framework for systematic data collection, identifying indicators for systematic assessment, and determining the relationships between the different result levels.

An alternative for this evaluation design is to take changes in reality as a starting point rather than the policy theory. Supporters of 'realist evaluation' advocate this kind of approach, which first devotes attention to the context in which changes occur and the mechanisms involved. Only in the second instance does it attempt to attribute changes to specific activities.

34

In recent years there has been increasing emphasis internationally on the need to provide stronger evidence of the effectiveness of government policy. Common evaluation practice, which generally provides plausible explanations of the relationship between interventions and their effects, is no longer considered adequate. There is a need to identify the net effects of interventions, if possible in quantifiable terms. That makes it necessary to conduct impact evaluations. The desire for harder impact studies is also expressed in the RPE 2006, which explicitly distinguishes ex-post impact studies as a separate category in addition to policy reviews. In its evaluation programming IOB therefore includes the implementation of a series of impact evaluations in appropriate policy sectors and themes. In practice, however, the options are not unlimited. 'Randomised control trials' (RCTs), in which a 'treatment group' is compared with a 'control group' and which in effect reproduce a laboratory situation, are not applicable in the great majority of cases. The most common alternative for RCTs is a quasi-experimental research design. Although this closely resembles a fully experimental RCT, it recognises that the researcher does not have full control over the composition of the treatment group and the context in which the measurements are made. These limitations are accounted for in the evaluation report.

For the sake of completeness it should be made clear that IOB's impact evaluations are also designed on the basis of a theory of how the interventions to be evaluated work. The most important difference is that this theory is not exclusively based on a reconstruction of policy, but is largely fuelled by existing scientific insights.

Qualitative and quantitative methods of data collection

Various methods have been developed to measure the effectiveness of interventions. Generally speaking a distinction can be made between qualitative and quantitative research methods. Qualitative research methods focus more on processes and emphasise the behaviour of actors. They are flexible, often because they have an eclectic approach, and can place findings in their cultural and political context. A qualitative approach can give a plausible explanation of why a certain intervention has or has not had an effect in a given context. Qualitative methods play an important role in determining internal validity. It is however not always possible to generalise the findings of qualitative research. Quantitative research, on the other hand, explicitly aims to make general statements (external validity). Statistical techniques make it possible to isolate the effects of an intervention in a complex environment. This method is, however, very data-sensitive. It requires a large quantity of reliable information to allow the formation of sample populations.

Quantitative and qualitative techniques are not interchangeable, but complementary. Quantitative techniques can provide a statistical basis for a more qualitative approach, while a qualitative analyse can explain relationships established by quantitative research. Without a theoretical framework and qualitative research into the mechanisms behind assumed relationships, statistical analysis has no substance. In its evaluations IOB applies both quantitative and qualitative techniques.

35

Methods of data collection

The validity and reliability of an evaluation stands or falls with the reliability of the data collection. Where the data is collected depends strongly on the subject of the evaluation. When development interventions are being evaluated some of the data will usually be collected in the partner countries concerned. If they are available, existing data bases can be used.

The most common forms of data collection are:

- **File and literature studies**

To obtain a clear picture of the field of study it is necessary to become familiar with the complete range of written sources available. This includes information relating to interventions, more policy-related documentation and academic literature. It requires working through the archives in great detail, preferably using a systematic list of points for attention. One important point for attention in a file study is to assess the reliability of the information.

- **Interviews with key informants**

Generally a distinction is made between structured, semi-structured and non-structured interview methods. Which method is applied depends on the role of the interview in the research and the nature of the information that is being sought. Especially in the initial orientation phase interviews will generally be open to allow as many different opinions and viewpoints to be collected from the interviewees.

- **Group interviews**

Various kinds of group interviews are used in evaluations. One of the most common forms is the focus group discussion, in which the interaction between the interviewees can also be a valuable source of information.

- **Direct measurement**

A lot of quantifiable information can be acquired by simple recording and counting. Especially at the level of outputs, this is the most logical approach. Direct measurement is particularly useful if data is available from a baseline study.

- **Direct observation**

Observation methods can be limited to the physical outputs of an activity, but can also relate to certain processes. Observation techniques (with or without the deployment of audiovisual aids) have always played an important role in field research. It should be added that they can be very time-intensive. Evaluations generally make use of a quick (more or less structured) tour of physical outputs, a 'transect', and/or attending one or more meetings of the target groups concerned.

