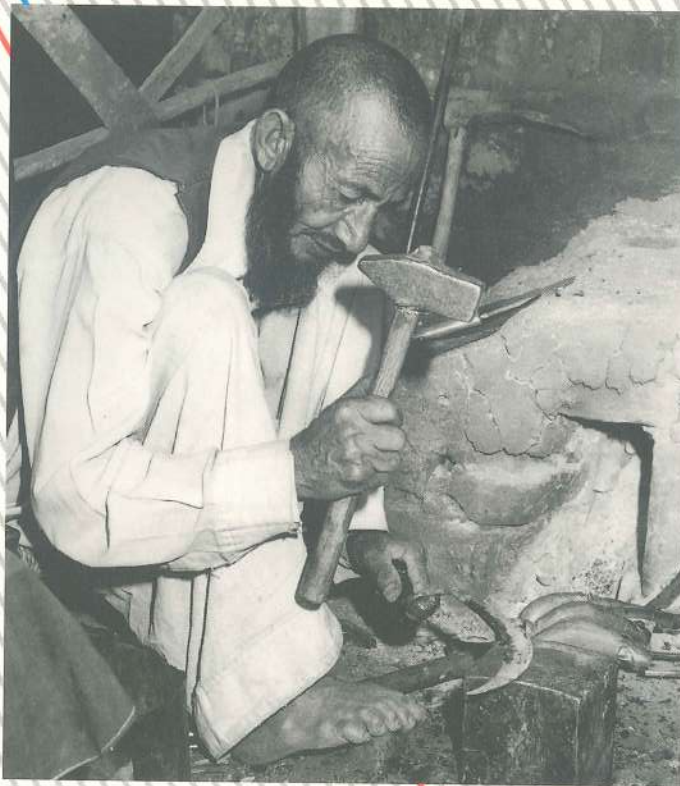# EVALUATION AND MONITORING

NETHERLANDS DEVELOPMENT COOPERATION

EVALUATION AND MONITORING

The role of project evaluation and monitoring
in Netherlands bilateral aid

Preface

It is crucially important to acquire insight in how development activities may be implemented successfully. Through monitoring and evaluation of aid implementation and its results information is gathered on which responsible and effective decision-making regarding policy implementation and formulation may be based.

This summary report is an assessment of project evaluation and project monitoring as practiced in Netherlands bilateral development cooperation. The report deals with evaluations carried out by the operational units of the Directorate General for International Cooperation, and is based on desk studies and field studies. The desk studies focused on the quality of evaluation reports; the field studies aimed at a detailed picture of the preparation and implementation of evaluations and the use of evaluation results. The role of monitoring in project management was studied as well, as were the perceptions and views of those involved in monitoring and evaluation.

In September 1993 a more extensive version of this report, written in Dutch, was presented to Parliament. The Minister for Development Cooperation's comment on the report is included in the present publication.

The evaluation study was directed and coordinated by Ted Kliest (Operations Review Unit), Willem Koot (Consultants for Culture and Management) and Bert van de Putte (Management for Development Foundation). Although many others contributed to this study, the Operations Review Unit bears sole responsibility for the contents of this report.

Director Operations Review Unit

Contents

As from 1 January 2002, all amounts in guilders must be read as
amounts in euros, based on the official exchange rate

EUR 1 = NLG 2,20371.

This means that each amount should be divided by 2,20371.

Abbreviations used

CIDA	–	Canadian International Development Agency
DAC	–	Development Assistance Committee (OECD)
DAP	–	Drainage Advisory Panel (Egypt)
DGIS	–	Directorate General for International Cooperation
DST	–	Directie Sectorbeleid, Speerpuntprogramma's en Technische Advisering (Spearhead Programmes Coordination and Technical Advice Department)
DST/TA	–	DST/Technical Advice Section
EU	–	European Union
FAO	–	Food and Agriculture Organisation
Finnida	–	Finnish Development Agency
IFAD	–	International Fund for Agricultural Development
IOV	–	Inspectie Ontwikkelingssamenwerking te Velde (Operations Review Unit)
MIDAS	–	Management Informatie en Documentair Administratief Systeem
NGO	–	Non-Governmental Organisation
NORAD	–	Norwegian Agency for Development Cooperation
ODA	–	Overseas Development Administration
OECD	–	Organisation for Economic Cooperation and Development
OED	–	Operations Evaluation Department (World Bank)
OOPP	–	Objective Oriented Project Planning
PC 1–5	–	Planning Commission Forms (Pakistan)
PFP	–	Productive Families Project
UNDP	–	United Nations Development Programme
USAID	–	United States Agency for International Development
WID	–	Women In Development

Summary

1   Background

Evaluations, in the Netherlands' development cooperation programme, take two forms. Some focus on specific themes, sectors, countries and programmes and are carried out by the Operations Review Unit (IOV). Others are limited to specific projects and are the responsibility of the operational units of the Directorate General for International Cooperation (DGIS).

Project evaluation is the occasional assessment of a current or completed project and the Evaluation Guidelines list three main purposes to it: management support, policy support and communication (between those concerned with implementation). Wherever possible the DGIS and the relevant institutions in each developing country shall be jointly responsible for the design and implementation of these evaluations. Independent expatriate and local experts should be employed to conduct them.

Monitoring is a general management tool which supplies those with the responsibility of managing the implementation of development policy with the up-to-date information necessary for sound decision-making. There are two kinds of project monitoring. Internal monitoring is part of project management and should provide project managers with the up-to-date information necessary for sound decision-making. External monitoring gives the donor and its counterpart institutions information about progress and, in the context of Netherlands' development policy, is made the responsibility of the embassy in the recipient country. Embassies may, when necessary, enlist the support of monitoring missions made up of external experts.

Adequate evaluation and monitoring are essential tools for supporting for the aid programme. Yet the question of how good and effective project evaluations are has been debated for some time. This study was prompted by the lack of real information in that debate.

This study is concerned solely with evaluations for which the operational units of DGIS are responsible. Evaluation is assessed in the light of the formal functions assigned to it in the Evaluation Guidelines. Monitoring is considered only in so far as it relates to evaluation. The investigation involved a combination of interviews, desk studies and field work. One hundred and eighty evaluation reports submitted between 1987 and 1990 were examined for the extent of their coverage, for their clarity and for their completeness. Thirty seven reports were chosen at random and the quality of their content was assessed. The perceptions and opinions of those involved in evaluation are analyzed. Field work was undertaken in Burkina Faso, Pakistan and Egypt (16 projects in total).

2   Main findings

A   *Project evaluation does not adequately fulfil all its allotted functions*

Project evaluation clearly supports project implementation, mainly because of its operational focus and its orientation towards the future. It also plays an important part in communications between the responsible counterpart institution, those engaged in project implementation, the embassy and DGIS. However, the results of project evaluations are not used structurally in formulating policies either for countries or for regions. Because use is not made of them for policy support, their contribution to the quality and effectiveness of aid is limited.

B   *Project evaluations do not satisfy methodological requirements*

Country desks, embassies and counterpart institutions seem unable efficiently to set up project evaluations. Evaluators are selected for their knowledge of the country or for their technical and sectoral expertise rather than for their experience in evaluation. Insufficient time is allowed for the job to be done well. Necessary project information is frequently lacking, usually because monitoring in projects has been inadequate. The effect of projects on their target groups is rarely subjected to proper analysis. Sustainability is often not assessed.

The absence of an explicit and coherent methodology, the lack of decent statistical material, of indicators and of standards for measuring performance means that assessments rest predominantly on the impressions and expert opinions of the evaluators. Consequently, the information in the reports can be both superficial and limited in scope.

Because time is so limited and because field visits give no opportunity for the structured collection of primary data, evaluators gather the information they need largely through interviews with project staff and government officials. This may result in an 'authority bias'.

C   *The involvement of counterpart institutions and evaluators from the developing countries is limited*

Because the DGIS usually initiates evaluations, counterpart institutions see them chiefly as a necessary step in DGIS's procedures for modifying or extending projects. This, combined with their limited financial, technical and personnel resources, means that counterparts have little interest in actively shaping evaluations.

Local evaluators, with few exceptions, play a subordinate role in relation to their expatriate colleagues. Since the design and preparation of the evaluation and the final reporting all normally take place in the Netherlands, local evaluators have no part in them. They do, however, play an important role as facilitators providing access to groups of respondents and as interpretors of local social and cultural phenomena.

D   *The role of monitoring is unclear*

Although, in some cases, significant steps have been taken to encourage the establishment of monitoring systems in projects, substantive DGIS guidance on the design and operation of such systems is lacking. As a result, project documents usually pay insufficient attention to the place of monitoring in management or to setting suitable indicators. Monitoring is also hampered by project objectives being defined too broadly and by the failure sufficiently to specify assumptions, activities and intended results.

External monitoring, by means of regular progress reports, contacts between project teams and embassies and, sometimes, monitoring missions consisting of external experts, is effective. It is, however, strongly donor-centered. Even when joint external monitoring arrangements exist, a parallel DGIS-oriented system of monitoring remains in operation.

The absence of proper guidelines for monitoring means that those responsible for the design of external monitoring arrangements have a large measure of discretion.

Whether external experts are deployed depends largely on the problems encountered in the project at issue, the embassy's capacity to supervise and the need of the embassy or country desk for an additional channel of information. In practice monitoring missions do the jobs neglected by others; this blurs the dividing line between project implementation and project supervision.

Conclusion

The main findings of this study point to deficiencies in evaluation and monitoring which limit the effectiveness of these two instruments. That these deficiencies are not unique to the DGIS is made clear by studies commissioned by other donors (see for example CIDA, 1991c, Scanteam, 1993, Finnida, 1991, USAID, 1992).

The problem springs from a variety of causes involving institutional and procedural matters. Institutionally, the frequent changes of staff and shifting project portfolios are among the factors which impede a balanced build-up of knowledge and experience. Procedurally, the lack of clarity between country desks and embassies about where operational responsibility lies creates difficulties; so too, does the ambiguous distribution of responsibility between the donor and recipient countries. To improve the quality of evaluation and monitoring it is first necessary to establish the right kind of institutional, procedural and substantive framework.

3   Recommendations

In future, attention must be paid to the functions that evaluation and monitoring are expected to fulfil. These functions should include a combination of learning, control and accountability. This study proposes some measures for improving the quality of these instruments and the way they are used within DGIS.

When it comes to the form which evaluation takes, greater diversity would enable the need for information, at both operational and policy levels, to be met more effectively. For example, a distinction could be made between reviews which are operationally-oriented and evaluations that are explicitly designed to consider issues of policy. Far more transparency and consistency is needed in the design and formulation of projects and more appropriate planning techniques would make for greater clarity. Evaluation and monitoring should be better integrated into the project cycle.

The study also suggests that baseline data should be generated during the formulation of projects and that outputs and effects should be monitored during their implementation. Specific impact studies could be planned before evaluations take place.

Responsibility for evaluation and monitoring should lie with the main users of the information they produce. The embassies should be responsible for monitoring and reviewing, while evaluations more to do with policy issues should remain the responsibility of the country desks.

The role of recipient countries in setting up project evaluations needs to be strengthened. The participation of counterpart organisations and of local evaluators in carrying out evaluations should be increased. Any effort towards greater joint responsibility for evaluation must also include measures to strengthen the evaluation and monitoring capacities in both government departments and private institutions in recipient countries. Where a developing country has established its own monitoring and evaluation capacity, it should be used to its fullest extent.

Chapter 1 Introduction

1.1 Background to the study

Improving the quality and effectiveness of aid is a central theme of the Netherlands' policy of development cooperation. Learning from the past plays an important role and information is obtained by evaluation. Evaluations take two forms:

- those carried out by the Operations Review Unit (IOV) which focus on specific themes, sectors and programmes;
- those which are carried out under the responsibility of the operational units of the Directorate-General for International Cooperation (DGIS). These evaluations are limited to specific development activities (mainly projects) and are governed by specific guidelines (Evaluation Guidelines).

The quality and effectiveness of evaluations carried out by the operational units have been debated for some time. It was the lack of adequate information on how these evaluations worked that prompted this study.

Questions addressed

Four principal questions are addressed:

1) What forms of evaluation can be distinguished in practice, and to what extent do they conform to the Evaluation Guidelines?
2) To what extent and in what way are the purposes of evaluation set out in the Guidelines achieved? What factors determine this? Are any other functions assigned to evaluation?
3) How far does current practice meet the wishes, expectations and capacities of the various parties involved (country desks, embassies, evaluators both in the Netherlands and in developing countries, counterpart organisations, executive agencies)?

4) How far does current practice within DGIS reflect the recommendations of the OECD Development Assistance Committee (DAC) and the views on evaluation current in the international donor community?

1.2   Design of the study

This study is concerned solely with the evaluations carried out by the operational units of DGIS in the context of the bilateral aid programme. It covers projects contracted out to consultants and those directly managed by DGIS. Project monitoring is dealt with only when it relates to evaluation.

The study analyzes the perceptions and opinions of those involved in evaluation, including DGIS staff in The Hague and at embassies in developing countries, evaluators, project staff and the personnel of counterpart organisations. Evaluation and monitoring are studied with the Evaluation Guidelines in mind.

The investigation involved a combination of desk studies and field work. First, 180 evaluation reports from the period 1987–90 were appraised using criteria such as coverage, clarity and completeness. This was followed by an assessment of the quality of the contents of a random sample of 37 reports.

Field studies were carried out in 16 development projects in Burkina Faso, Egypt and Pakistan. The aim of these studies was to obtain a detailed picture of the preparation and implementation of evaluations and the use of evaluation results. The role of monitoring in project management was also studied. The three countries were chosen for their varying capacities for local evaluation, for the differences in their supervisory structures for project implementation and for the number of evaluations carried out in recent years. The choice of projects was based on the quality of the evaluations, on the supervisory structure of project implementation, and on the existence of internal monitoring arrangements. Projects were chosen which would encompass the full range of evaluation studies and monitoring systems. The desk and field studies were supplemented with material obtained in interviews with experts covering general aspects of evaluation and monitoring.

The study sought neither to assess the projects themselves nor to duplicate their evaluation. It makes no comment on the accuracy of the information generated by project evaluation, but focuses instead on the coverage and depth of the information and on how it is used.

The study was supervised by a committee made up of external experts and DGIS staff. Responsibility for the content of this report rests with the IOV.

1.3 Terminology

The terms 'evaluation' and 'monitoring' are used in widely differing ways. For the purpose of this study they are defined thus:

- Evaluation is the occasional assessment of a current or completed project or programme with a view to determining the relevance of its objectives, its effectiveness and efficiency, its immediate and wider effects and its sustainability. Evaluation produces information from which lessons can be drawn for use in decision-making at project and policy levels.

The functions of evaluation, listed in the Evaluation Guidelines, are management support, policy support and communication. For the purposes of this study these functions are interpreted thus:

- Management support is geared to the needs of the parties responsible for guiding project implementation. These are the project agency or contractor, the embassy and (at some distance, under the system of delegation) the relevant operational unit within DGIS.
- Policy support entails the supply of information on development activities for the purpose of adjusting existing policies and formulating new policies. Information from bilateral project evaluations is of particular relevance to the policy plans of the country or region, as well as to aid policy in general. Policy support is thus geared, in the first instance, to the relevant operational units and regional directorates.
- The function of communication is to ensure that evaluations lead to the exchange of relevant information between DGIS units, embassies, projects and counterpart organisations.

Monitoring is a general management function which supplies those with management responsibilities with the up-to-date information necessary for sound decision-making. The specific form taken by monitoring is determined mainly by the structure and responsibilities of the level of management for which the information flow is intended. Two levels of monitoring are distinguished:

1) Monitoring aimed at project management is internal to the individual project and should keep project managers informed of progress in implementation. This internal monitoring is primarily, but not exclusively, concerned with aspects of the implementation of the project itself.

2) External monitoring carried out by the donor (and the counterpart organisation) in the context of project supervision should provide information about the progress of the project concerned. Its aim is the early identification and analysis of factors which may affect the planned progress of the project and the achievement of its goals. This form of monitoring is based on progress reports and contacts with projects and is concerned with the progress of the project in a general sense and with the institutional framework within which it is being implemented. Where necessary, embassies can enlist the help of external experts (monitoring missions).

Chapter 2 Monitoring and Evaluation in Development Cooperation

There is growing interest in the evaluation of aid projects and programmes. Donor organisations are sponsoring complex activities throughout the world in hugely differing social, political, cultural and economic environments. This, together with a shift towards a more pragmatic approach to development assistance, as against the largely ideologically inspired approaches of the past, has led to a sharper focus on the effectiveness and efficiency of aid activities.

In the past two decades the number of interventions of a long-term, multisectoral and multidisciplinary nature, mainly aimed at strengthening the institutional framework of developing countries, has grown. The implementation of these interventions is increasingly flexible, which makes any necessary adjustments simpler to introduce. This increases the need for thorough analyses of the factors determining success or failure.

Where other agencies have direct access to, or even come under pressure from, those whom they seek to benefit, this is not the case in development cooperation. Target groups are far away and poorly organised and do not generally see aid as a right. They are, therefore, unlikely to adopt a critical attitude. Communication between the many different projects in the field and the headquarters of the donor organisations is often inadequate.

Where informal communication is made difficult by, for example, the great distances involved, more formal procedures and instruments such as monitoring and evaluation take on greater importance. They constitute an important link between the determination of policy and its implementation.

2.1   Theoretical aspects of monitoring and evaluation

Monitoring

In order to derive full benefit from the greater flexibility in the design and execution of interventions, which the programmatic approach to planning offers, a continuous flow of information is essential. It can be obtained by adequate monitoring.

Monitoring was seen initially as a form of evaluation and no clear distinction was made between information generated by it and that emerging from evaluation studies (Rossi & Freeman, 1985). Gradually, however, it became clear that the differences were considerable. In the literature monitoring is seen as separate from, but closely related to, evaluation (Casley & Kumar, 1987). In the context of development cooperation, despite attempts at standardisation, it is defined in various ways (see, for example, IFAD, 1984; FAO, 1984; Casley & Kumar, 1987). Van de Putte (1991) notes that as yet the concept has no generally accepted or consistent theoretical underpinning.

Casley and Kumar (1987) see monitoring as specifically to do with implementing projects and programmes. In their view monitoring is in the first place an activity internal to a project or programme. It is essential to effective management and hence a vital tool for project managers.

Monitoring development activities is not, however, limited to the level of project implementation: 'as an integral component of the management function, and hence as an essential part of good management practice, monitoring needs to be conducted by those responsible for project/programme implementation at every level of the management hierarchy' (IFAD, 1984, p. 20). Within donor organisations, management at various levels monitors to gather information on activities in progress. Data generated at project level flow, by means of reporting systems, from 'lower' to 'higher' levels in the management hierarchy, becoming increasingly aggregated as they do so.

Monitoring often focuses on the use of project inputs and activities, which are usually clearly specified and easily measured. On the other hand, the monitoring of outputs, effects and impacts of project activities is less well developed (Casley & Kumar, 1987). The bias towards input measurement has meant that in practice monitoring has developed into a strongly control-oriented instrument of management. Use is often made of forms on which the inputs used and activities carried out must be shown. The comprehensive and standardised monitoring systems of such donors

as USAID and UNDP are of this control-oriented nature (UNDP, 1987; USAID, 1987a,b). While this is of great importance, information systems which only meet administrative reporting and accounting requirements represent a limited interpretation of monitoring. Broader forms involving the measurement of effects on the target group are increasingly seen as important (see, for example, Casley & Kumar, 1987; Frerks, Thomas & Tomesen, 1990; Van de Putte, 1991), but they are relatively little used. The design of such systems within projects seems to be difficult.

Evaluation

Evaluation is seen as a major research instrument in the service of government for determining the effectiveness of policy and its implementation.

For a long time the literature on evaluation was dominated by debate which focused on the methods to be used in evaluation studies and on their object, function and scope. Since the start of the 1980s the emphasis has gradually shifted towards more practical issues. Because evaluation is insufficiently recognised as an instrument of management, fitting evaluation into the policy process produces organisational and management problems. Concern at the use made by policy-makers, and those implementing policy, of the results of evaluations is growing. Questions are being raised about the underlying rationale of evaluation (see, for example, Hoole, 1978; Weiss, 1988; Patton, 1990).

In a critique of the evaluation methods used in development cooperation, Marsden and Oakley (1991) suggest that it is never a 'neutral' research instrument. They argue that its use can never adequately chart the complex reality: evaluators need to penetrate that reality by explicitly examining the position and role of the various actors involved, directly or indirectly, in the project. This means determining the relative strengths of project managers, counterpart organisations, target groups (and factions within them) and donor agencies. To arrive at an understanding of the reasons for the success or failure of an intervention, an appreciation of the complex social, economic, political and institutional realities within and around projects is needed.

2.2   Evaluation and development cooperation

The 1970s and 1980s brought growing scepticism about the effectiveness of development aid. Donor organisations responded by promoting evaluation (OECD, 1988;

Stokke, 1991). It became institutionalised as part of the policy and planning cycle of development activities and evaluation units were set up. Several factors are at work in the support for evaluation, and first among them is scepticism. Evaluation is seen as offering an insight, and its results can be used by donor organisations and by politicians both to justify aid allocations and in answer to the question whether, and to what extent, aid actually works. Second, the actual aid activities take place in developing countries and are as a rule undertaken by institutions which differ from donor organisations in their administrative structures and decision-making processes. The need for control is considerable, and evaluation is seen as instrumental in this respect (Berlage & Stokke, 1992).

A third rationale for evaluation, after political legitimation and control, is that it allows lessons from the practice of development cooperation to be drawn: analysis of and feedback from past experience is an essential part of the policy cycle. The Expert Group on Aid Evaluation, established in 1982 by the OECD/DAC, emphasised this function by drawing attention to the importance of integrating evaluation into aid planning (OECD, 1988).

Compared with the importance that donors attach to evaluation, enthusiasm in the recipient countries is generally less marked. Berlage and Stokke (1992) see several factors at work here. The evaluation of development projects and programmes may lead to criticism of government bodies and result in reduced aid. Evaluation is initiated mainly by donors and carried out with help from foreign experts, with the result that developing countries and those implementing projects see it as a one-sided instrument of control. The diversity of evaluation approaches is also confusing for recipient countries. The net result of these factors is a measure of dissatisfaction with current evaluation practices.

Despite these circumstances and the lack of funds and expertise, recipient countries increasingly see the importance of evaluation. Thus individual donors, particularly through the DAC, are working to strengthen their evaluation capacity (see, for example, OECD, 1988 and 1992). Moves are also being made through the DAC towards some standardisation of the approaches and methods used by donors. The DAC Expert Group on Aid Evaluation promotes, among other things, inter-donor coordination in the area of evaluation, and it is an accepted principle that aid activities should be evaluated jointly by the donor and the recipient.

The evaluation of development aid has undergone a number of changes over the years. In the 1960s and 1970s project-oriented evaluation was dominant. Most studies were relatively simple in design, particularly those considering unisectoral

activities developing physical infrastructure and improving agricultural and industrial productivity. Evaluations of such interventions naturally tended to concentrate on analyses of economic costs and benefits. The advent of multisectoral projects aimed at target groups, the increased role of institutional support in aid activities and the move towards a more process-oriented approach to implementation brought a change in the nature of evaluation. Determining the immediate and wider effects of such interventions calls for complex research methods and multidisciplinary evaluation teams (see, for example, Binnendijk, 1989). The 1980s brought more and more new forms of development aid, among them programme aid and macroeconomic support, whose evaluation also demands a special approach.

2.3   DAC principles for monitoring and evaluation

The *DAC Principles for Effective Aid* (OECD, 1992) stresses the importance of monitoring and evaluation: 'Technical cooperation requires systematic monitoring and evaluation to give managers and policy-makers full information for decision-making, effective implementation and public accountability' (OECD, 1992, p. 64). The DAC Expert Group on Aid Evaluation has sought to define the meaning of monitoring and evaluation more precisely.

Monitoring is defined as 'a management function which uses a methodical collection of data to determine whether the material and financial resources are sufficient, whether the people in charge have the necessary technical and personal qualifications, whether activities conform to work plans, and whether the work plan has been achieved and has produced the original objectives' (OECD, 1988, p. 36). Monitoring should not be confused with control, but should be seen as the continuous collection and analysis of information which will determine the progress of a project and serve to underpin any adjustments that may need to be made in its implementation.

The importance which the DAC attaches to monitoring is explained partly by the recognition of omissions in evaluations. Many of them are hampered by a lack of project-level information and monitoring could supply much of the information required, giving a clearer picture of both the intended and unintended effects of a project. For that reason alone, the contribution of adequate monitoring to evaluation is very important.

Evaluation is defined as 'an assessment, as systematic and objective as possible, of an ongoing or completed project or 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' (OECD, 1992, p. 132).

The DAC devotes considerable attention to the principles and objectives of evaluation. It gives two main objectives. First, evaluation has a function in the improvement of development policy in general, and of policy for specific interventions in particular. It helps 'to improve future aid policy, programmes and projects through feedback of lessons learned' (*ibid*, p. 132). Second, the data it generates on aid activities serve 'to provide a basis for accountability, including the provision of information to the public' (*ibid*, p. 132).

Impartiality and independence are major principles in evaluation. The ideal arrangement for it in the organisational structure of donor institutions is one in which responsibility for the evaluation and for the operational units involved in planning, managing and implementing development activities is separated. The monitoring of activities in progress is primarily the responsibility of operational units.

In the DAC's view the evaluation of aid activities is an integral part of the policy process and the scheduling of aid activities. Donors should take the information needs of various levels of management within the organisation into consideration and the feedback of findings to each of these levels is essential. It is desirable that evaluation results be transmitted in aggregate form to central managers and to policymakers.

2.4 Organisational aspects of monitoring and evaluation

Donors have developed monitoring and evaluation largely independently of one another. Organisational arrangements, therefore, differ.

Monitoring

Most donors have a decentralised structure consisting of headquarters and field offices. While the extent of overall decentralisation and delegation differs, monitoring is usually decentralised. In practice most donors divide responsibilities between headquarters staff and field personnel or embassy staff fairly flexibly. The following examples illustrate this diversity.

USAID and the European Union put the accent on monitoring in the countries where aid activities are located. Other donors, like Canada (CIDA), operate a 'mixed' monitoring structure in which a supervisory team, drawn from officials at headquarters and from the field office or embassy, is set up for each project. CIDA frequently makes use of local project steering groups for the purpose. The British Overseas Development Administration (ODA) has organised its monitoring system on a regional basis: headquarters formulates policy and decides on allocations, while the regional ODA offices are responsible for supervising and monitoring projects in their region.

Most donors use their own staff to monitor projects, but external experts, usually from the donor country, may also be involved.

Evaluation

Evaluations of projects and programmes can help decision-makers either to continue or replicate aid activities. Programme, theme and sector evaluations, which usually extend over several projects, are of particular value to general policy formation. These two forms of evaluation should complement one another.

Partly as a result of this distinction, most donor organisations operate a dual structure. Evaluations focusing on current interventions are often integrated with the regular activities of operational units, while those at programme or sector level tend to be carried out by central evaluation units.

Each donor calls for its own pattern of relationship between the central evaluation unit and the operational units responsible for decentralised forms of evaluation. For example, the World Bank's Operations Evaluation Department (OED) has support, coordinating and control functions. It also carries out its own evaluations, mainly of a sectoral or thematic nature. Operational units are responsible for the evaluation of continuing projects and programmes.

In the case of CIDA, project evaluations are generally supervised by *ad hoc* Project Evaluation Steering Committees recruited from representatives of the central evaluation unit, the operational unit responsible for the project in question, the sectoral/technical unit concerned and, where desired, the recipient country. As a rule the actual evaluations are undertaken by external experts, drawn, where possible, from the recipient country.

Within the British Overseas Development Administration, evaluation is centrally organised and the Evaluation Department is responsible for both project and programme, theme and sector evaluations. Operational units are regularly asked which activities are due for evaluation. Here, too, the evaluation itself is undertaken by external experts.

A major distinction made by USAID lies between evaluations of individual activities (project evaluations) and wider-ranging studies (sector studies, thematic and multi-project evaluations). Project information is gathered routinely by continuous monitoring and evaluation for which the field offices are responsible. Interim evaluations of current activities are wide in scope, combining problem-solving, progress control, assessment of effects, and determination of the relevance of project goals and of the assumptions underlying implementation. The findings about project implementation, set out in Project Completion Reports, may lead to this kind of evaluation. In certain cases *ex post* evaluations are carried out. Experts from recipient countries are included in the teams to which evaluations are sub-contracted. Like its management structure, USAID also employs a highly decentralised evaluation system. Responsibility for project evaluations is shared between the field office and the regional bureau in Washington. Depending on its size the field office may have its own unit, or staff member, with special responsibility for evaluation. The Centre for Development Information and Evaluation is responsible for initiating evaluations extending across a number of projects and provides support to the regional bureaus and field offices.

As in the case of monitoring, effective feedback from evaluation is vital to policy development and to the management of aid. Studies carried out by the OECD and the European Union show that donors use a variety of mechanisms and instruments for this (OECD, 1990; Cracknell, 1991), including the circulation of evaluation reports, the compilation and circulation of synthesised reports and the use of short summaries of evaluations showing their objectives, methods and main findings. In the case of thematic, sectoral and programme evaluations with significant policy implications, seminars are often organised to promote feedback. Several donors use computerised information systems combining project data and information from evaluation and monitoring.

How feedback is organised depends on how evaluation itself is organised. As a rule feedback from individual project evaluations is geared primarily to operational decision-making; evaluations which are wider in scope (sectoral and programme evaluations) tend to be used at policy level. Combinations of the two are also found. While feedback is sometimes 'voluntary' (that is, it is merely seen as desirable that

activities of a similar nature are considered when new activities are formulated), some donors make a formal link between evaluation and decisions about new activities.

2.5   Experiences with evaluation

In the 1970s evaluation was beset by problems, among them inadequate definitions of the purpose of evaluation and of project goals, and the difficulty of finding the right balance in data collection and analysis (Binnendijk, 1989). Donors also experienced organisational and management problems, largely because evaluation was not (or not sufficiently) recognised as a management instrument. The consequence was that it took place in isolation from management. From the outset evaluation was a donor-centred activity, so that both recipient countries and those directly involved in project implementation saw it as a mechanism of control.

The 1980s brought changes in the practice of evaluation, reflecting among other things a more differentiated application of principles and techniques, moves towards coordination and steps to bring it into line with the systems used in developing countries.

Donor's concerns about the effectiveness of evaluation have prompted studies of its role in practice and on the use made of its results (CIDA, 1991c, Scanteam, 1993, Finnida, 1991, USAID, 1992). They show that evaluation is frequently undertaken as part of decision-making concerning individual projects. Most evaluations are carried out by experts in the relevant technical field, but only a minority of them have specific expertise in evaluation. Hence, evaluations are more concerned with the achievement of technical objectives than with social and economic goals. They are generally snapshots of projects in progress, which rather than analyzing their immediate and wider effects, emphasise implementation and management. Participation by evaluators from developing countries is usually limited. Effectiveness is normally measured by the use of general observation and interviews. One widely used method of gathering data is that of interviews with people in key positions in the project and within the organisations responsible for implementation. Investigators' 'expert opinions', which are often impressionistic in nature, are also common. Despite a growing emphasis on the use of their results, evaluations frequently fail to comply with design, implementation, content and reporting guidelines.

Recently attention has also been paid to the methods used by donors to feed the information obtained from evaluations into management and policy. A review of

these methods, conducted by the DAC Expert Group on Aid Evaluation (OECD, 1990), found a growing emphasis on feedback but a wide diversity as regards the systems and instruments used to provide it.

A study commissioned by the EU into the incorporation of monitoring and evaluation into decision-making by European donors (Cracknell, 1991) found that, in recent years, there had been improvements in the use both of evaluation and of the information obtained. The study also found wide differences in the nature of the evaluation and feedback mechanisms and the intensity with which they are used.

Chapter 3   Monitoring and Evaluation in the Netherlands' Bilateral Cooperation: Policy and Procedures

3.1   Institutional framework

Four departments within the DGIS are responsible for the implementation of the bilateral aid programme. They are the three regional departments (for Africa, Asia and Latin America, subdivided into country desks), and the department concerned with spearhead programmes and technical advice (DST). The regional departments have overall responsibility for implementing country and regional policies. The country desks and the development cooperation sections in the Netherlands' embassies in developing countries are responsible for the aid implementation. Projects and programmes are contracted out to a large extent. Necessary technical support is provided by the Spearhead Programmes Coordination and Technical Advice Department (DST) or is obtained from outside consultants.

As regards the way bilateral aid is organised, the procedures, the division of responsibilities for policy implementation and staffing arrangements have all changed over the last decade.

3.2   The project cycle

The DGIS uses a project cycle consisting of six major steps: identification, formulation, appraisal, allocation, implementation and completion. Procedural guidelines exist for each of these.

Identification determines whether projects accord with overall development objectives, with the specific objectives of the developing country concerned and with the Netherlands' development policy. Their technical, institutional and financial feasibility are also assessed. The results are described in an identification memorandum drawn up by the embassy concerned.

At the formulation stage, proposals are further elaborated to allow for assessment of all their technical, institutional, economic and financial aspects. The relevant authorities in the recipient country are responsible for formulation, which may be carried out independently or with the help of external experts. The embassy is generally closely involved in an advisory capacity. This phase leads to project proposals.

Proposals are subjected to appraisals in which they are examined for their consistency with policy. Assessments of their effectiveness, efficiency, sustainability and management requirements, and technical, institutional, economic and financial feasibility are included. Appraisals are carried out by the DGIS unit concerned (the country desk) and internal technical advisors are involved. Results are laid down in appraisal memoranda.

Once an appraisal memorandum is finalised, an allocation is agreed with the country or organisation concerned. This step includes the aims of funding and possible conditions governing its provision.

The donor's main role in the implementation of the project is supervisory. This responsibility rests with the aid sections of embassies in recipient countries; it includes regular monitoring by embassies of projects in progress. External experts may be involved in monitoring if this is considered necessary.

When a project is completed an evaluation follows and responsibility for this belongs to the country desks. The central issue in this phase is whether the original objectives of the activity have been achieved and to what extent the criteria of effectiveness and efficiency have been met. If the activity is not to be continued an end-of-project report is drawn up. On the basis of this document it may be decided to conduct an *ex post* evaluation at some later time.

3.3   Organisational aspects of project implementation

The country desks are responsible for formulating policy. Embassies, on the basis of information from the project, supervise and support implementation. In turn, the embassy provides the relevant DGIS section with information about implementation. The extent to which a country desk and an embassy are involved in a particular activity depends on the agreements reached with the developing country or organisation concerned.

3.4   Deployment of personnel

In 1987 Ministry staff and Foreign Service personnel were integrated. As a rule new staff are recruited 'from below' and are selected on the basis of their suitability for deployment to the full range of activities in the Ministry. Since 1987 a large proportion of Ministry personnel have been subject to rotation between the various sections of the Ministry and embassies abroad. Postings usually last from three to four years.

Before integration, the staff of the DGIS were mainly people with a broad experience of, and an affinity for, development cooperation. But the second half of the 1980s saw an increasing inflow of generalists with little or no knowledge and experience in this field.

Since development cooperation requires specialised knowledge which often cannot be got by 'recruiting from below', there is also a so-called 'lateral' inflow of development specialists. These include the sectoral experts for the Spearhead Programmes Coordination and Technical Advice Department (DST), particularly the Technical Advice Section (DST/TA), and the technical specialists attached to embassies.

During the 1980s the Technical Advice Section was relatively small, but since 1991 both its personnel and its sectoral coverage have increased. Since the establishment of the Sector Specialists Programme in 1986, the number of technical specialists at the embassies has increased substantially.

Both country desks and embassies can employ external advisors in the various phases of the project cycle. These advisors are widely used to provide technical support and policy-related advice, either on an *ad hoc* or on a more permanent basis.

The country desks and embassies determine the need for external advisory services, depending on need and on the availability of specialists. This means that DGIS supervisory and advisory structures take many forms (DGIS, 1989). A common feature is the lack of any rigid separation between policy advice for DGIS, technical support for projects, and the monitoring and evaluation of projects.

3.5   Monitoring and evaluation in the project cycle

Monitoring

Guidelines exist for project and programme evaluation, but not for monitoring. The Evaluation Guidelines state that monitoring includes 'regular checks on, and

assessments of, the implementation of the activity' (DGIS, 1982a). No precise indication of the function, objectives and requirements of monitoring can be found in these guidelines or in any other policy document. Monitoring within projects, geared in the first instance to providing information for project managers, is left entirely to the project teams themselves.