36

- **Surveys**

Surveys aim to collect a large quantity of standardised data from a large population. Such data is very suitable for statistical analysis. Surveys make use of previously drawn up questionnaires which are preferably tested first and which can be in written or verbal form. They are essential in determining the starting situation (baseline) and are indispensable in impact evaluations. Because surveys can be time-consuming and costly, impact evaluations often use the results of previous questionnaires or modules added to pending surveys, so that they can serve multiple purposes.

- **Case studies**

Case studies are in-depth studies of a small number of specific interventions. They offer the possibility of considering the perceptions of stakeholders or reconstructing histories. Case studies are a valuable supplement to broader surveys, because they provide an opportunity to examine and explain relationships that become apparent from statistical analysis in greater detail. If an evaluation is based exclusively on a series of case studies, it is essential to determine to what extent the selected cases are representative.

6.7 Reporting

The final report of an evaluation is approved by the IOB Director, after she/he is satisfied that the comments of the reference group, internal peer reviewers and the IOB Departmental Meeting have been incorporated adequately. IOB then sends the report to the minister or ministers concerned to be forwarded to parliament. On request, IOB can also provide the minister with additional information on the findings of the evaluation. If an evaluation is addressing a politically very sensitive issue, IOB will inform the minister concerned in good time on the progress of the evaluation and, as far as possible, the expected outcomes.

In principle, all IOB evaluation reports are accessible to the public. This rule can only be deviated from in exceptional cases. An evaluation might, for example, be very confidential in nature or may only serve an internal learning function. The decision not to publicise an evaluation report must always be taken before the evaluation starts.

The report should take account of the fact that the findings of the evaluation are to be presented to different stakeholders with various interests. The basic principle is that IOB reports are designed to account to both parliament and society, and to fulfil a learning function. That sets high standards for the reporting. Parliament benefits from a short and concise presentation of the main findings and conclusions. Accounting for the results and the deployment of the public funds required to achieve them has priority. Policy departments will be interested in the extent to which the report provides information that offers opportunities to adjust policy formulation or implementation. In all cases it is important to aim for a transparent presentation of findings, analysis and assessment.

37

IOB evaluations identify a number of burning issues that, in the light of the evaluation findings, require attention. IOB therefore avoids making concrete recommendations, putting itself in the position of the policymaker and thereby of a butcher inspecting the quality of his own meat. Furthermore, IOB works according to the OECD/DAC definition of evaluation, which refers only to 'lessons learnt', and the guidelines for policy review in the RPE 2006.

In the interest of transparency the Terms of Reference should always be attached to the evaluation report as an annex. The research questions should be systematically addressed in the report, also if some of them cannot be answered. If necessary, additional information should be provided in the annexes on the research methods employed and any problems that arise during implementation of the evaluation.

If there are fundamental and substantiated differences of opinion between IOB and the stakeholders regarding the main conclusions of the report, these should be explicitly mentioned in the report. The decision to this will be taken by the IOB Director.

Many evaluation reports are a synthesis of subsidiary studies. Especially where these subsidiary studies can also be read independently it is important to make them accessible for policy departments or other stakeholders. They often contain detailed information that is not included in the main report. The evaluation reports are published both in print form and on the Ministry of Foreign Affairs website. Subsidiary studies are always available on the website and are in some cases also published in print form. In the case of very extensive evaluations, in the interests of accessibility, IOB produces a summarised version of the report. Both the summary and the full text are submitted for inclusion in OECD/DAC network for evaluation services' evaluation inventory. Completed impact evaluations are also entered in the data base set up by 3IE and NONIE for evaluations of this kind.

7

Quality control

The IOB inspector is responsible for conducting independent evaluations. The quality of the work carried out is, however, a concern for IOB as a whole. The IOB management has its own responsibility in this context, as it approves the evaluation reports. To prepare for this decision-making process, IOB has a quality control system comprising five instruments: internal peer reviewers, reference groups, the departmental meeting, the advisory panel and the self-evaluation.

7.1 Internal peer review

Quality control benefits from intensive cooperation and a systematic intercollegial monitoring. IOB therefore has a system that ensures that inspectors can rely on substantive support from colleagues when conducting evaluations. Start-up reports, Terms of Reference, subsidiary studies and the draft final report are discussed with two internal peer reviewers, fellow inspectors, before being submitted to the departmental meeting. The inspector chooses, in consultation with the IOB Director, which colleagues are to act as internal peer reviewers. The Director appoints internal peer reviewers per evaluation. They are selected and appointed on the basis of the following criteria:

- coaching of the inspector;
- internal readership as a learning function for recently appointed inspectors;
- evaluation experience;
- affinity with the evaluation subject and expertise in the theme concerned;
- avoiding conflicts of interest;
- an equal distribution of the workload within IOB.

Not all criteria are equally important and a specific criterion can be decisive.

The task of the internal peer reviewer is to give advice to the inspectors. This advice is not binding but inspectors do have to indicate what they do with it. Lasting and fundamental differences of opinion must be solved in consultation with the IOB Director.

Inspectors must include the comments and advice of internal peer reviewers and a short report of their discussions with the readers in the file on the evaluation.

The internal peer reviewers are not part of the reference group for the evaluation.

As the IOB Director is responsible for approving evaluation products he/she cannot be an internal peer reviewer.

The decision whether to make use of internal peer review in the event of joint evaluations between IOB and the evaluation services of other donors is left to the IOB inspector responsible for the evaluation, in consultation with the IOB Director.

7.2 Reference groups

A second form of quality control over IOB evaluations is reference groups. These groups, which are created for each evaluation, are partly external, consisting of representatives of the policy departments involved in the evaluation, organisations and/or local bodies, and external experts. A reference group is chaired by the IOB management and meets at key moments during the process (e.g. to discuss the draft ToR, partial reports and the draft final report). The inspector is responsible for reporting on reference group discussions and/or providing summaries of written comments by reference group members.

For evaluations that entail intensive cooperation with researchers in a specific partner country, a sounding board group is set up in the country concerned. The form the group takes will vary from case to case.

42 For joint evaluations with other donors a Dutch reference group is usually not set up. There is a joint steering group comprising representatives of the participating donors. In the case of large-scale, complex evaluations a smaller management team is set up from within the steering group to run the evaluation on a daily basis. The steering group is then primarily concerned with quality control. In addition, with these evaluations, the research bureau conducting the evaluation is expected to arrange its own quality control. This is one of the aspects taken into account when assessing tenders.

7.3 Departmental meeting

The IOB departmental meeting plays a part in quality control at two moments during the implementation of an evaluation: when discussing the Terms of Reference (or sometimes the inception report) and during discussion of the draft final report. The ToR are always discussed at a full departmental meeting. Before this meeting it must be clear to the inspector which colleagues have submitted comments on the Terms of Reference and the substance of the comments, in general terms, so that he/she can prepare for the discussion. Draft final reports are then discussed once again in a separate final meeting between the inspector and the internal peer reviewers, chaired by the IOB Director, before the Director approves the report. This final meeting, which is open to all IOB staff members, has two purposes: it offers the opportunity for a final assessment of the quality of the product, and helps colleagues to learn from each others' experiences. It also ensures that the final product of an individual inspector becomes a product of IOB as a whole.

7.4 Advisory panel

To increase the usability and use of evaluations in the field of development cooperation the Minister for Development Cooperation made a pledge to parliament to establish an Advisory Panel of prominent external experts. The panel, with its secretariat at IOB, will play a role in quality control by monitoring the usability of evaluations and offering advice on request or on its own initiative. It will become operational in the course of 2009.

7.5 Self-evaluation

Every evaluation concludes with a critical review of how it proceeded, with the aim of improving the approach to future evaluations. These self-evaluations are written by the inspector, sometimes together with a research assistant and the internal peer reviewers. The Terms of Reference for the evaluation determine the format of the self-evaluation. By systematically going through the various components of the ToR, the self-evaluation provides a clear insight into the successful and less successful elements of the evaluation. It should in any case address: (1) the extent to which the research assignment could be implemented, (2) any differences between the planned time path and the actual duration of the evaluation, and (3) differences between estimated and actual costs. Self-evaluations are discussed internally and periodically at the departmental meeting.

8

Communication and feedback

Communication and feedback relate to the feedback of evaluation results, with a view to the transparency of the evaluation process, the accountability function and the integration of lessons learned in future policy formulation and implementation. Feedback is not limited to the phase following completion of the evaluation. Consultation with and feedback to stakeholders take place throughout the whole period of implementation: during the preliminary study, in relation to the Terms of Reference, and during discussion of subsidiary studies and the draft final report. Where necessary, consultation and feedback is also provided to the partner countries. While feedback during the evaluation is important at implementation level, after the evaluation has been completed it is more important at political and more general level.