Reporting

The Reporting Guidelines (DGIS, 1982b) define two kinds of reports: one for progress, the other for status. Progress reports, showing the current state of implementation, are compiled by the implementing agency. They enable the embassy to provide any necessary guidance. Requirements depend on the mode of implementation and, in addition, specific agreements may be reached about reports on individual activities. Status reports are brief summaries of progress reports, drawn up, to a standard pattern, by the embassy. They are sent at regular intervals to the country desk, together with copies of the progress reports and a summary of any decisions that have been taken or are planned. The country desk uses status reports to check progress against policy.

Both kinds of report fulfil five functions (DGIS, 1982b). The first is as a tool of management: 'it involves the regular compilation of a clear picture of progress made and may provide grounds for short-term planning adjustments and starting points for decision-making'. Second, reporting promotes continuity and clarity by 'recording progress made and experience gained'. This is particularly important in view of the frequent changes in the personnel of those sections involved in implementation. Reporting is also necessary in compiling information and in communication: 'all parties involved in the activity (recipient institution, embassy, project agency or contractor, Ministry) are to receive a full and clear picture of the current position and of relevant matters affecting implementation'. Reporting provides a means of 'checking how far the progress being made is in line with the implementation schedule, to which the content of reports must always be related'. The final function relates to evaluation: 'reporting has a function for inspection and evaluation teams, giving them an insight into the history of projects and facilitating evaluation. This is not however the function for which it is primarily intended'.

Reports are required at regular intervals. The content of progress reports is flexible, since their purpose is to check how far projects are proceeding in line with implementation schedules. Reporting on the status of individual activities is part of the embassies' wider task of reporting on the progress of aid programmes

in particular countries as a whole. Reporting is synchronised with the bilateral consultations and mid-term reviews with the recipient country. Status reports are forwarded to the country desk concerned and may be supplied to the recipient country.

Evaluation

Decentralised evaluation is usually limited to individual development activities, mainly projects, and is an important management instrument. Evaluations are chiefly interim assessments of progress and are made, for example, on completion of a particular phase. They are carried out by experts, from the Netherlands and developing countries, who are not involved in project implementation. Evaluations should primarily support decision-making concerning the intervention at issue. Evaluation is also intended to contribute to the adjustment of current policy at programme, sectoral, country and thematic levels.

Decentralised evaluation is mainly geared to the information needs of the operational units. Results can also be put to direct use in making decisions about possible new activities or continuing existing ones. This form of evaluation also allows the feedback of results to the agencies concerned, both in the Netherlands and in developing countries. Around 150 projects and programmes are evaluated in this way each year.

The purpose of evaluation as detailed in the Evaluation Guidelines is not primarily one of control and verification, it is first and foremost a learning process. Lessons can be used both in the project itself and in a wider context.

The Evaluation Guidelines define evaluation as follows: 'Evaluation forms part of the management process. It enables an assessment to be made, on the basis of an activity analysis which is as objective as possible, of the extent to which the original appraisal may need adjusting' (DGIS, 1982a). The scope of evaluation extends to the scrutiny of project design as well as implementation and the formulation of an opinion (*ibid.*). The purpose of evaluation is to analyze and assess an activity's effectiveness and to determine its relevance to the needs it addresses and the significance of any foreseen or unforeseen side-effects.

Wherever possible, DGIS and institutions in the developing country concerned have the responsibility to jointly design and conduct the evaluations. Procedures are laid down in detail. They cover timing, planning and preparation, the formulation of

terms of reference, the composition of the evaluation team, the actual conduct of the evaluation and the follow-up activities.

Besides procedural requirements the Evaluation Guidelines include specifications governing the form of reports and a checklist for the compilation of the terms of reference, position papers and work programmes of evaluation missions. Evaluation reports are to be drawn up in the working language of the developing country concerned. They need to follow a standard pattern, with a brief general and factual description of the project followed by findings, assessment and recommendations.

Chapter 4 Results of the Desk Study

Altogether 180 reports from the period 1987–90 were analyzed. A checklist of criteria covering the terms of reference, the composition of the evaluation team, the evaluation methodology and the aspects covered was used for the purpose. The results are shown in section 4.1. A random sample of 37 reports was subjected to an examination based on criteria derived from the Evaluation Guidelines, on the DAC principles, and on the requirements for evaluation generally accepted in the literature. Findings are given in section 4.2.

4.1 General characteristics of evaluation reports

The reports

The bulk of the reports related to projects in Asia and Africa; only a limited number concerned projects in Latin America. The spread of reports across the regions was broadly in line with the estimated numbers of projects under implementation during 1987–90 (see figure 4.1).

	Current projects	Evaluation reports	
Asia	256	99	
Africa	212	73	
Latin America	55	8	
Total	**523**	**180**	

Current projects and evaluation reports in percentage

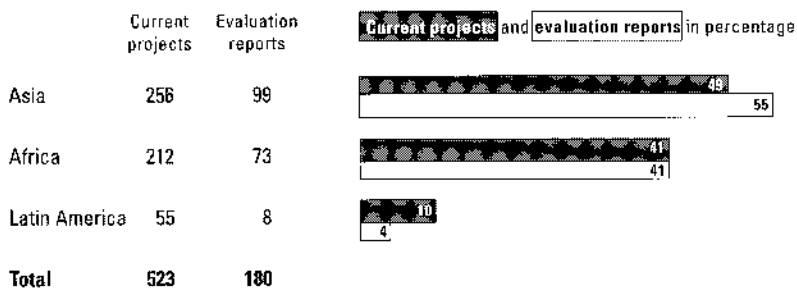Asia: 49 / 55
Africa: 41 / 41
Latin America: 10 / 4

Figure 4.1 Evaluations by region, 1987–1990

Most evaluations were carried out while the project concerned was implemented, either in the middle of a project phase or at its end. End-of-project and *ex post* evaluations were rarely found (see figure 4.2).

	Number of evaluations	Percentage
Mid-phase	55	39%
End of phase	85	60%
End of project	2	1%
Ex-post	1	–
Total	**143**	

*) Information based on **143** evaluation reports

Figure 4.2 Moment of evaluation in project cycle

Characteristics of evaluated projects

The speed with which effects emerge depends on the nature of the project, but it is clear that evaluations carried out shortly after the commencement of a project reveal little about its effects and wider impact. They are best carried out when the effects have become visible and its wider impact can be determined.

The points at which the 180 evaluations occurred were checked (figure 4.3) and it was found that 30 per cent of them were undertaken within two years of the start of activities. Usually this is too soon for solid information about the effects of the projects and their sustainability to emerge. A similar percentage of evaluations took place when the projects concerned had been in progress for six years or more. At this point the effects of projects and their impacts on target groups are usually visible and their likely sustainability can be assessed.

	Number of evaluations	Percentage
0- 2 years	41	29%
3- 5 years	53	38%
6- 9 years	25	18%
10-14 years	21	15%
≥ 15 years	1	–
Total	145	

*) Information based on **145** evaluation reports

Figure 4.3 Timing of evaluation in relation to start of project

In a relatively large number of projects the contribution from the Netherlands' Development Budget was considerable, over 50 per cent of them exceeded ten million guilders. It is clear that the majority of evaluations related to projects of considerable financial significance.

Most of the projects were concerned with health, housing, rural water supplies, education and training, community development, land settlement and development, women in development, regional development, family planning, small-scale enterprises and so on. In other words, the projects were directed at clear target groups. The primary sector was less strongly represented and projects in it were mainly to do with infrastructure or research.
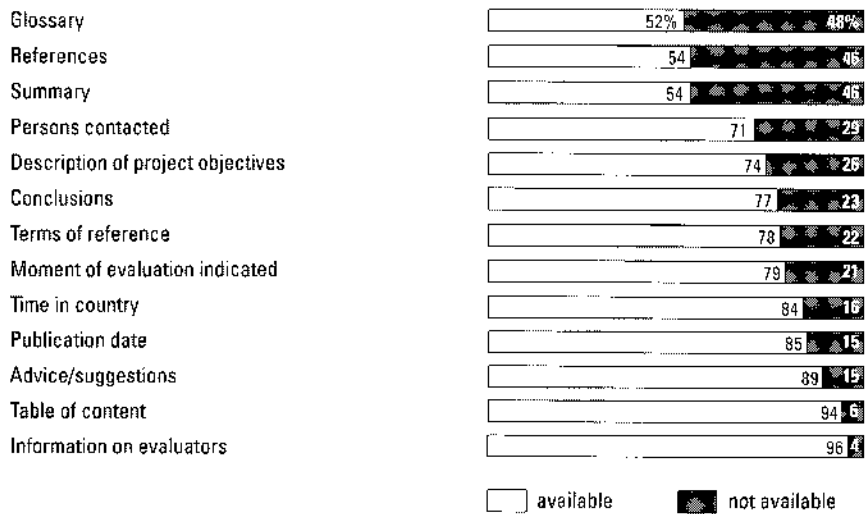
Comprehensiveness and comprehensibility

The reports gave a reasonable amount of information about the point reached in the project cycle when the evaluation was carried out; in most cases the country, the project and the publication date of the report were also given. In some 80 per cent of the reports the terms of reference for the evaluation was reproduced, either in the main text, or attached as an appendix. Over 80 per cent of them showed the time taken up by on-the-spot investigation. Almost all included a contents list. Only 32 reports (18 per cent) included a description or justification of the methodology used. Three quarters of them presented conclusions and 90 per cent included recommendations. Major omissions were journey reports, explanations of terms used and lists of documents consulted; summaries too were often missing (for details see figure 4.4).

Most reports indicated the size of the evaluation team, but information about their professional and institutional background was incomplete. Only 60 per cent of the reports showed which member of the team acted as leader.

Terms of reference

An evaluating team's terms of reference will affect the course and content of an evaluation, the manner in which findings are reported and the content of reports.

The topics mentioned in the terms of reference presented a diverse picture. They were either long lists of points to be covered by the evaluators or summaries of general aspects to be studied. Terms of reference were commonly not drawn up in

Item	available	not available
Glossary	52%	48%
References	54	46
Summary	54	46
Persons contacted	71	29
Description of project objectives	74	26
Conclusions	77	23
Terms of reference	78	22
Moment of evaluation indicated	79	21
Time in country	84	16
Publication date	85	15
Advice/suggestions	89	15
Table of content	94	6
Information on evaluators	96	4

☐ available   ■ not available
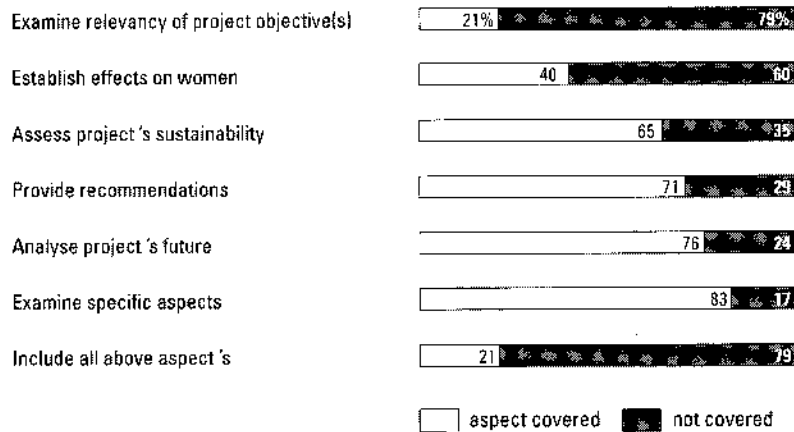
*) Information based on **180** evaluation reports

Figure 4.4 Basic information in evaluation reports

accordance with the Evaluation Guidelines: the justification, occasion and purpose of the evaluation were often not made explicitly clear.

Women in development, the relief of poverty, the promotion of economic self-reliance, environmental protection and the sustainability of development activities are all major themes in the Netherlands' development policy. They should all be included in the terms of reference for evaluation reports.

In general, the terms of reference devoted limited attention to the relief of poverty and the promotion of economic self-reliance. One in five called for an examination of the relevance of original project objectives. Two thirds mentioned sustainability as a specific topic for consideration, but there was no general requirement that sustainability be included among the criteria used in evaluation. In a minority of cases the terms of reference demanded a consideration of the projects' effects on the position of women. Except where projects were directly concerned with environmental matters virtually none of the terms of reference mentioned the need to study their impact on the environment (see figure 4.5).

Guidance in evaluation procedures was lacking; methodology was indicated in only 20 per cent of the cases. Evaluations were largely geared to the needs of both local and the Netherlands' authorities, and paid little explicit attention to the needs of target populations. It was rare to find any indication in the terms of reference of the

Examine relevancy of project objective(s)	21%	79%
Establish effects on women	40	60
Assess project's sustainability	65	35
Provide recommendations	71	29
Analyse project's future	76	24
Examine specific aspects	83	17
Include all above aspect's	21	79

☐ aspect covered ■ not covered

*) Based on analysis of **140** Terms of Reference

Figure 4.5 Scope of terms of reference

extent to which members of the target group should be consulted in the course of field visits, or how investigations focusing on the target group should be carried out.
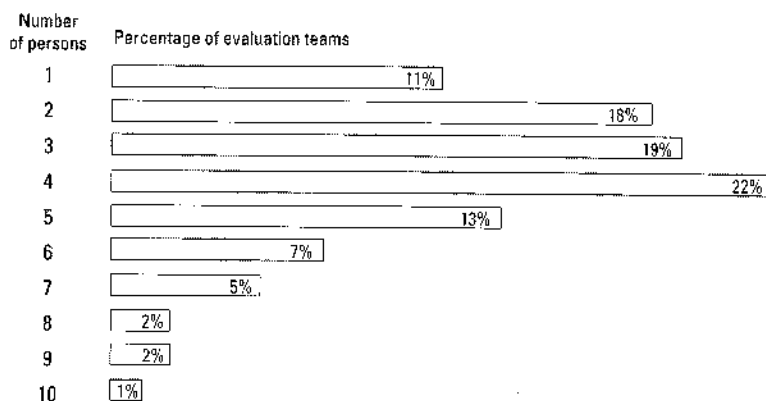
In most cases no guidance was given to evaluators about which feedback procedures should be followed. Where debriefing was specifically mentioned, those to be involved were generally institutions (embassy, project managers, counterpart organisation); feedback to project target groups was only occasionally mentioned.

Evaluation teams

With one exception evaluations were carried out by teams sent out to the projects concerned. The duration of missions varied, but the great majority of them (78 per cent) lasted between two and three weeks. In size they ranged from one to 24 members, but the average was four; almost 60 per cent were made up of from two to four people (see figure 4.6). No relationship was found between the size and length of missions and the financial significance of the projects.

In 27 per cent of the missions the teams consisted exclusively of evaluators from the Netherlands, two per cent comprised local experts only and the remainder were undertaken jointly by Dutch and local evaluators. In general there was a slight under-representation of local evaluators in the joint missions.

In those cases in which the information was given, over half the evaluators had a technical background and the social and economic disciplines were in a minority.

Number
of persons   Percentage of evaluation teams

Number of persons	Percentage of evaluation teams
1	11%
2	18%
3	19%
4	22%
5	13%
6	7%
7	5%
8	2%
9	2%
10	1%

*) Information based on **172** evaluation reports

Figure 4.6 Composition of evaluation team

A little over a third of the Netherlands' evaluators were university employees, a quarter worked in small specialised consultancy firms, and just over ten per cent were government personnel. DGIS staff were only occasionally present in these missions, either as evaluators or as resource persons. Most local evaluators were civil servants, usually in the ministry in the recipient country responsible for coordinating the project. One in three local evaluators worked in the private sector.

Team members were mostly men, only ten per cent of the evaluators were women and they were often responsible for the assessment of the Women in Development (WID) aspects of projects. The leaders of 85 per cent of the evaluation missions were Dutch. Among the team leaders, social and economic experts were slightly over-represented .

Available data were analyzed to find out if a core group of evaluators was involved in several different missions. Most evaluators (62%) were found to have taken part in one mission, and only 15 per cent had participated in three or more. Of the 103 named team leaders, twelve had headed more than one evaluating mission under the Country and Region Programme in 1987–90, and three of those had done so more than three times. There is no obvious core group of evaluators.

Costs and activities

The time and costs involved in a number of evaluations were determined. Staff time provided either by the Netherlands or by the recipient country averaged 98 person

days. This was broken down across the three phases of the evaluation process (preparation – 8%, work in the developing country – 77%, reporting – 15%). These are averages from which the figures for individual missions may differ markedly. However, in no case were local evaluators involved in preparatory activities, and their role in drafting the final report was minimal since this usually took place in the Netherlands after the evaluation mission had terminated.

The average cost of an evaluation amounted to some 85,000 guilders. This suggests a total cost for all 180 evaluations of the bilateral aid projects between 1987 and 1990 of just over 15 million guilders – less than one per cent of the total funds allocated to the projects concerned.

Form and content of reports

The length of the main text of reports varied considerably. Almost 50 per cent of those studied were between 25 and 50 pages in length and almost a quarter exceeded 50 pages. With four exceptions, all reports were written in the working language of the recipient country.

The Evaluation Guidelines list five essential elements to be included in reports. These are:

1) A general description of the project covering the problem addressed, short- and long-term objectives, project design and the actual course of implementation.
2) An account of the significance of the project covering its general significance (impact on development in general, results for target group/s), its significance for the recipient country (relevance to local priorities, potential for continuation by the recipient country, quantification of its effects on, for example, employment), its relevance to the Netherlands' development policy (compatibility of results with the twin-track policy of poverty alleviation and economic self-reliance and with the criteria laid down in the policy programme under which the project was funded).
3) A consideration of the effectiveness of the intervention.
4) An analysis of its efficiency.
5) A presentation of conclusions and recommendations.

Only six per cent of the reports covered all these requirements and 85 per cent covered fewer than three. Among those particularly neglected were effectiveness, efficiency and relevance.

Methodology

The extent to which particular research methods were used (field visits, comparative evaluation methods, financial analysis etc.) was also examined. Information on which methods were used, and why, was generally inadequate – only one report in five set them out. In 139 reports it was said that evaluation teams had made field visits. Only one report in six referred to the use of comparative research methods and, of the evaluations examined, only ten per cent included a financial or an economic analysis.