The extent to which completed evaluations fulfil the learning function depends on a number of factors. In the first place, it is a matter of the usability of the evaluation itself. This is determined by:

- the clarity with which the external aim (i.e. outside the evaluation itself), for which the evaluation results can be used, is specified;
- the clarity and completeness with which the essence of the evaluation and especially its main findings are presented in the evaluation report and the summary;
- the completeness with which all research questions are answered by the conclusions;
- the extent to which the results of the evaluation suggest actions that lie within the capacity of the recipient of the report.

45

In the second place learning from evaluations depends on the extent to which the recipients and users are receptive to the lessons drawn from the studies. This is related to the institutional context and the organisational culture. IOB has no influence on this and it is not IOB's responsibility. It is the direct responsibility of the management of the organisation (the ministry or other organisations that receive IOB reports).

IOB does however have a responsibility to communicate on evaluation reports. As communication with policy departments is crucial to increase the learning effect, it is crucial that evaluations are discussed in the relevant staff meetings. With regard to external communication, IOB draws up a communication plan for each evaluation, in consultation with the ministry's Information and Communication Department. IOB increases the accessibility of evaluation results and ensures their wider distribution by preparing media presentations, giving lectures and interviews, writing articles, and organising conferences and seminars.

IOB evaluations are also an instrument for rendering accountability to parliament. Within three months of the report being approved by the IOB Director, the minister sends it to parliament, usually together with a policy reaction. The text of the reaction, which is usually drafted by the policy department concerned, contains indications on

how the results of the evaluation are to be acted upon in the light of future policy. In this way, accounting to parliament also contributes to the learning process.

In the case of evaluations in the field of development cooperation the reports are also published in the partner countries, offered to the government and made widely known through the organisation of meetings. With a view to mutual accountability IOB invests systematically in such feedback to the partner countries.

Evaluation studies published by the Policy and Operations Evaluation Department (IOB) 2004-2009

Evaluation studies published before 2004 can be found on the IOB website at www.minbuza.nl/iob

- 297 2004 Over solidariteit en professionalisering**
Evaluatie van Gemeentelijke Internationale Samenwerking (1997-2001).
isbn 90-5328-341-2
- 297 2004 On solidarity and professionalisation**
Evaluation of Municipal International Co-operation (1997-2001).
isbn 90-5328-342-0
- 298 2004 Onderzoek naar de kwaliteit van in 2002 afgeronde decentrale evaluaties**
Eindrapport.
isbn 90-5328-344-7
- 299 2005 Een uitgebreid Europebeleid**
Evaluatie van het Nederlands beleid inzake de toetreding van Midden-Europese landen tot de Europese Unie 1997-2003
isbn 90-5328-347-1
- 300 2005 Aid for Trade?**
An Evaluation of Trade-Related Technical Assistance
isbn 90-5328-349-8
- 301 2006 Van Projecthulp naar Sectorsteun**
Evaluatie van de sectorale benadering 1998-2005
isbn 90-5328-351-x
- 301 2006 From Project Aid towards Sector Support**
An evaluation of the sector-wide approach in Dutch bilateral aid 1998-2005.
isbn 90-5146-000-7
- 302 2006 Evaluatie van het Nederlandse mensenrechtenbeleid in de externe betrekkingen**
isbn 90-5328-352-8
- 303 2006 Dutch Humanitarian Assistance An Evaluation**
isbn 90-5328-352-8
- 304 2007 Evaluatie van de vernieuwing van het Nederlandse onderzoeksbeleid 1992-2005**
isbn 978-90-5328-353-0
- 304 2007 Evaluation of the Netherlands' Research Policy 1992-2005 (Summary)**
isbn 978-90-5328-353-0
- 305 2007 Impact Evaluation: Water Supply and Sanitation Programmes Shinyanga Region, Tanzania 1990-2006**
isbn 978-90-5328-354-7
- 306 2007 Chatting and Playing Chess with Policymakers**
Influencing policy via the Dutch Co-Financing Programme
isbn 978-90-5328-355-4
- 307 2008 Beleidsdoorlichting seksuele en reproductieve gezondheid en rechten en hiv/aids 2004-2006**
isbn 978-90-5328-358-5
- 308 2008 Het Nederlandse Afrikabeleid 1998-2006**
Evaluatie van de bilaterale samenwerking
isbn 978-90-5328-359-2
- 308 2008 Het Nederlandse Afrikabeleid 1998-2006**
Evaluatie van de bilaterale samenwerking (Samenvatting)
isbn 978-90-5328-359-2
- 308 2008 The Netherlands' Africa Policy 1998-2006**
Evaluation of the bilateral cooperation (Summary)
ISBN 978-90-5328-359-2

- 309 2008 Het Vakbondsmede-financieringsprogramma**
Een evaluatie van steun gericht op versterking van vakbonden en vakbonds- en arbeidsrechten
isbn 978-90-5328-357-8
- 309 2008 The Netherlands Trade Union Co-Financing Programme**
An evaluation of support for trade unions and trade union and labour rights (Summary)
isbn 978-90-5328-357-8
- 309 2008 El Programa de Cofinanciamiento Sindical**
Una evaluación del apoyo orientado al fortalecimiento de sindicatos y de derechos sindicales y laborales (Resumen)
isbn 978-90-5328-357-8
- 310 2008 Clean and sustainable?**
An evaluation of the contribution of the Clean Development Mechanism to sustainable development in host countries
isbn 978-90-5328-356-1
- 311 2008 Impact Evaluation: Primary Education in Uganda**
isbn 978-90-5328-361-5
- 312 2008 Impact Evaluation: Primary Education in Zambia**
isbn 978-90-5328-360-8
- 313 2008 Xplore-programma**
isbn 978-90-5328-362-2
- 314 2008 Primus inter pares**
Een evaluatie van het Nederlands EU-voorzitterschap 2004
isbn 978-90-5328-364-6
- 315 2008 Impact Evaluation: Support to Rural Water Supply and Sanitation in Dhamar and Hodeidah Governorates, Republic of Yemen**
isbn 978-90-5328-364-6
- 316 2008 Be our guests**
Beleidsdoorlichting Nederland als gastland van internationale organisaties en Samenvatting
isbn 978-90-5328-370-7
- 316 2008 Be our guests**
Policy review on hosting international organisations in the Netherlands and Summary
isbn 978-90-5328-371-4
- 316 2008 Be our guests**
Examen de la politique menée par les Pays-Bas en tant qu'État hôte des organisations internationales (Sommaire)
isbn 978-90-5328-372-1
- 317 2008 Sectorsteun in milieu en water**
isbn 978-90-5328-369-1
- 318 2008 Samenwerking met Clingendael**
Evaluatie van de subsidieovereenkomst tussen de ministeries van Buitenlandse Zaken en Defensie en Instituut Clingendael
isbn 978-90-5328-365-7
- 319 2008 Meer dan een dak**
Evaluatie van het Nederlands beleid voor stedelijke armoedebestrijding
isbn 978-90-5328-365-3
- 320 2008 Het tropisch regenwoud in het OS-beleid 1999-2005**
isbn 978-90-5328-374-8
- 321 2009 Maatgesneden Monotoring 'Het verhaal achter de cijfers'**
Beperkte beleidsdoorlichting Medefinancieringsstelsel 2007-2010
isbn 978-90-5328-374-5
- 322 2009 Draagvlakonderzoek**
Evalueerbaarheid en resultaten
isbn 978-90-5328-375-2
- 323 2009 Preparing the ground for a safer world**
Evaluation of the Dutch efforts to control landmines and explosive remnants of war 1996-2006
isbn 978-90-5328-377-6

Published by:

Ministry of Foreign Affairs
PO Box 20061 | 2500 EB The Hague | The Netherlands
www.minbuza.nl

Layout: Eindeloos, The Hague
Printer: OBT, The Hague

© Ministry of Foreign Affairs | October 2009 | Order code: BZDR666z/E

This document describes the evaluation policy and guidelines for evaluations of the Policy and Operations Evaluation Department (IOB) of the Dutch Ministry of Foreign Affairs. In publishing these guidelines, IOB wishes to provide transparency to third parties: colleagues at the Ministry of Foreign Affairs, the Dutch parliament, organisations and institutions affected by IOB evaluations, users of the evaluation results, and interested citizens. IOB will periodically review its policy and the guidelines and, if necessary, revise them.

Policy and Operations Evaluation Department | IOB | October 2009 | Policy and Operations Evaluation Department | IOB | October 2009 | Policy and Operations Evaluation Department

Published by:

Ministry of Foreign Affairs

PO Box 20061 | 2500 EB The Hague | The Netherlands

www.minbuza.nl

© Ministry of Foreign Affairs | October 2009 | Order code: BZDR6662/E