4.2   Quality of evaluation reports

The quality of the contents of evaluation reports was assessed using the following criteria:
- inclusion of a clear and complete project description (rationale, short- and long-term objectives, resources used, activities, target groups, outputs);
- an account and a justification of the investigative methods used, the means of data collection and the use of data-gathering instruments;
- use of secondary information from progress and monitoring reports;
- concern with the effects of the projects, particularly those which were gender specific, on the target population;
- concern with the integration of the project into the institutional framework of the counterpart institution and attention paid to all actors involved in the project (counterpart institution, target groups, embassy and DGIS);
- consideration of the effectiveness, impact and sustainability of projects and their conformity with policy;
- logical presentation of lessons, conclusions and recommendations.

Project description

Project descriptions included in evaluation reports left much to be desired. Half of them failed to mention the reasons for which projects were undertaken and the other half did so inadequately. Problems and the ways they were addressed were covered in a fragmentary fashion. Reports commonly gave a general description of problems affecting the sector concerned, rather than an analysis tailored to the specific project. Most evaluation reports distinguished short- and long-term project objectives. In some cases objectives were incorrectly grouped and were often confused with project activities.

In many reports the description of project activities and the resources used was unclear and incomplete. Project budgets were dealt with very briefly if at all; often no indication was given of the financial contribution made by the counterpart institution.

In 26 of the 37 sample reports the project description gave no information at all about the beneficiaries and in the remainder it was given only in general terms. Thus it was unclear how the target groups differed from the populations at large. Criteria by which the target groups were defined, their size and composition were generally not given.

Approaches and strategies were seldom clearly stated in the project descriptions. Accounts of the organisational structure, institutional set-up and orientation of projects were lacking or incomplete. Information on project size and the composition of the staff, both foreign and local, was often incomplete. The physical, social and economic features of the project area and its location within the country were not always described.

Methodology

Few reports included an account of the evaluation methodology used. Most missions combined interviews, consultation of documents and field visits. Only rarely had evaluators spoken to all the parties involved in projects, their contacts were mainly with higher officials (project directors and the Netherlands' project leaders, officials of the counterpart ministries, staff in the Netherlands' embassies); only occasionally had discussions been held with a wider range of project personnel. Reports rarely mentioned substantive consultations with national or regional planning bodies or the relevant country desk in The Hague. Nowhere were reasons given for the choice of interviewees.

Most missions made field visits, but neither the locations nor the project activities visited were indicated. Reports did not reveal whether field visits included meetings with representative members of target groups or visits to representative locations. The use of questionnaires, interview plans or checklists was not mentioned, which suggests that virtually all interviews were conducted in an unstructured fashion. Where the evaluation itself was preceded by some form of investigation, its design and mode of implementation were not shown.

Project results and effects were given in general, qualitatively descriptive, terms such

as 'impressive results', 'useful', 'major progress' and 'satisfactory'. Such statements were not backed up with arguments or facts. Project success was sometimes measured in terms of outputs achieved, with some 30 per cent of reports basing their judgement on a combination of a description of the project's performance, a summary of its outputs and an indication of its effects. Only 16 per cent included a quantified description of effects, sometimes using indicators.

Many missions referred in their reports to the limitations of evaluation research. Major constraints were the lack of baseline data on the pre-project situation, the lack of data for the purpose of financial analysis, unspecified target groups, the lack of monitoring data, the limited time available for evaluation, and the lack of indicators for determining the results and effects. Some evaluation teams showed an awareness of the danger that analyses based on incomplete and unreliable information could produce a distorted picture.

Use of monitoring data as secondary information

The reports generally gave very little information on how those in charge of projects themselves determined and recorded progress, results and effects. Regular progress reporting, to which most reports made indirect reference, appeared to be the main source of secondary information. Others included monitoring forms, reports on meetings, supervision and inspection reports published by the projects themselves, project-generated statistics and evaluations undertaken by project staff. In a few cases evaluators specified the source from which secondary information had been drawn. No judgements on the quality of such information were offered. Occasionally it was reported that progress-monitoring systems had been set up within a project or that information produced by such monitoring was available. The ways in which internal management information systems functioned was hardly ever evaluated. The apparent lack of concern with internal monitoring arrangements and project-generated information was hard to reconcile with evaluators' view that their work was hampered by a lack of data.

Target group perspective

Project descriptions provided only very general accounts of target groups. Reports of projects geared to institutional development did not identify the intended beneficiaries. This was also true of projects aimed, directly or indirectly, at end users like consumers, farmers etc. Some evaluations tried to identify and analyze the effects

on these groups, but the lack of baseline information meant that any conclusions were very uncertain.

Some evaluation teams recognised that target groups were not homogeneous, distinguishing differences in effects within the target group and in the extent of their participation in project activities. How far projects had increased or decreased differentiation within target groups was not addressed. Where effects were reported there was little attempt at quantification. Causal links between project activities and any changes and effects noted by evaluators were seldom made.

A small majority of the reports described the impact of projects on the position of women, usually in separate chapters or sections of the report. As a rule these effects were related to the performance of household tasks or to work done by women outside the home (increases or decreases in workload in consequence of techniques introduced by the project, changes in women's employment). Effects in the form of changes in women's role in policy-making, administration and management or in decision-making processes within the family were not considered. The treatment of the role of women in some reports seemed perfunctory; they offered general statements about the position of women but gave no account of how it might have been affected by the project.

Target groups seldom played an active part in evaluations. Indeed, evaluation reports gave the impression that evaluators focused on authorities rather than on the opinions of target groups. Few gave an indication that the views of officials had been corroborated by comparing them with those of target groups.

Organisational setting of projects

While many projects were organisationally complex, few reports managed to describe the various interconnections clearly. The internal structures of the counterpart organisations were generally shown in diagrams, but descriptions of the organisational designs of the projects, their position in relation to counterpart organisations and divisions of responsibility were incompletely analyzed.

Reports generally gave a fragmented account of internal project organisation and staffing. The different actors involved in a project – like the donor and the counterpart organisation, the project team, the target groups – were by no means always shown in full, so there was no systematic analysis of the project's position *vis à vis* each of them. While some attention was given to possible divergences among the

various actors' goals, expectations or interests, this did not prompt any systematic analysis of their relative positions. There was occasional mention of differences of opinion or conflicts of interest among the actors over policy and implementation.

Consistency with policy, efficiency, impact and sustainability

A small majority of missions gave some general consideration to the relationship of projects to the current policies of the developing country concerned, some of them focusing specifically on the question whether the recipient country regarded continuation of the project as desirable. Little if any attempt was made to show how far projects reflected the Netherlands' policy, most evaluators evidently taking the view that projects were by definition in line with it.

The amount of attention devoted to the efficiency and effectiveness of projects varied. Efficiency in design and implementation was determined in various ways. A few evaluations asked the question whether the results justified the efforts, sometimes these judgements were reinforced by quantitative information. In most cases, however, evaluators lacked the data, indicators and standards needed for quantitative analysis and therefore limited themselves simply to giving impressions. In general evaluations showed little concern with efficiency, only occasionally were comments made on whether project inputs had been deployed at minimum cost. Some reports included remarks on more efficient alternatives, but they tended not to be underpinned with figures or analysis.

While virtually every evaluation looked at whether projects had achieved their objectives, in most cases the analyses of effectiveness and impact were unsystematic and incomplete. Most reports considered whether planned outputs had been achieved, but these assessments were generally presented without quantitative data and, hence, were inadequate. For the most part, however, the concern was with short-term goals and whether they had been or were likely to be achieved; little attention was focused on long-term effects.

Some reports mentioned factors which made it difficult to determine effectiveness and impact. There were elements like inadequate project preparation, flawed analysis of the problems addressed, a lack of standards against which results could be measured and inadequate descriptions of starting positions.

The DAC definition of sustainability (the capacity permanently to maintain a certain minimum level of net benefit once financial, institutional and technical support has been withdrawn) was used to determine how far evaluations were concerned with it.

The following aspects of sustainability were borne in mind in assessing the evaluation reports:

1) the viability of project activities in relation to prevailing macro-economic conditions;
2) the long-term financial viability of project activities, expressed in terms of the extent to which resource flows for maintenance, current costs and depreciation are guaranteed;
3) the scope for the transfer of activities to counterpart organisations or local people;
4) the degree of political support for the project and its institutional base;
5) compatibility of the project with the carrying capacity of the local environment.

Sustainability was dealt with inconsistently in evaluation reports. Aspects frequently covered were the scope for the transfer of activities to the counterpart, long-term financial viability and viability in relation to prevailing macro-economic conditions. Judgements about sustainability tended not to be backed up by solid arguments.

Presentation of conclusions, recommendations and lessons

Most reports included a chapter with conclusions and recommendations. In some, recommendations were presented separately. Conclusions were broadly in line with the evaluation's findings although they were sometimes wrapped in woolly language.

The use and presentation of recommendations in the reports was variable. Some fully and systematically presented recommendations separately, while others included them in the main text. Recommendations tended to be general and not always consistent with findings and conclusions. In some instances findings and conclusions were not translated into recommendations despite their suitability for the purpose, and a number of reports clearly identified problems without suggesting possible solutions. Most recommendations could not be directly translated into practical measures.

Few evaluation reports made the point that drawing lessons from project experience was a major aspect of evaluation. Although many evaluations brought out lessons with implications stretching beyond the level of the individual project, they were not systematically and explicitly presented in the text.

Chapter 5   Results of the Field Studies

The role of monitoring and evaluation was investigated in projects in Burkina Faso, Egypt and Pakistan supported by the Netherlands. Sixteen of them were studied. The results are summarised below.

5.1   Burkina Faso

5.1.1 *Background*

The Netherlands' programme of aid to Burkina Faso aims at promoting food production and food security, improving water supplies and restoring ecological balance through reforestation, energy saving and anti-desertification measures. Policy is implemented by means of a multisectoral approach in a number of specific regions; this approach encompasses integrated agricultural development programmes in the Volta valleys and agricultural, forestry and water-supply activities on the Mossi plateau in central Burkina Faso.

Many projects in Burkina Faso are implemented under the direct management of the DGIS. Their design is process-orientated, with emphasis on local participation. Efforts are made to anchor development projects in governmental and semi-governmental structures, and regional government agencies get institutional and financial support to this end.

The five projects, with some of their main features, are listed in table 5.1.

Monitoring and evaluation in Burkina Faso

The monitoring system used by the Burkinabé government is based on monthly and quarterly reporting by the projects and by executive agencies. Reports are forwarded

to the central research and planning departments (Directions des Etudes et de la Planification) of the sectoral ministries and then passed on to the planning ministry.

Monitoring has been badly affected by the many administrative changes of the 1980s. Government, both central and provincial, lacks the resources and expertise to conduct or commission evaluations. As a result monitoring and evaluation are both undertaken by the donors funding projects and programmes. Government agencies participate where possible and where donors consider this necessary. Their involvement in the design of monitoring and evaluation is largely passive.

Table 5.1   Characteristics of projects studied in Burkina Faso

Name of Project	Duration	The Netherlands' Contribution *Dfl. millions*	Evaluated in	Evaluation interval *years*	Remarks
Sensibilisation et Formation des Paysans autour des Barrages	1981–94	19.1	1984, 1987, 1991	3–4	
Hydraulique Villageoise dans le Boucle du Mouhon	1981–93	43.4	1983, 1986, 1992*	3–6	Annual monitoring from 1986
Urbanisme Ouagadougou	1982–90	7.2	1989	N.A.	6 monthly support-cum-monitoring missions
Valée du Kou	1980–93	18.9	1983, 1984, 1988, 1992*	4	
Programmation et Exécution du Développement Intégré Kaya	1982–4	45.3	1986, 1990	4	

*The evaluation took place after the field study was completed.

Evaluation capacity

The growth of consultancy services in the private and semi-governmental sectors is mainly the result of the demand from donors for local expertise for use in project formulation, supervision and evaluation. A market in social, economic and

technical advisory services is gradually developing. This growth is helped along by the government sector's increasing inability to absorb the better educated, so that those with a higher education are more likely to seek work in the private sector. Most consultant firms are small, often with a permanent staff of only two to four plus a larger number on short-term contracts; some are associated with foreign firms.

Opinions about the quality and capacity of the consultancy services differ. Donors differ in the extent to which they involve local experts in project preparation, implementation, monitoring and evaluation.

5.1.2 *Evaluation*

Planning and design

Most evaluations were envisaged in the project documents. They were nearly always initiated by the country desk. Where project documents provided for evaluation, the decision to proceed seems to have been taken almost automatically, without reference to the current situation as regards the project. Even though their planned date was known well in advance, they were not prepared in good time. For example, the formulation of terms of reference and the choice of evaluators showed signs of having been done hastily.

Formulation of the terms of reference was usually delegated to one of the parties involved in the project (the embassy, the country desk, the project team or the supervisors). Country-desk staff felt that they lacked the necessary time, expertise and knowledge of the project, and most project teams thought themselves too close to projects to be sufficiently free from bias. No substantive alterations or additions to the draft terms of reference were made by the embassy, the counterpart or the country desk during the process of finalisation.

There was no clear link between the terms of reference for the evaluation and the content of the appraisal memorandum. This was so even when the memorandum included clear external conditions, critical assumptions or even conditions for approval. In the mid 1980s the terms of reference emphasised effectiveness and efficiency; this was followed later by a growing interest in impact and sustainability.

The terms of reference did not clearly indicate the limits of the planned evaluation. For example, should it include the working procedures of project staff, the deployment of resources, the impact on the target group, or the relative roles of counterpart institution, embassy and country desk?

No clear pattern was evident in the selection of evaluators. Expatriate evaluators were usually chosen by the country desk, with advice from the embassy and from the Technical Advice Section of DGIS. The main criteria were kinds of expertise, regional knowledge and reputation within DGIS; expertise in evaluation and skills in communication were held to be less important. The Burkinabé authorities often had great difficulty in finding suitable local evaluators, and their names and professional backgrounds were often not known until shortly before the expatriate mission members arrived in the country.

Burkinabé evaluators were usually government officials while their expatriate colleagues were more likely to be consultants not directly involved with the project. In practice this seems to have caused few problems in the evaluation itself, but the differences were often apparent in the formulation of recommendations for subsequent phases. Expatriate evaluators did not feel that their Burkinabé colleagues were independent, while the latter often saw the former as a buffer between the Dutch and Burkinabé authorities.

Preparation

Preparation for field work by the evaluators was limited and the expatriates were unfamiliar with the Evaluation Guidelines. In only one case had the country desk drawn up a work plan for the evaluation, including a methodology. Broadly formulated terms of reference meant that much was left to the evaluators' discretion. The Burkinabé evaluators did not get the terms of reference and other relevant documentation until the start of fieldwork and therefore were ill-prepared for it.

Implementation

Nearly all evaluations lasted from two to three weeks, irrespective of the type and complexity of the project and the nature and extent of the terms of reference.

Shortage of time meant that evaluators had to work quickly, and so the Burkinabé members were unable to catch up on the information shortfall caused by their lack of preparation. They often played a less active role during the evaluation, working instead as facilitators, resource persons or responding to the views of their expatriate colleagues.

Weak monitoring arrangements within projects meant that much of the data needed for evaluation was lacking, so evaluators were forced to try to gather it for

themselves. This overloaded their programme and left little time for analysis and reflection.

Box 5.1.1 Programme of the 1991 mission to evaluate the *Sensibilisation et Formation des Paysans autour des Barrages* project

Agreement to leave the design of the evaluation to the team after it had arrived in Burkina Faso was reached in discussion with them at DGIS. The country desk did, however, stress two points in the extensive terms of reference:
– the analysis of the project's institutional setting
– the assessment of the economic feasibility of the irrigation system
It was agreed that the mission would spend three weeks in Burkina Faso and that the Dutch members would spend two days preparing themselves before departure and three reporting after their return to the Netherlands.
On arrival in Burkina Faso it was discovered that the Burkinabé evaluators had not received the terms of reference and the basic project documentation. Three days had to be put aside to allow them to familiarise themselves with this material. An analysis of the current situation and proposals for the next phase, compiled by the project management, had also to be studied. The fourth day was a national holiday. On the fifth, the evaluators discussed the terms of reference among themselves. On the sixth day a work plan was made and fieldwork began on the seventh. Shortage of time limited the visit to the project to two days. Interviews with the embassy and the implementing agency, at national and at provincial level, then took place. Drafting began after ten days and debriefing took place on the seventeenth day (ministry, embassy, project management). The remaining three days were used to finalise the report.
Shortage of time also meant that the points stressed by the country desk were not given adequate attention. According to the mission leader, frequent discussions of the institutional framework of the project produced little of value: 'We Dutch knew too little of the local situation ... problems were glossed over by project managers ... at the ministries everything was kept very vague.' An economic analysis of the irrigation systems seemed to be taboo. Dutch mission members tried to discuss the issue, but local evaluators, project managers, embassy and Burkinabé authorities did not think an analysis was desirable. On the other hand, the government authorities wanted to go on subsidising rice production. But the Dutch project staff, realising how dubious the economic feasibility of the project was, resisted its analysis. The net result was that economic viability was not addressed.
The evaluation failed to throw light on the country desks' two points and, as a result, endorsed practically all the project team's proposals for the next phase. Given the time available for the evaluation, and the scope of the terms of reference, it is doubtful if the mission could have dealt thoroughly with the points in question.

In practice missions usually limited themselves to assessing progress (see box 5.1.1), often linking this to the desirability of extending the project. Analysis of the efficiency of project activities was frequently rendered impossible by lack of data. The relevance of projects to developmental needs tended to be described in very

general terms, especially where politically sensitive issues were involved. Impact was generally dealt with impressionistically, reflecting the lack of baseline information and of monitoring data.

It was often difficult for evaluators to strike a pragmatic balance between rational and analytical research and an interactive approach. The roles to be played by, for example, expatriate project staff and local personnel were often unclear, particularly at the start of an evaluation. Were they objects of study, resource people or active partners in a joint investigation? The problem was aggravated when evaluation was combined with advising about, or formulating, a subsequent project phase. Should the mission arrive at its recommendations using only abstract criteria such as developmental relevance, or should it try to achieve broad political and practical support from relevant parties? If the latter, the implications of what was recommended for each of those parties (project, counterpart institution, embassy, country desk) remained as a question.

Towards the end of the evaluation, when the time came to draw up recommendations for the next phase of the project, evaluators often met their 'patrons' informally: the Burkinabé evaluators with officials of the relevant ministry and the (usually Dutch) mission leader with embassy staff. Where the signals about policies and priorities diverged, tensions and differences of opinion could arise between Burkinabé and expatriate evaluators, usually about the future of the project.

Reporting and feedback

Final discussions at the project level were often brief and somewhat ceremonial with little substantive content. Almost without exception project staff were dissatisfied with this kind of debriefing. Having invested considerable time and effort in the evaluation, they did not see the evaluators completing their work.

The evaluators' provisional recommendations were the main focus of debriefing, which involved the counterpart institution and the embassy. Although differences between provisional and final recommendations could not be distilled from evaluation reports, those evaluators who were interviewed said that recommendations were often amended as a result of the debriefing.

Project goals were usually couched in very ambitious terms in the project documents, so evaluation reports often looked at what had gone wrong, rather than at successes and the lessons to be drawn from them. This negative tone was reinforced by the fact that DGIS expects evaluators to produce 'critical' analyses.

In practice evaluations were nearly always limited to checking progress against the implementation schedule. The project's setup, including the funding channel and the method of implementation, its integration into the counterpart organisation etc., were seldom analysed and assessed thoroughly. An analysis of major decisions made by the central counterpart organisation, the embassy or the DGIS concerning project implementation, was frequently lacking.

Evaluators' recommendations for subsequent project phases were the central focus of the follow-up process, both within the DGIS and in discussions between the country desk and the embassy. Findings were not subjected to further analysis or used for purposes beyond the individual project.

5.1.3 *The functions of evaluation*

The Evaluation Guidelines list three functions for evaluation: management support, policy support and communication.

Management support

Project modification in the light of evaluators' recommendations, *inter alia*, is probably the most essential function of the evaluations under review and one that recurs regularly.

When the Burkinabé authorities explicitly gave a role to evaluation, it was to 'create the conditions for continuation of the project'. The country desk mainly saw evaluations as a contribution to the appraisal memorandum for the subsequent project phase.

Combining evaluation with recommendations for the next project phase presents no problems when there is prior agreement on its broad outline. However, when opinions about the continuation of a project differ substantially, the combination has its drawbacks. Part of the responsibility for negotiating project design is delegated to the evaluators and they may be subjected to conflicting interests. In that case the conflation of *ex post* evaluation with *ex-ante* formulation may undermine objectivity and subordinate the lessons to be learnt to questions of formulation.

Policy support

In Burkina Faso none of the cases examined was 'drawing lessons for the programme of which the project forms part'. Nor were any indications found that the evaluation

reports were later used to formulate policies. The lessons remained limited to the project concerned and, because of their operational focus, evaluations contained little that could be of wider relevance.

Communication

There was no obvious evidence that evaluations led to an improved mutual understanding of the principles and objectives of a project. Where substantive differences existed between the Netherlands' and Burkinabé authorities over the future of a project before the evaluation, they tended to be obscured rather than brought out by the vague compromises formulated in evaluation reports.

Even so, evaluations have produced indirect dialogue, as the two parties respond (often independently) to the proposals of the evaluation mission and the subsequent formulation mission.

5.1.4 *External Monitoring*

The intensity with which projects were monitored varied. Some were subjected to minimum control, others were intensively supervised by the embassy. Supervision also varied, depending on the kinds of problem arising in a given project and the division of responsibilities within the embassy. The latter depended on the formal allocation of functions either to policy staff or to sector specialists and on the actual allocation of the project portfolio. It was not always entirely clear what role was being played by individual embassy staff members at particular times; nor was the gradually altered division of responsibilities between embassy and country desk always clear.

Monitoring was often approached from different angles by the embassy and the country desk. Where the latter linked monitoring chiefly to objectives laid down in the project document, the embassy tended to relate progress to what was necessary and feasible in a political and institutional environment that was, in project terms, subject to frequent and unpredictable change.

The ways in which the embassy in Burkina Faso and the country desk view one another's roles also has an effect. From the embassy's point of view, the country desk was part of the large DGIS bureaucracy. The tightening up of procedural requirements have slowed it down, made it too remote from the field and rendered it excessively dependent on external consultants. For its part, the country desk

often saw the embassy aligned with the projects and the Burkinabé authorities and thus felt it had to check proposals closely against DGIS policies and priorities. Communication between the embassy and DGIS could, as a result, be difficult. Project supervision by external consultants then acquired an extra function as a second channel of information for DGIS.

The country desk assigns the following functions to external monitoring:
– identifying features common to projects in different countries;
– advising the country desk about assessment and evaluation;
– advising the embassy about project identification;
– advising the counterpart about project formulation;
– advising and supervising project management.

Combining these functions is more efficient since the same knowledge and experience is reused in the different phases of the policy cycle (see box 5.1.2). There are, however, drawbacks which were illustrated by the experience in Burkina Faso. Where the objectives, priorities and interests of project managers, counterpart, embassy and country desk diverge, the external advisor risks becoming embroiled in a situation in which roles conflict. If the same advisor is engaged in the different phases of the project cycle, her or his more private interests may come to play a significant role. Finally, relying on the same advisor may mean that the dividing lines between of the specific responsibilities of the counterpart, the embassy and the country desk become blurred.

5.1.5 *Internal monitoring*

DGIS has no guidelines or standards for monitoring within projects. In practice the design, establishment and role of internal monitoring depends mainly on the individual expertise and vision of the expatriate project leaders and embassy staff. Staff changes sometimes result in changes to internal monitoring arrangements.

With regard to the projects studied, internal monitoring developed in the following manner between 1985 and 1991. Initially, the objectives set out in project documents were too broad to allow progress to be measured. To the extent that organised monitoring occurred, the focus was chiefly on inputs, physical activities and, to a limited extent, outputs. It was seldom concerned with results and effects. From 1989 onwards, however, monitoring systems were established in several projects. The absence of guidelines and specific operational knowledge proved to be a constraint.

Box 5.1.2   Supervision of the *Urbanisme Ouagadougou* project

The *Urbanisme Ouagadougou* project was identified in 1982 by an external expert of the University of Amsterdam and was, from the outset, supervised by the same person. In the project document his role is described as *Conseiller Technique Principale*. In practice the expert has the following roles:

1) Monitoring the technical aspects of the projects for the country desk and the embassy.
2) Intermediary between project team, counterpart, embassy and country desk. The expert's long association with the project has enabled him to build up a relationship of trust with all actors. This accelerates decision-making.
3) Technical advisor to project management.
4) Coordinator of field research carried out by Dutch and Burkinabé research assistants and students. This produces valuable information for the project.
5) The project's institutional memory. The expert has a major role in ensuring continuity: in ten years he has worked with eight Burkinabé ministers, seven Burkinabé directors-general, five officers at the embassy and six officers at the country desk.

The supervisory function is mainly exercised through short missions. Their terms of reference are written by the expert himself and approved by the country desk and the embassy. His very clear reports are mainly concerned with institutional aspects, government development policy and the project's links with other urban development activities in Ouagadougou.

The expert also has to balance the project's interests and those of the University of Amsterdam. According to the Burkinabé project coordinator he has always succeeded. The expert says that it was never a problem.

The 'advisory' function benefits from short missions, which help to maintain the expert's prestige and to ensure that problems are addressed and decisions reached with the minimum delay.

From 1982 to 1989 the advisor had to liaise between all parties involved in the project. Since 1989 the embassy was increasingly active, but without any formal changes to the arrangements for supervision by the expert.

The expert also played a part in the project's evaluation in 1988. The draft terms of reference were agreed between him and the counterpart before submission to the embassy. He also advised the DGIS on the choice of the mission leader.

In this specific case all concerned are agreed that this is a workable and efficient arrangement. It is probably relevant that the University's interest in the project is limited and that the expert is strongly motivated to ensure that the project succeeds. He has also taken the trouble to build up a close network of personal contacts which is used for the benefit of the project.

5.2   Pakistan

5.2.1 *Background*

Major components of Netherlands' project aid to Pakistan include support for rural development in the poverty-stricken province of Baluchistan and the North West Frontier Province, support for small-scale industry in the North West Frontier Province and the Punjab, and support for various thematic activities. As well as regionally targeted activities, the aid programme includes several nationwide projects which are aimed at strengthening the institutional capacity of a number of national organisations. They are largely contracted out to Dutch consultants who advise local institutions. Several projects are carried out by national institutions; some are co-financed with international and multilateral agencies.

Six projects were studied; their main features are summarised in table 5.2.

Table 5.2   Characteristics of projects studied in Pakistan

Name of Project	Duration	The Netherlands' Contribution *Dfl. millions*	Evaluated in	Evaluation interval *years*	Remarks
Animal Husbandry In-Service Training Institute	1988–92	4.4	1989, 1991	2	
Pata Integrated Agricultural Development	1986–95	39.1	1987, 1990	3	From 1992, 6-monthly monitoring
Quetta Sewerage and Sanitation	1986–93	27.8	1992*	N.A.	Annual monitoring
Matric Education	1985–92	2.0	1988	N.A.	Annual monitoring
Development and Utilisation of Human Resources	1987–93	8.8	1992*	N.A.	Annual monitoring
Pak-Holland Metal	1986–95	23.4	Internal evaluation 1988	N.A.	Annual monitoring

*The evaluation took place after the field study was completed.

Pakistani planning, monitoring and evaluation procedures

Pakistan has a comprehensive system for planning, monitoring and evaluating projects. It is based on a series of structured forms, the Planning Commission Forms (PC 1-5), which in principle play an important role in the administrative control of projects. In practice, however, the system has many shortcomings (see Al-Jalaly, 1991). The project planning form (PC 1) is widely used, because project budgets cannot be approved without it. Project activities cannot be significantly modified without altering the PC 1, which is a time-consuming procedure.

In theory, the project cycle operated in Pakistan ensures the efficient and integrated performance of planning, implementation, monitoring, adjustment (of project schedules or design) and evaluation of results. In practice, however, these links are not made, and little if any qualitative information is collected in the framework of the project cycle.

Monitoring is very much geared to physical progress, financial control and accountability. Information gathered at project level is sent through the provincial departments to central level (the ministry concerned and the Planning Commission). Each level is supposed to use this information for its own planning purposes and in supervising policy implementation.

Attempts have recently been made at central level (the Project Wing of the Planning Commission, and the Audit Office) to monitor at least some of the more important national projects and to evaluate their effects. The activities concerned included a training project, supported by the Netherlands, aimed at the development of performance auditing within the Audit Office. With the help of donors (among them the Netherlands and the UNDP), efforts are also being made at the provincial level to promote institutional development which will reinforce planning, monitoring and evaluation within the Planning and Development Departments.

Evaluation capacity

Pakistan has considerable evaluation capacity in the private sector, but the universities, with their emphasis on academic research, play little part in policy or project-oriented research. There has long been a considerable number of Pakistani consultancies in the civil engineering and construction sector; since the start of the 1980s a market has also gradually developed for advisory services in the social and economic field. Demand by donors has led to the establishment of growing numbers of small consultancies operating commercially. These consultancies are involved in project formulation, implementation and evaluation.

The evaluation and management capacity of government departments is not held in high regard among the consultancies. Several reasons are given, including the inadequate field experience among civil servants, their lack of knowledge and experience of monitoring and evaluation and their predominantly bureaucratic ethos. For its part, the government largely relies on reputable research agencies funded or part-funded by the state in project formulation, evaluation and monitoring, rather than on the more expensive private firms.

Donors operating in Pakistan and the government have differing views about the quality of the local consultancies. Neither their capacity nor their degree of independence is unanimously regarded as adequate. For their part, the consultancies perceive differences between the various donors as regards the use of local expertise in project formulation or evaluation. In cases where Pakistani consultancies are involved, it is, as a rule, in combination with foreign firms. This arrangement seems to work well.

In short, Pakistan possesses a considerable reservoir of local expertise able to take part in project preparation, implementation and evaluation.

5.2.2 Evaluation

Planning and design

Although the timing of evaluations was usually laid down in project documents, their preparation was left till the last moment. Sometimes this meant that the counterpart institution was unacquainted with the terms of reference or was unable to provide an expert to take part in the mission.

The country desk usually requested external consultants, the project team or the consultant responsible for project implementation to draft the terms of reference. These were usually either formulated in fairly general language (see box 5.2.1) or comprised lists of points about the operational aspects of project implementation. Matters whose significance went beyond the individual project were not given prominence.

The selection of evaluators showed signs of haste. They were usually not contacted until shortly before work was due to start. While the evaluators dispatched by DGIS were commonly experts experienced in the relevant field, in no case were they sufficiently familiar with DGIS's Evaluation Guidelines. They lacked practical as well as theoretical knowledge of evaluation methods and approaches.

Box 5.2.1   Terms of reference for the evaluation of the Matric Education Project in 1988

Terms of reference:
1) To evaluate the progress of the project to date, to advise on any modifications to the monitoring/evaluation procedures and to help in training of the evaluation coordinator, if necessary.
2) To advise on the follow-up of the recommendations laid down in the evaluation report of the first semester (phase 1) of the Women's Matric Education project, with particular reference to new project areas.
3) To help in the formulation of a job description for an advisor in the field of educational methodology in distance learning, and to assess the possibility of employing a local expert for this job.
4) To discuss the findings of the mission with the concerned staff members of the Alama Iqbal Open University in a workshop to be held at the end of the mission.
5) To submit a report at the end of the consultancy.

Preparation

The limited time allowed to the evaluators meant that preparation was seldom thorough. Often no clear methodology was proposed before the start of the evaluation, position papers were not written before the fieldwork, and sometimes it was not possible to bring together the expatriate evaluators to prepare them before their departure for Pakistan.

Two of the five evaluations studied were carried out jointly with Pakistani experts. The latter did not have the opportunity to prepare themselves sufficiently for the evaluation.

Implementation

In one of the jointly executed evaluations the Pakistani participants did play an active part, as did the Pakistani project director. In the other their role was limited.

The collection of information in Pakistan was largely based on interviews with senior project staff and with counterpart organisations directly involved with the project. Field visits were too brief to allow for more than impressions. None of the five missions made either a structured study of the project's impact on the target group or a detailed investigation of the sustainability of its activities.

One evaluation used an explicit methodology: the core of the evaluation was a workshop using the Object-Oriented Project Planning (OOPP) approach. This participative technique looks first at the identification and formulation (design) of project activities and provides a framework for judging the relevance of the original objectives. The workshop gave the evaluators and the project staff the opportunity for an intensive exchange of ideas on fundamental aspects of the project, such as its objectives, approach and execution.

Exchanging ideas while evaluation missions are at the project site can be very intensive. Project staff, who often work in professional isolation, feel a great need for contact with other professionals. Such contacts, often informal in nature and not directly reflected in evaluation reports, may be of great significance to the project. In three of the five missions there are indications that they played a more important role than would appear from the official reports and related correspondence.

Evaluations were normally limited to assessing progress in the implementation of projects and devoted relatively little attention to their effects on the target population or to their general design. Surprisingly, they were also often concerned only with the current project phase. Projects often enjoy long histories and their effects and their potential sustainability can be assessed.

Reporting and feedback

Debriefings were normally held with the project team, the counterpart organisation and the embassy, but project staff thought them largely ceremonial. Their intensive contacts with evaluators during their stay at the project site meant that the meetings were more by way of being a summary than an analysis which might bring forth new factors for consideration.

By the end of their stay in Pakistan the evaluators had generally drafted a preliminary report. These were anything from a largely complete final report to a debriefing memorandum of a few pages listing principal conclusions and recommendations. Reports were generally finalized in the Netherlands, without input from the local evaluators.

Feedback at the project level has recently improved. At the instigation of the embassy, the project team must draw up a compliance report a few months after completion of the evaluation showing how far their decision-making reflects the mission's recommendations.

Role of the Pakistani authorities

Although the Pakistani government has its own project evaluation procedures, little or no use appears to have been made of them in the projects studied. The Pakistani counterparts were only marginally interested in the evaluations. Interviews also showed that the authorities stressed the control and verification function of evaluation rather than its learning function. When an evaluation is carried out at a time of indecision over the project's future, the counterpart usually becomes more actively involved in its planning and design.

5.2.3 *Functions of evaluation*

There was often a direct link between evaluation and initiation of the decision-making process for project extension, and counterparts thought of it as a necessary precondition in DGIS's procedures. Evaluation as a form of control was seen mainly in terms of checks on the relevance to policy and on the effectiveness of project implementation. Financial control was not listed among the functions of evaluation. Project teams welcomed an evaluation because it enabled them to exchange ideas with the evaluators. Trouble-shooting, as a function of evaluation, was emphasised, particularly in those cases where problems, especially of an institutional nature, could not be solved without jeopardizing the relationship of the project team with the counterpart organisation.

The exchange of ideas as an aim of evaluation was mentioned relatively often, both by the evaluators and by expatriate project staff; trouble-shooting was cited mainly by expatriate project staff.

The functions of evaluation understood by those involved largely reflect those formally given to it by the DGIS: management support, policy support and communication. The extent to which they were realised is discussed below.

Management support

When the recommendations of the missions were compared with their follow up, it was clear that recommendations geared to aspects internal to the project were often put into effect by project management. In this sense evaluation missions fulfilled a management-support function (see box 5.2.2). There were, however, clear indications that many such recommendations were based directly on ideas that already existed within the project. In such cases the recommendations are reflections

Box 5.2.2   Role of the 1987 mission to evaluate the Pata Integrated Agricultural Development Project

Between 1986 and 1990 the Pata project was concerned mainly with the use of groundwater for irrigation. Project activities involved for the most part the construction of wells for irrigation and to a lesser extent measures to improve production methods.

The Pata project was evaluated twice during its first phase. The first evaluation, in 1987, was prompted by the fear that the components of the project proposal to do with groundwater exploitation and agricultural development were not in balance. In addition, the project team regarded the project's short-term goals as unrealistic. It was decided to have the project evaluated by independent experts to establish a more realistic timetable.

The main purpose of the evaluation was to examine activities and to decide to what extent the short-term goals were realistic. The evaluators were also expected to judge the effectiveness of the project's organisational structure and to assess the scope for using surface water for irrigation purposes.

Two Dutch agriculture and irrigation experts formed the evaluation team; there was no Pakistani participation.

The evaluators suggested deferring groundwater drilling and paying more attention to agricultural development. The Pakistani authorities were initially reluctant to accept the proposed change, but after consultation with the embassy the shift was agreed. As a result of the evaluation the plan for the first phase was significantly modified.

of ideas which have sometimes already been presented elsewhere (for example, in progress reports).

There are no clear-cut procedures for the adoption of evaluation results either by projects or by DGIS. The adoption of recommendations is largely left to the agency implementing the project. At the end of an evaluation, a sort of negotiating process seemed to take place between evaluators, the project team and sometimes the embassy and the counterpart organisation; this was intended to arrive at recommendations acceptable to everyone.

Evaluation missions generally have a two-fold task: the evaluation itself and the presentation of an outline for the next project phase. Sometimes they are even asked to formulate the new phase in detail. As the recommendations are, of course, based on the findings and conclusions of the evaluation, missions thus often become major users of their own recommendations.

Policy support

The role of evaluation in underpinning policy was less clear. Since evaluations were generally geared closely to the operational aspects of project implementation, they produced little usable information of relevance to general policy.

Communication

Only two of the evaluations studied proved to have played a major part in communications between the various parties. One was the evaluation in which the participative OOPP approach was used. This led to intensive communication between the project staff and the counterpart organisation. The other case was an evaluation of a project which was executed without expatriate staff. Since the reporting by the local contractor was inadequate, the evaluators acted as intermediaries between the embassy and the project agency.

The importance attached to the communication function seems to be greater where informal contacts between project staff and embassy were more difficult because, for example, the project site was far from the capital or no expatriate staff were employed.

5.2.4 *External monitoring*

The embassy is responsible for supervising projects in progress and systematically monitors them, using a variety of channels. Some of these are relatively unstructured and informal, while others have a clearer structure and are more formal. Two formal monitoring channels were investigated: regular visits to projects by monitoring missions and regular progress reporting.

Monitoring missions

Four projects included in the study are reviewed regularly by monitoring missions, which usually consist of experts involved for a long period in the project concerned. These experts also write commentaries on technical and project progress reports and provide advice to the country desk and the embassy on a wide range of matters.

The main formal purpose of the missions is to keep the embassy staff informed of the progress of the project and aware of any relevant problems. In practice, however, their role extends well beyond this. Their range of concerns and the depth of their

investigations varies widely. Much depends on how the various parties involved in the project (project staff, project supervisors for the donor and the counterpart organisation, country desk) operate.

Monitoring missions tend to fill the gaps left by any of these parties. They also deal with unclear demarcations of responsibility and authority between those working on the project. For their own part, they must cope with ambiguous terms of reference. In addition, a long-term relationship between the monitoring experts and the project complicates their position. They have objectively to determine progress despite often having a strong affinity with the project and seeing themselves as advisors to its staff.

Progress reporting

The embassy also uses the regular progress reports from project teams to monitor implementation. These should be submitted quarterly or half-yearly and should be structured in line with the Reporting Guidelines. Project teams comply with these requirements to a reasonable degree.

Information contained in progress reports was rarely new to the embassy staff and any problems mentioned in them were probably the subject of earlier consultations between the embassy and the project team. It was exceptional for the embassy to respond to the contents of progress reports; this may be due to the fact that they were effectively no more than administrative formalisations of earlier contacts.

The Pakistani system of project monitoring

Monitoring by the Pakistani authorities is designed mainly to check on project expenditure and on the achievement of quantified objectives; only to a very limited extent is it suited to monitoring non-quantifiable objectives and activities. Monitoring involves the completion of a form (PC 3) which is intended to give a general picture of progress. This form is normally used for the larger projects only. In it, project inputs and results are expressed in percentages. Information generated in this way does not lend itself to substantive analysis.

The Pakistani system is largely independent of the progress reports prepared by project teams for the embassy. Only occasionally are progress reports used to help complete Pakistani monitoring forms. The Pakistani counterpart institutions and the embassy show little interest in each other's monitoring systems and the two operate side by side.

5.2.5 *Internal monitoring*

During a seminar on monitoring and evaluation, organised by the embassy in 1989, it was made clear that establishing internal project management information systems was very important. The embassy has since taken further steps to promote internal monitoring, supplying project staff with information and offering assistance with the design of monitoring systems.

Partly in response to the embassy's initiative, project teams have devoted much attention to monitoring; this is particularly true of expatriate project staff. However, these efforts, while sometimes considerable, have not yet resulted in effective internal monitoring. Some systems have been set up for parts of projects only, and most of the remainder have been working only for a brief period. Their effectiveness in either case has yet to be established. That apart, the limited success achieved in establishing and operating internal monitoring systems is explained by a variety of factors, the most important of them are:
– Few project staff members have practical experience of monitoring.
– DGIS does not lay down operational guidelines for internal monitoring, so monitoring varies widely from one project to another.
– Project workloads are relatively heavy, and priority is given to doing the job. Establishing and maintaining monitoring systems is lower down the list.
– Many project activities generate unquantifiable results (for example, institution building). These are difficult to fit into a monitoring system.

Interest from counterparts in internal monitoring of a more substantive nature is limited. Their organisations are subject to the normal reporting procedures of Pakistani bureaucracy which are geared mainly to finance.

5.3   Egypt

5.3.1 *Background*

Since 1985 project and programme aid have formed parts of the Netherlands' bilateral aid programme in almost equal measure. Project aid has mainly gone to supporting national and regional activities. Some years ago it was decided to concentrate the regional projects in the governorate of Fayum. Rural development projects there are aimed at activities such as improving water management and horticultural and poultry production. In the governorate of Damietta project aid is limited to one activity in the field of health care. Support is given to national

projects involving water management, non-farm employment in rural areas, food production and health care. Most of the project aid takes the form of technical advice and projects are carried out with the help of both expatriate and local experts.

Monitoring and evaluation in Egypt

The government sector is very bureaucratic and decision-making is highly centralised. Coordination might be possible at regional level (the governorates), but the fact that most sectoral agencies at this level communicate mainly with their respective ministries in Cairo forms a major obstacle. Planning mainly involves macro-economic factors. Each ministry has its planning division, but they do little by way of planning, evaluation and monitoring at project level.

Steering committees (usual in locally financed activities as well as in donor-funded projects and programmes) are seen as a way of enabling the government to supervise interventions. They are regarded as useful forums which, among other things, simplify communication and coordination between government agencies. They also act as advisory bodies to projects and as an instrument of decision-making and control by government. Project appraisals and evaluations are often contracted out to public research institutions and universities. The government makes hardly any use of local private consultancies.

Evaluation capacity

Egypt has long been familiar with private civil engineering consultancies, but social, economic and cultural development consultants have been thinner on the ground. In the second half of the 1980s they increased in number considerably. Consultancies are usually small firms, which sprang up in response to demand from donors.

The private consultants see the monitoring and evaluation capacity of the Egyptian government, at central and regional level, as inadequate. In their view, project monitoring by the government is purely to do with finance and to provide checks on outputs. This form of monitoring is applied mainly to infrastructural work and is not aimed at estimating the effects of, or learning from, projects.

There are differences of opinion among donors about the quality and independence of the private consultants. Private consultants who have been involved in evaluations for donors are usually positive about working on an equal footing with foreign experts. A distinction is nevertheless made between the roles of Egyptian and non-Egyptian team members. Local evaluators often act as facilitators for the mission

or as resource people; their job is mainly to supply information on the sometimes complex local institutional and social context of the project to the foreign experts participating in the mission. This role is seen as crucial to the success of evaluations. The combination of foreign experts and Egyptian experts who are familiar with local conditions and have contacts in government is seen by the Egyptian consultancies as practical and effective.

The private consultants do not see universities and public research institutions as competitors because university researchers are in their opinion not equipped to carry out short-term policy-oriented investigations.

Five projects were chosen for on-the-spot study in Egypt; their main features are listed in table 5.3.

Table 5.3   Characteristics of projects studied in Egypt

Name of Project	Duration	The Netherlands' Contribution *Dfl. millions*	Evaluated in	Evaluation interval *years*	Remarks
Re-use of Drainage Water	1982–94	7.7	1988	N.A.	Monitored by Egyptian-Netherlands steering committee
Horticultural Development, Fayum	1984–93	3,4	1986, 1990	4	At times, monitoring by steering committee
Productive Families	1985–92	8.1	1987, 1991	4	Monitored by steering committee
Damietta Primary Health Care	1985–93	7.3	1988, 1990, 1992*	2	Monitored by steering committee
National Potato Cultivation	1981–94	3.8	1985, 1990	5	Monitored by steering committee

*The evaluation took place after the field study was completed.

5.3.2 *Evaluation*

Planning and design

Nearly all the evaluations were envisaged in the project documents. They were generally initiated by the country desk and, although they had often been planned as early as the formulation stage, they did not, as a rule, take place on time. The effect has been that any further project phase could not start directly from the current phase, and interim phases lasting anything from a few months to a year had to be interpolated. This gives rise to uncertainty among project and counterpart staff and renders the project less effective.

Terms of reference for evaluation missions were usually first drafted by the country desk, which consulted the Technical Advice Section (DST/TA) on their content. In all cases they were submitted to the counterpart institution which, however, did not usually respond substantively. Only where the project was affected by some crisis did the counterpart take a more active role in formulating the terms of reference.

Tasks were generally made up of elements prescribed in the Evaluation Guidelines. However, terms of reference varied widely in their detail and precision. While in the past they were sometimes seen by evaluators as too general, they have, in recent years, become more specific with the addition of non-evaluative tasks and comprehensive lists of points to be covered.

Dutch evaluators were normally experts in the relevant field but did not have any specific expertise in evaluation. They had evidently been selected for their technical background and knowledge of the country. It is possible that the country desk's or DST/TA's familiarity with particular evaluators may also have played some part.

The Egyptian evaluators were usually government officials working in public institutions which, while not directly involved in project implementation, maintained indirect links with the project or the agency coordinating it. Their expertise varied, so that their role in the evaluation process ranged from being very active to rather passive.

Preparation

Preparation time for missions was generally limited to a few days, with one exception. This was an evaluation prompted by serious problems arising during the implemention of a particular project; in this case thorough preparation was undertaken.

Dutch participants in evaluation missions generally felt that the time allotted to preparation was inadequate. The Egyptian evaluators were allowed little or no preparation time.

Implementation

Most of the evaluations were carried out jointly by international and Egyptian experts. The period spent in Egypt was usually two to three weeks and duration was relatively independent from the complexity of the project or the workload consequent on the terms of reference.

The position of the Egyptian evaluators was not entirely satisfactory since they were not always clear about what was expected of them. They usually acted as facilitators and interpreters of the often complex social and economic situations concerning the projects, although sometimes their role was more generative and substantive.

Expatriate experts generally felt that missions were too short to allow them time to familiarise themselves with the often complex social, cultural and institutional structures of the projects. Consultation with the Egyptian evaluators did not offer a complete solution.

Cooperation between project teams and evaluation missions was generally good, although work was hindered by the lack of adequate monitoring data. Because embassy staff did not involve themselves in the field phase, the evaluators had maximum freedom to reach their own conclusions.

The counterpart institutions often saw the evaluators as negotiators. While this partly reflects the ethos of the Egyptian bureaucracy, it was also a consequence of the fact that for the Netherlands the main question they had to address was whether projects should be extended, and if so, how. As the evaluators had no formal mandate to enter negotiations on the future of projects, they operated in an ambiguous situation in which they had to shape their own roles as intermediaries.

The effectiveness of project implementation was the principal subject of evaluation and no thorough analysis of the efficiency of the use of project resources was undertaken. Evaluations were not retrospective to any great extent and they were usually limited to the activities of current, or recently completed, project phases. They were often geared solely to the operational aspects of a project, taking its design, objectives and underlying assumptions simply as given. That the original formulation was, in many cases, flawed played a part. Confusion about project goals,

activities and resources and a lack of operational indicators made the evaluator's job particularly difficult. The effects and impact of the project on the target group were not systematically analysed in those evaluations which were examined.

Reporting and feedback

The missions generally ended with meetings with the project team, the embassy and the counterpart institution. Most debriefing meetings were felt to be satisfactory by those concerned. It was not clear, however, whether this feeling related to the result of the mission's recommendations (in many cases that the project should be continued) or reflected an agreement with the analysis presented by the evaluators.

A provisional report was compiled by the evaluation missions while they were still in Egypt. They usually gave a brief description of the mission's activities and a summary of its main conclusions and recommendations. This document provided the basis for the discussions held at the final meetings. Evaluation reports were generally finalised in the Netherlands, with no active input from the Egyptian evaluators.

In the light of the evaluation, the embassy and the country desk consulted each other on the substantive aspects of the report, the recommendations and more generally matters relating to the evaluation exercise; the recommendations for the subsequent project phase lay at the heart of these consultations. The findings were not subjected to further analysis or used for purposes beyond the individual project, so that the lessons learnt remained limited to the project concerned.

5.3.3 *Functions of evaluation*

There were marked differences of opinion between the expatriate and Egyptian evaluators with regard to the purpose of evaluation. The former emphasised the lessons to be learned by those involved in the project, and the latter saw evaluation as a series of checks on progress. This attitude reflects the Egyptian bureaucratic ethos which is geared largely to inspection and control and leaves little room for critical examination of the original project objectives, which are generally taken as given.

For the Egyptian counterpart institutions the main purpose of evaluation was to provide the opportunity for reflection and negotiation on possible project extension. This was reinforced by a combination of the tasks of evaluation and formulation in the same mission. Evaluations initiated by the DGIS were sometimes called

'mid-term reviews' when the Egyptian authorities were told that they were scheduled. This strengthened their impression that projects were to be extended whatever the outcome. As the counterparts saw mission findings as major factors in negotiations around the next project phase, missions were often seen as the 'Netherlands" party in the process. This clearly made the role of the Egyptian evaluators difficult. Missions had no formal mandate to conduct negotiations, and meeting counterpart expectations demanded great tact, especially from the mission leader.

Management support

Project teams usually accepted mission recommendations or gave reasons for not doing so. It was unclear, however, whether the adoption of recommendations reflected their technical merits or whether the project teams were too dependent on the evaluation findings and recommendations to be able to distance themselves from them. Extensions are, of course, also in the interests of the teams. Some of the evaluations studied had a considerable impact on the broad direction taken in the subsequent project phase.

Policy support

The evaluations of the five projects selected were not geared to policy development at a level higher than the individual project, whether in terms of policy aspects of relevance to the Egyptian counterparts or of more general aspects of development cooperation between the Netherlands and Egypt. The experience gained from the evaluations was not translated, at least not a in transparent fashion, from the level of project implementation to that of policy formulation on a wide scale.

Communication

Some evaluations fulfilled a clear communication function, notably in those projects where cooperation among the various parties was less than satisfactory. In such situations the intensive consultations which resulted from the presence of the mission were of great assistance.

5.3.4 *External monitoring*

Steering committees

Steering committees were involved in four of the five projects. The committees generally comprised the Egyptian project manager, the expatriate team leader, rep-

resentatives of the counterpart institutions associated with the project and a senior representative of the counterpart ministry. As a rule the ministry representative functioned as the 'general project director', was senior to the Egyptian project manager and chaired the steering committee. The embassy was represented by an observer on one of the steering committees associated with the projects.

Committee functions varied from one project to another, but usually they played an important role in the management of projects and sometimes in day-to-day decision-making (see box 5.3.1).

Box 5.3.1 Role of the Drainage Advisory Panel in the Re-use of Drainage Water Project

The Drainage Advisory Panel (DAP) is a forum for cooperation between Egypt and the Netherlands. The Panel, which initially comprised 10–12 Egyptian and Dutch experts, served as an independent advisor to the Egyptian Public Authority for Drainage Projects. Later it became the main counterpart for the Drainage Research Institute established in 1977. The DAP gradually acquired more operational functions (studies, consultancies, training, etc.). In 1983, at the instigation of DGIS, the Panel returned to its original advisory role, which had been primarily that of advisory body, and focused *inter alia* on the monitoring of the Re-use of Drainage Water Project and its technical supervision.

The DAP met once or twice a year. Long agendas limited the time it could devote to the Re-use of Drainage Water Project. Reports of this technically complex project were often not submitted by the project team until the last moment. The situation became complicated as ever more of the project activities were carried out in the Netherlands (the development of the drainage model). Tensions developed between the Egyptian and Dutch project staff, and scepticism grew on the Egyptian side as to the usefulness of the project results. In 1988 the country desk decided to evaluate the project. The results of this evaluation helped prompt the establishment of a separate steering committee to supervise the technical aspects of the project and to check progress. The committee comprised an Egyptian and a Dutch expert (both members of the evaluation mission), the Dutch and Egyptian project directors and a senior civil servant within the counterpart ministry, by whom it was chaired. The effect was to push the DAP into the background.

The steering committee had close contacts with the project: in 1989 there were four meetings, each lasting three days or more. These contacts meant that the committee had a strong influence on the project, on cooperation within it and on the links with the counterpart institutions concerned. Regular contacts were also maintained with the embassy.

No more meetings of the steering committee took place after the original project termination date (September 1989), even though the project was extended. Since then, the embassy has had a more direct monitoring role. The DAP, too, continues to monitor the project at arm's length.

The steering committees, which are a feature of project implementation in Egypt, can be regarded as a purely Egyptian phenomenon reflecting the highly centralised organisation of their bureaucracy. An active committee can facilitate the supervisory role of the embassy and can be regarded as a major instrument of external monitoring. They could make another important contribution by providing a channel for linkage and interaction between the project and the counterpart institution responsible for formulating and implementing policies which go beyond the individual project. This could make it relatively easy for the experience gained within projects to be used in policy formulation. It was impossible to establish whether the committees do in fact act in this way.

Reporting

Project teams kept the embassy abreast of their activities by means of regular progress reports. The reports were then forwarded, sometimes with comments, to the country desk. This form of reporting seems to have filled an administrative function rather than meeting current management needs. Problems affecting project implementation with which the embassy could help, either by offering advice or by suggesting solutions, were communicated to it verbally or in writing. Interventions by the embassy or the country desk went ahead regardless of whether regular progress reports had been published.

Monitoring missions

In three of the five projects, regular missions formed part of the embassy's monitoring practice; the work of the missions included advising the project team on substantive matters.

5.3.5 *Internal monitoring*

Most projects studied did not have well-developed internal monitoring systems. Only one featured a monitoring system which recorded not only activities and outputs but also certain effects on the target group (see box 5.3.2). Elsewhere, structured monitoring was mainly concerned with the use of inputs and with financial aspects of implementation.

Box 5.3.2   The internal monitoring system of the Productive Families Project (PFP)

The PFP has two main activities, the supply of small-scale credit to the poor and the provision of technical training in non-agricultural income-generating activities. It is carried out by an NGO, the Productive Families Association.

The project has developed from social activities to a more economic role. This shift was accompanied by the definition of monitoring indicators and of data to be gathered by internal monitoring. The project has several monitoring instruments:

- progress reports, compiled at village level and consolidated at higher levels;
- impact studies, carried out by a consultant;
- field visits and meetings, with senior staff regularly visiting villages and having frequent informal contact with both field workers and the target group.

Credit is the main focus of monitoring and the data gathered by it is the core of progress reports to higher levels.

A separate steering committee monitored progress and policy development. An unusual aspect of the project is that staff work at several hierarchical levels. All of them are familiar with the workings of the project organisation at various levels, and this facilitates feedback to the relevant decision-makers and planners. The use of several monitoring and reporting channels is a safeguard against misunderstandings in communication and helps to ensure effective internal checks on data flows.

The combination of an internal monitoring system that works well and a management unit skilled in policy formulation, makes effective response to changing circumstances possible. That the project was clearly formulated is also important: the objectives are clear to everyone and the collection of data on progress is almost automatic.

Chapter 6   Discussion

This chapter discusses the findings of the desk and field studies in relation to the various steps in the evaluation process. It considers whether the functions assigned to evaluation by the DGIS are realised and includes a brief account of the incorporation of DAC principles into DGIS evaluation practice. As there has been no desk study, the discussion of monitoring is based solely on data from the field investigation.

6.1   The evaluation process

Evaluation can be divided into several steps: planning, design, preparation, execution and feedback.

6.1.1 *Planning and design*

Evaluation proper is preceded by provisional planning. Whether and at what stage a project can expect to be evaluated is normally decided when it is formulated or appraised. Whether evaluation should take place at the time proposed can be decided as the project proceeds. There was no evidence that country desks planned for evaluations which might occur over a span of several years.

The first step in evaluation is its design. Its purpose, the questions to be addressed and the framework in which it is to be carried out must all be settled. This step also includes the selection of the evaluators.

Design is primarily the responsibility of the country desks, which normally consult the counterpart (through the embassy). For the more technical aspects of evaluation design they can enlist the help of the Technical Advice Section and of external advisors. The country desks influence evaluations in two major ways: by setting the terms of reference and by choosing the (Dutch) evaluators.

In practice, terms of reference are often drawn up by third parties, who may be advisors involved with the project, the embassy or, quite often, the project team itself. The staff of country desks made clear during interviews that they do not see themselves as sufficiently well informed about projects to be able to draft satisfactory terms of reference, nor do they have the time.

Terms of reference were found, both by the field investigations and from the analysis of the evaluation reports, either to be formulated in very general terms or simply to be a heterogeneous 'shopping-list' in which elements of advice, formulation and problem-solving regularly played a larger part than evaluation. Since evaluations are often oriented towards the future (the formulation and appraisal of proposals for subsequent project phases), they are considerably influenced by the question of how the project should be continued. This is at the expense of analysing the experiences of the project and identifying the lessons to be learnt from it.

Terms of reference tend mainly to focus on project implementation. Central issues in the policy of the DGIS like sustainability and the effects of the projects on target groups frequently receive only cursory coverage. Neither does the DGIS give direction about the research methods to be used. That decision is left entirely to the evaluators themselves, who are thought to be the experts in these matters. With some exceptions, evaluators had not been given the Evaluation Guidelines.

Analysis showed that the more comprehensive the terms of reference, the better the eventual report and that the haphazard way in which terms of reference are put together is one explanation for the mediocrity of the reports.

The choice of evaluators is based mainly on their technical qualifications and their familiarity with the region concerned. Relatively little importance is attached to experience of evaluation, and familiarity with evaluation methodologies or DGIS evaluation procedures plays no demonstrable part at all. This runs counter to the demand that evaluators should adopt an appropriate method of investigation; they cannot do this without some intervention or substantive guidance from the client.

Responsibility for bilateral projects rests jointly with DGIS and the recipient country; evaluation is thus also a joint responsibility. But authorities in the recipient countries see an evaluation as a necessary element in the Netherlands' procedures for adjusting or extending projects. Local authorities have a very limited interest in the design of evaluation studies and their involvement is usually confined to accepting terms of reference and appointing local evaluators. The exceptions are

to be found in projects where problems have arisen which threaten either their extension or even the aid relationship with the Netherlands.

6.1.2 *Preparation*

Preparation involves collecting information on the project concerned and choosing the approach to be used. These tasks are primarily the responsibility of the evaluators.

Dutch evaluators are normally given two to three days in which to prepare the mission before leaving the Netherlands. In this time they must visit the DGIS to discuss the terms of reference and to get the project documentation. Sometimes they must take part in discussions in the Dutch office of the consultant involved in implementation and undertake some logistical preparation. Only in exceptional cases is any thorough study of the relevant documents, called for by the Evaluation Guidelines, possible. Nor is a position paper on the project to be evaluated written. There is little if any discussion of the methodology to be used either between the client and the evaluators or among the evaluators themselves.

Local evaluators, however, have even less time: for them the evaluation often starts the moment their Dutch colleagues arrive. The counterpart institutions also seem to attach little importance to thorough preparation.

Country desks think that this brief preparation is adequate, and most evaluators seem happy to accept the situation. The field study found only a few instances of more intensive preparation, organised at the instigation of the evaluators themselves.

Little real information about the projects and their results so far is available to the evaluators in advance. This, together with deficiencies in project planning and design makes preparation even more difficult. The use of techniques which give the project plan a transparent structure (for example, the logical framework approach) is the exception rather than the rule. Clarity in project design and planning is essential both for monitoring and evaluation. Documents provide insufficient information about the initial situation facing projects and are muddled in their contents: long and short-term objectives are confused, are often set out as activities, and the assumptions underlying the choice and design of projects are seldom made clear.

6.1.3 *Execution*

Evaluation is the responsibility of evaluators, both Dutch and local. It is exceptional for DGIS or the embassy staff to take an active part.

Five major factors affect evaluation: the methods used, the availability of information, the working relationship between the evaluators and the project staff, the role of counterpart institutions and of local evaluators.

Evaluation methods

Few evaluation reports gave any indication of the mission's approach or of its method of collecting data and those which did demonstrated the weakness of these methods. Thus lacking a sound methodological foundation, the findings are largely impressionistic, particularly when they try to describe and assess the effectiveness and impact of the projects. These subjective expert opinions, substituting for an investigation, make it difficult to verify the findings and make the reports less convincing.

Availability of project information

A lack of project information is one of the greatest obstacles to effective evaluation. Even though internal monitoring arrangements are not well developed, information could be got from project teams or from outsiders and combined with the written studies produced by the project team. But during preparation, time is too short for a thorough study of any information that might be obtainable.

For the most part, evaluators get the project information they need from interviews with project staff, civil servants and the staff of the counterpart institution. Field visits were an almost universal feature of the evaluation missions covered by this study. These, often brief, visits offer few opportunities for the structured collection of primary data, discussions with the target group or visits to different project locations. This means that no more than a broad impression of a project's effects on its intended beneficiaries can be gained. It also means heavy reliance on the views of the official authorities involved. In exceptional cases, preliminary studies among the target groups by experts not involved in the project were conducted before the evaluation.

Relationship between evaluators and project staff

Perhaps the most important factor governing the conduct and the direct effects of an evaluation is the relationship developed between the evaluators and the project staff. Although each is unique in reflecting personal views and preferences, certain common elements can nevertheless be seen.

Both project managers and evaluators are usually experts interested in the project's technicalities; as a result managers, and other project staff, often see evaluations as an opportunity for an exchange of views with experienced colleagues rather than for examining the project and the context in which it operates. 'Professionally isolated' teams working in remote areas are especially likely to feel the need for these exchanges.

Evaluators need to develop a good working relationship with the project team if they are to advise on a project's future. Because a mission which spends only a brief time in a project area cannot build up a complete picture of its implementation and of its social, economic and institutional context, evaluators are partly dependent on project staff, both for information and opinion. Project staff usually have clear views about the strong and weak points of their project and its future lines of development. In reaching their conclusions and recommendations, evaluators are naturally influenced by the views current within a project. While this may undermine their independence, it enhances the likelihood of recommendations being accepted and hence the effectiveness of the evaluation for project management.

Role of the counterpart

With some exceptions, when they examine projects with a view to possible adjustment or extension, local authorities have little interest in playing an active part in evaluation. As a result, and given the Netherlands' clear desire for it, the counterpart institutions see evaluation as a requirement imposed by DGIS's procedures for project adjustment or extension and have little interest in possible lessons for their own policy development. Counterpart institutions are generally well disposed towards evaluation, which they see as an opportunity for negotiating the future of the project. This, in turn, reinforces the tendency for evaluators to concentrate on the future.

Role of local evaluators

Local evaluators usually play a limited role in joint missions. This is partly because they are not attached to them until a late stage and so have no opportunity to

prepare. A heavy workload and shortage of time make it very difficult to compensate for any consequent shortfall in knowledge. It is rare for them to play a part in the compilation of the final report and debriefing at DGIS, both of which take place in the Netherlands.

Local evaluators may chiefly be facilitators and supporters, by no means unimportant roles. They provide access to certain groups of respondents and act as interpreters of local social and cultural phenomena.

6.1.4 *Effects*

The effects of evaluations are achieved by the process of informing interested parties (debriefing, reporting). Findings are used at two levels: the project (management and the institutions involved in supervision) and policy-making (priority-setting within DGIS and by the counterpart institution).

Feedback of results normally starts towards the end of a mission. This is to ensure that everyone is properly informed of the mission's findings. At this stage the evaluators are, of course, the chief actors, but the counterpart and, particularly, the embassy are also important. After the initial findings, conclusions and recommendations, the report is used in reaching decisions about the future course of the project. This is the job of the country desk and, through the embassy, of the counterpart institution.

Process of feedback

Feedback takes the form both of official debriefings and of informal consultations between evaluators and the interested parties. A preliminary version of the evaluation report and, frequently, a summary of the main conclusions and recommendations, provide discussion material for the debriefings.

The final days of a mission are very hectic, involving intensive consultations between mission members, project managers, embassy staff and the counterpart institution. These are often negotiations and during them the local evaluators move into the background, their place is commonly taken by the counterpart institution directly concerned. By contrast the Dutch evaluators, and particularly the mission leader, have an increasingly important role. This consultation usually comes to provisional conclusions, which are frequently presented at a final debriefing with senior officials of the coordinating ministry. Where such meetings are held at a senior level of

government they are somewhat ceremonial in nature. Evaluators' conclusions and recommendations naturally awaken expectations regarding the future of the project concerned and provide the basis for further consultations between embassy and counterpart institution.

Reporting, normally completed in the Netherlands, concludes the evaluation. Little time is allowed for compiling reports and many of them show signs of having been written in a hurry. They are often difficult to read, poorly structured and largely inaccessible to outsiders. This helps to explain why note is seldom taken of the results of evaluations outside the circle of those directly involved.

Uses of evaluation results

Evaluation results play a major part in decisions about the future of the projects. For projects to make use of them communication between evaluators and project staff must be good, the results must reflect the ideas of project staff, the evaluators must consider the future, and results should be geared towards project implementation. The use made of them within DGIS is decided partly by the scope the evaluation reports offer to underpin and legitimise decision-making. Appraisal memoranda for subsequent project phases usually refer *en passant* and in very general terms to evaluation results.

There are no procedures guaranteeing the use of evaluation results. In the case of a current project phase, for example, it is not clear how projects should respond to those recommendations formulated by evaluators and endorsed by the Netherlands and the recipient country. The embassy in Pakistan has instituted a follow-up procedure: some time after the issue of an evaluation report the project team is asked to submit a report showing how far they have complied with the mission's recommendations. Where projects are extended, it is customary for these recommendations to be included in the project document for the next phase; as that phase is implemented the project team must take account of them.

It is difficult to know exactly what effect evaluation results, which go beyond the individual project to thematic, sectoral and country policy, actually have. Neither long-term policy plans nor annual plans make any explicit mention of them. Interviews with the DGIS staff make clear that evaluation findings are rarely used in adjusting or in formulating policy. Changes are initiated by adjustments in the Netherlands' overall development policy. In any case the operational focus of project evaluations means that few lessons with a wider application can be drawn. Shortcomings in the presentation of results also limit the scope for wider-ranging effects.

6.2   Realisation of the formal functions of evaluation

The Evaluation Guidelines identify three functions with regard to evaluation: management support, policy support and communication. They are defined more fully in Chapter 1, but there are others. Each party involved in evaluation and monitoring sees these functions differently. These differing perceptions are presented in the following sections.

6.2.1   *The actors' view*

Those involved are usually aware of all the functions of evaluation and monitoring, but interviews and conversations during the field visits made it clear that each party ascribed functions to evaluation that reflected their own background and position. The specific circumstances of a project also played a part. Five functions can be distinguished, some of which coincide with those defined in the Evaluation Guidelines.

1) Evaluation is often seen as advisory and the evaluators' main role as getting the project team and counterpart institution to look at fundamental issues relating to the project; this is management support.
2) Some people stressed the ability of an evaluation mission to take up issues or problems which the project team cannot easily tackle without jeopardising the relationship with the counterpart institution or the embassy; this is communication.
3) Evaluation is frequently described as a learning process. Analysis of a project's design, implementation and results by independent experts is regarded as an essential contribution to the sum of knowledge about it. Evaluation should lead everyone thoroughly to re-examine the original project design and the manner in which implementation has been supervised. Conclusions from this should contribute to more effective and efficient design and implementation. This is a learning function, which is a combination of management and policy support. It relates first to the project but can, in principle, be extended to similar activities elsewhere or to policy development.
4) Nearly all respondents point to the part that evaluation plays in making decisions about extending the project. This is an administrative function which, with its legitimating effect, forms part of management; it is neither recognised as such in the Evaluation Guidelines nor referred to explicitly in the other procedures of the project cycle.
5) Evaluation is also seen as having an audit and control function. This is not

mentioned in the Guidelines, but nearly all interviewees involved in project implementation or working in counterpart institutions refer to it and regard it as perfectly legitimate; the evaluators, the DGIS and embassy staffs give it less importance.

6.2.2 *Formally-assigned functions*

This section considers to what extent evaluations comply with the functions mentioned in the Evaluation Guidelines.

Management support

While the field investigation identified major differences between evaluations, in many cases they clearly support project management. However, the ways in which evaluation affects a project's future are often not so obvious and depend on the quality of communications between evaluators and project staff. The conclusions and recommendations of an evaluation are more likely to be accepted where there are good working relationships and mutual respect. Sometimes the recommendations formulated by evaluators actually emerged from the project team. Where project staff can see their own contribution to a mission's findings and conclusions, they find it easier to accept the results. This is true even when the tone of the report is critical. Evaluation tends to concentrate on project implementation, and the adoption of recommendations to do with operational matters is largely a matter for the project management.

Since most missions are also expected to advise on subsequent project phases and, indeed, sometimes to formulate them, they are the primary users of their own output. An evaluation report also provides the country desk with an 'objective' contribution to the appraisal of a future phase for the project, together, in some cases, with a plan for it drawn up by the project team themselves. In those cases where the various parties are in broad agreement about the future phase before an evaluation takes place, it mainly serves to legitimise decisions.

Policy support

The Evaluation Guidelines mention policy support as a function of evaluation. In practice little happens. This is partly because most evaluations deal with project implementation, and partly because the experience gained with and the lessons from evaluation have not, in the past, been given systematic consideration. There

are no decision-making or consultation procedures which require evaluation results to be considered, either in the implementation and extension of current projects or in the identification, formulation and appraisal of new ones. Nor have any mechanisms been established enabling experience in project implementation to be used in general policy development or in the adjustment of country and region policy plans.

The narrowness and superficiality of project evaluation are rooted in its design and preparation and are reflected in the terms of reference. The evaluators' background and experience reinforce a restricted focus. Most of them are experts who have built up their knowledge and experience in project implementation; it is far less common for them to have any experience of policy evaluation and policy development beyond the level of the individual project. Their interest is chiefly in technicalities; policy principles are generally taken as given and not questioned.

Since many evaluations concentrate on the achievement of direct and concrete results, questions about sustainability and effects on the target group often get pushed into the background. Analyses of Netherlands' aid policy or of policies to do with the general development of the recipient country play only a secondary role. Efficiency in project implementation also tends to be neglected.

There are however exceptions to this. Interviews with country desk staff showed, for instance, that wide-ranging and thorough evaluations of long-term programmes, or groups of similar projects, are sometimes carried out. In these studies explicit provision is made for analysing policy and for examining the immediate and further effects on the target group and the sustainability of the activities in question. This approach has implications for the duration and the consequent cost of evaluation and for the intensity of the preparation and fieldwork involved.

Communication

An evaluation mission is a major event for project staff. They see the substantive assessment, which it inevitably involves, as more significant than their administrative contacts with the embassy and the counterpart institution. The findings may put project extension at risk, and everyone will, therefore, take steps to secure and promote their own interests. Consequently intensive consultations usually take place before, during and after the mission. The exchange of ideas begins when the terms of reference are compiled and the evaluators are selected; position papers and, perhaps, proposals for a subsequent project phase may be drawn up before the evaluation takes place. The mission's findings are the subject of continuing

consultations at the debriefing stage; the responses of DGIS and the recipient country to the evaluation report are the subject of the final round of consultation. Evaluation missions can act as catalysts in stimulating debate on fundamental matters relating to a project. Sometimes a mission has to work in a situation in which the donor, the project staff and the counterpart institution hold different views on the project's future. Then it can find itself involved in negotiations in which objectivity may be hard to reconcile with political and strategic pressures. It is then that the communication function of evaluation is most important.

Missions sometimes set out quite deliberately to promote or to restore communications when they have broken down between the parties involved in a project affected by poor working relationships.

6.3 Application of the DAC evaluation principles

The principles and purposes of evaluation formulated by the DAC were summarised in Chapter 2. A major principle is that evaluation should form an integral part of policy-making and of planning and implementation. In developing their evaluation instruments and in planning evaluations donors should take account of the information needed by different levels of the organisation: getting the result from an evaluation to both management and policy-makers is vital.

Objectivity in evaluation depends partly on clear objectives for policy, programmes and projects. In the design and implementation of evaluations their purpose and scope should be set out and the users of the results should be identified. A description of the methods and criteria by which aid activities are judged is essential. So, too, is the availability of sufficient time and resources. The DAC emphasises the need to plan and carry out evaluations jointly, or in coordination with, the developing countries.

The desk and field study findings discussed in the previous sections show that the practice of bilateral project evaluation departs markedly from the DAC principles.

6.4 Monitoring

6.4.1 *External monitoring*

External monitoring takes place in the context of the supervision of project implementation. Supervision is delegated to the aid sections of embassies, which have

various instruments for this purpose at their disposal. These are regular working contacts between project and embassy staff, regular progress reporting by projects, monitoring missions and, in certain cases, project steering committees.

Progress reporting

Project teams submit regular progress reports. Their frequency, usually quarterly or half-yearly, depends on the projects' circumstances. The reports rarely contain up-to-date management information which is normally given by telephone, by letter or during mutual visits. They are really a formal record of known facts and have an administrative function. Where day-to-day contact is more difficult, progress reports may contain more up-to-date management information. Embassy staff were generally satisfied with them.

As a rule the embassy sends the country desk copies of the reports, sometimes accompanied by comments, information on action taken or suggestions for future measures. This is often done in status reports on the aid programme prepared for policy consultations and mid-term reviews and submitted by the embassy to the country desk.

Monitoring missions

Regular monitoring by external experts is an important tool for embassies in supervising projects. It is usually undertaken by research institutions or private consultancies contracted for the purpose. This form of monitoring is used widely in Pakistan and, to a lesser extent, in Burkina Faso and in Egypt.

The absence of a procedural framework for this form of monitoring makes for a considerable measure of freedom in the structuring the role, function and content of monitoring missions and in working out the operational details. Annual missions to the projects concerned has been the norm, but visits are also made at shorter intervals. Whether external experts are employed in project supervision depends largely on the problems affecting a project, its size and complexity, the embassy's capacity to exercise adequate supervision and the country desk's need for an additional information channel. The views and attitudes of the embassy and country desk regarding project supervision play an important role.

Monitoring missions do not always have a clear remit and their members have often played a part in identifying or formulating the project concerned in the first place. They often perform tasks neglected by other parties involved in a project.

There is evidence from the field studies that missions fulfil a variety of functions *vis à vis* the projects: some emphasise technical support for the project team, while others are more concerned with verification. A few missions also collected information for policy development beyond the individual project. Many of them act as troubleshooters, solving problems and mediating between institutions directly or indirectly involved in project implementation.

This wide range of functions blurs the dividing line between implementation and supervision. In principle it is the job of monitoring missions to check progress against the project's schedule; their findings then provide a basis for recommendations for the subsequent period. These become binding once they are endorsed by the embassy and by the counterpart institution (and, frequently, the country desk). Both the embassy and the missions themselves attached considerable importance to the advisory function of monitoring; project teams may be less appreciative.

Steering committees

In Egypt the field study found that, along with progress reporting and direct contacts with project teams, the embassy gave a major supervisory role to steering committees. These are familiar and widely-used instruments in Egypt which are not limited to donor-funded development activities. They are usually made up of representatives from the project, from the counterpart institution and, occasionally, from the donor. The functions of steering committees and the scope of their activities vary widely: some concentrate on progress and thus become monitors, while others also intervene in day-to-day project management and so form a sort of second tier of management. Counterpart institutions play a major part in project steering committees in Egypt.

The role of the counterpart

The policy statement *A World of Difference* (DGIS, 1991) says that a guiding principle must be the integration of project monitoring into the recipient country's control structures. This implies, at the very least, a joint role in monitoring for the embassy and for the counterpart. The study found that, in practice, this almost never happened.

Only in Egypt do the embassy and counterpart institutions cooperate in external monitoring through the steering committees, although regular supervision by the embassy also takes place in parallel with the work of the committees.

Progress reports from the project teams are geared solely to the embassy's supervision of the project. Counterpart institutions may also get the reports, but make no use of them in their own monitoring procedures, which are mainly for checking expenditure.

In each of the three countries covered by the field investigation, the country desk and the embassy have organised monitoring exclusively to meet the needs of the donor. This is either because the counterpart institution's procedures do not meet DGIS reporting requirements or because its monitoring system is not working. In practice monitoring missions are exclusively DGIS tools and normally consist only of expatriate experts.

Moves are being made under the programmes of cooperation with both Pakistan and Burkina Faso to establish or strengthen local monitoring – so far with no structural results. In Burkina Faso support is given to the planning and monitoring unit in the Ministry of Water which will supervise and evaluate all the water supply projects funded by the Netherlands. In Pakistan the evaluation and monitoring capacity of the provincial planning departments of Baluchistan and the North West Frontier Province are being strengthened. These activities are coordinated, wherever possible, with a UNDP project for improving monitoring and evaluation at provincial level elsewhere in the country. Long-term cooperation with the Pakistani Audit Office in the area of performance auditing is also taking place.

6.4.2 *Internal monitoring*

Internal monitoring serves, in the first instance, to give project managers relevant and up-to-date information on progress; an effective internal system is closely geared to implementation and, to a lesser extent, to factors in the project's environment. A second function is to generate information for evaluation purposes. Establishing and maintaining an internal monitoring system is the responsibility of the project agency or contractor.

Arrangements for internal monitoring varied widely among those projects studied. A few had elaborate systems while others paid little attention to monitoring. Most of them, however, had partial monitoring systems, usually related to one or more specific activities. Despite the often considerable efforts devoted to setting up internal monitoring arrangements and the encouragement occasionally given by the country desk or the embassy, effective monitoring systems existed only in a minority of the projects.

There are a number of reasons for this. First, insufficient attention was devoted to monitoring during the formulation stage. Also, imprecisely and inaccurately described objectives, assumptions, activities and expectations make it difficult to design internal monitoring systems. Another difficulty is that neither project agencies nor contractors know much about designing and establishing them.

DGIS does not provide implementing institutions with clear guidelines for establishing internal monitoring systems. Neither has monitoring in practice yet given rise to informal norms and standards. Internal monitoring, which seems to be a relatively new concept within DGIS, is not yet seen as a precondition for effective management and is, therefore, not regarded as a vital part of any project. Pioneering ventures such as a workshop on forms of monitoring organised by the Netherlands' embassy in Pakistan bear the stamp of individual initiative. Without wider acceptance and a procedural framework, it is doubtful if such ventures will prove sustainable or be followed up.

Chapter 7   Conclusions and Recommendations

7.1   Conclusions

Project evaluation has been operationalised only to a limited extent. The Evaluation Guidelines recognise only one type of evaluation, which must satisfy several different requirements. Except for the Guidelines, which combine procedural and some substantive elements, no additional support is given to the design, planning and execution of evaluations.

The length (2–3 weeks), focus (project implementation) and methodology (analysis of documents, interviews with project staff and officials of relevant institutions) of evaluations are all becoming increasingly standardised. Within the country and regional programmes, evaluation has become the assessment of projects by experienced experts familiar with the relevant technical field. In other words, it is closer to a form of peer review than to objective, verifiable and soundly-based research.

The use of project evaluation by the DGIS leaves to be desired: insufficient care is given to formulating terms of reference and the design and preparation of evaluations show signs of haste. Procedures that would help to shape the process are not consistently applied. In the management of bilateral aid projects an 'evaluation culture' has grown up. It is marked by frequent evaluations, a large measure of discretion for the staff and evaluators in their design and conduct, a dominant position for the Dutch members of the missions, superficiality and a failure to learn lessons which have a relevance that goes beyond the individual project.

7.1.1 *Evaluation does not adequately fulfil all its allotted functions*

Evaluation fulfils different functions for each party concerned. Each of them, therefore, imposes specific demands which do not, as a rule, coincide with the qualitative

requirements for evaluation laid down in this study. The limited accessibility of evaluation reports, for example, is not a problem if it contains arguments of use to those directly concerned. Also, the vaguely defined evaluators' role and the limited breadth and depth of their terms of reference has advantages. The result is that an evaluation offers something to everyone. Describing an evaluation as a 'lucky dip' in which every actor gets something useful may be an exaggeration, but only a slight one.

Evaluations administer and legitimise DGIS decisions (country desk and embassy) about possible modifications or extensions of projects.

Counterpart institutions in developing countries see evaluation as a necessary element in DGIS's procedures for project modification or extension.

Project agencies and contractors see evaluation mainly as the assessment of project execution and as an instrument of verification for the donor. Evaluation also affords them the opportunity to engage with independent experts in technical discussions of the project. They, too, see it as an instrument for decision-making about project extension.

Expatriate evaluators emphasise the learning function of evaluation, while evaluators from developing countries stress its control function.

Evaluation clearly supports project implementation, mainly because of its operational focus and its orientation towards the future. It also plays an important part in communications between the responsible counterpart institution, those engaged in project implementation, the embassy and DGIS.

There is little if any feedback from project implementation to policy formulation; indeed, there is evidence that projects are seen as isolated activities. Since the procedures for formulating policy and making decisions about new projects do not require the experience gained with comparable projects in the country, region or sector concerned to be considered.

Even though policy support is mentioned in the Evaluation Guidelines as one of the functions of project evaluation, it receives little attention in practice. Evaluations are occasionally carried out which are specifically aimed at drawing conclusions for policy support purposes, but the experience of most project evaluations plays little part in policy formulation. Their contribution to the quality and effectiveness of aid in a broad sense is, therefore, limited.

The focus and content of evaluation are determined by the interests of the various stakeholders. None of them regard checking the extent to which underlying policies are being realised as a primary purpose of project evaluation. Consequently, the design, preparation and implementation of evaluations show insufficient concern with policy matters. The importance which Dutch aid policy gives to issues such as gender, target-group orientation, environmental protection and sustainability is only partially reflected in project evaluations.

Various institutional and organisational factors influence the feedback process within DGIS. There is no repository or 'institutional memory' for the findings, recommendations and lessons set out in evaluation reports. The reports themselves are seldom read outside the circle of those immediately involved, no summaries are produced, and the reports are not classified and stored in a way which would make them easily accessible.

The place of embassies and country desks in project management emphasises administrative procedures. A corporate ethos has grown up in which priority is given to administrative control rather than optimising the use of development-related knowledge and experience possessed by many members of staff.

Furthermore, the build-up of this knowledge and experience is hampered by the regular rotation of staff and heavy workloads. The deployment of sector specialists at embassies and the reinforcement of the technical expertise of DGIS's Spearhead Programmes Coordination and Technical Advice Department have not ensured the effective use of project evaluation.

7.1.2 *Evaluations do not satisfy methodological requirements*

Little is done in evaluations to check if projects are in line with Netherlands' development policy or the policy objectives of recipient countries. The attention devoted to efficiency and effectiveness also varies widely. Thorough assessments of project management structures are frequently lacking and the effect of projects on target groups is rarely subjected to proper analysis. There is often no assessment of sustainability.

The absence of a coherent methodology, the lack of decent statistical material, of indicators and of standards against which performance can be measured means that assessments rest predominantly on the impressions of the evaluators. As a result, the information contained in evaluation reports can be both shallow and narrow.

The methodological deficiencies of evaluation are caused by various interrelated factors. The country desks, embassies and counterpart institutions seem unable to put the necessary care and effort into planning, organising, preparing and implementing evaluations. The selection of expatriate evaluators is based mainly on their technical expertise and familiarity with the region concerned – knowledge and experience of evaluation are not thought to be very important.

Shortage of time means that evaluations are rarely thoroughly prepared and carried out. Effective evaluation is made difficult by defective project formulation in which the analysis of the problem is generally inadequate, objectives are often not clearly specified, and baseline information is lacking. Planning techniques to enhance the transparency and quality of project design are infrequently used. As little has been done to develop monitoring in projects, there is a lack of basic data on achievements and their effects on target groups. Since field visits provide no opportunity for the structured collection of primary data, evaluators obtain the information they need largely through interviews with project staff and government officials which may result in an 'authority bias'.

7.1.3 *The involvement of counterpart institutions and evaluators from the developing countries is limited*

The Netherlands' development policy on cooperation with recipient countries in the planning, design, implementation and use of evaluation has, in practice, had only a modest impact. The DGIS usually initiates evaluations, so counterpart institutions, because they see them chiefly as elements in DGIS's procedures for modifying or extending projects, often have little interest in it. Evaluators are regarded as negotiators acting for the donor. The fact that counterpart institutions do not play an active part in evaluation also reflects their limited financial, technical and personnel resources.

With occasional exceptions the role of local evaluators is subordinate to that of their expatriate colleagues. They do not take part in the design and preparation of the evaluation or in final reporting and debriefing at the DGIS, since these activities normally take place in the Netherlands. Local evaluators tend mainly to act as facilitators.

7.1.4 *The role of monitoring is unclear*

The lack of success in internal monitoring is explained by a number of factors.

– There is as yet no general policy governing the desirability, purpose and design of monitoring systems within bilateral development projects. In some cases (for example, Pakistan), however, significant steps have been taken to encourage the establishment of such systems.

– There is no substantive DGIS guidance on the design and operation of internal project monitoring systems. The Evaluation Guidelines refer to monitoring but do not give it content, and the Reporting Guidelines are concerned solely with monitoring as a supervisory tool for embassies. Project agencies and contractors thus have no solid basis for developing internal monitoring systems. In practice there is also little or no exchange of ideas based on experience of the kind built up, for example, in Pakistan.

– Country desks and embassies do not see project agencies and contractors as responsible for establishing and operating effective internal monitoring systems; where they exist they are thought to be special rather than matter of course.

– Very few people approached in this study had any ideas about or knowledge of the purpose, design and functioning of monitoring systems. Those who were interested in internal monitoring gave their own account of the job which it could or should do; this was true both of country desk and embassy staff and of project personnel. Only in one or two cases was the commonly encountered interest in monitoring translated into practice. But without broader acceptance and a procedural framework, it is doubtful if systems developed in the relative isolation of individual projects will be followed up or prove to be sustainable.

Project documents usually pay insufficient attention to the place of monitoring in management or to the determination of suitable indicators. Broadly defined objectives and the lack of specified assumptions, activities and intended results hamper the design and working of monitoring.

External monitoring, by regular progress reporting, contacts between project teams and embassies and, sometimes, monitoring missions of external experts are effective. The picture that emerges from the study is one of a strongly donor-centred activity in which the role of counterpart institutions is small. Where joint external monitoring arrangements exist, a parallel DGIS-oriented system of monitoring remains in

operation. The embassies' role in the supervision of project implementation seems not to vary with differing circumstances: in-house and contracted-out projects are both subject to the same reporting requirements and monitoring missions are used in both cases.

The procedures for the project cycle provide neither an unambiguous definition of supervision nor operational guidelines. The practice of supervision goes beyond what is suggested in the relevant procedures, which is that project execution should be checked against the agreements embodied in the commitment decision. The embassies' role in projects is generally more extensive, encompassing a substantive advisory and supervisory function which seems to be effective. This effectiveness, in part, reflects the interest and expertise of individual embassy staff members and the working agreements between the embassies and the country desks.

The absence of a proper policy for monitoring means that those responsible have a large measure of discretion in the design and detail of external monitoring arrangements. Whether external experts are deployed depends largely on the problems encountered within the project concerned, the embassy's capacity to undertake supervisory functions and the need of the embassy or country desk for an additional channel of information.

In principle monitoring missions should measure the project's actual progress related to the initial planning schedule. In practice they do the jobs neglected by other parties.

It was found that evaluation and monitoring missions fulfil virtually identical functions in communication and management support. During interviews with the DGIS staff the point was often made that the quality of information generated by monitoring missions which visit a project routinely is superior to that produced by non-routine evaluations. In practice, monitoring missions combine *ex post* checking for the embassy with an advisory function for the project, thus blurring the dividing line between project supervision and implementation.

However, when it comes to scrutiny, the two types of mission differ in breadth and depth. Monitoring missions are parts of the regular routine of project supervision and are geared to the practical aspects of implementation; evaluation is a non-routine activity and differs from monitoring by virtue, for example, of the assumed independence and objectivity of the evaluators. According to DGIS staff, this explains why evaluation has an important role in deciding whether project extensions are justified.

To sum up, the main findings of this study point to deficiencies in monitoring and evaluation which limit their effectiveness. Before considering the causes of these deficiencies and possible remedies, two points need to be made.

Studies carried out by other donors (CIDA, 1991, Scanteam, 1993, Finnida, 1991, USAID, 1992) show that these deficiencies are not unique to DGIS. Despite the considerable experience that some other donors have in evaluation, their more elaborate procedural and substantive framework for monitoring and evaluation, and their repertoire of instruments and methods, they, too, fail to fulfil the relevant requirements.

In the light of this study and comparable exercises carried out for other donors, one may ask whether too much is expected of monitoring and evaluation. Neither of these two instruments delivers all that it should in terms of policy support and in the contribution made to the quality of development aid. Can monitoring and evaluations carried out at project level reasonably be expected to make such a contribution? If so, evaluation must then be set in a proper policy framework and have the necessary methodological scope and analytical depth.

With hindsight it may be said that the DGIS's current limited procedural and substantive framework cannot support the high expectations from the results of project evaluation. Nor can the inadequate knowledge of evaluation among all concerned and the manner in which the instrument is used. In practice, project evaluation, with its operational focus and superficial nature, can scarcely be expected to make a substantial contribution to policy support and quality assurance.

These deficiencies have a variety of causes that involve institutional and procedural issues. Institutional issues include the frequent changes of staff in the country desks and the embassies, and shifting project portfolios which impede the balanced build-up of knowledge and experience in the field of development cooperation. The lack of clarity regarding the operational responsibilities of the country desks and embassies on the one hand and the diffusion of responsibilities between the donor and recipient countries for 'project ownership' on the other is a procedural issue. These factors are linked. The right kind of framework, institutional as well as procedural and substantive, is a first requirement if the quality of monitoring and evaluation is to be improved and these instruments are to be used effectively.

7.2   Recommendations

Recommendations solely to do with procedural and substantive matters can bring only limited benefits when they are not put into a wider institutional perspective. Although this study did not include an analysis of the overall institutional framework of development cooperation, several issues concerning the future shape of monitoring and evaluation must be addressed before more concrete recommendations are presented.

The central issue must be the functions to be assigned to these instruments. Thinking on aid management and accountability has changed since the start of the 1980s. Little is said in the Evaluation Guidelines of 1982 about project evaluation in relation to contemporary questions of control and accountability. This must be allowed for in the debate, and control and accountability must take their place alongside the functions already assigned to evaluation, both operationally and in policy support.

Monitoring and evaluation contribute little to quality assurance. Given the results of this study and the similar experiences of other donors, debate about what can realistically be expected is clearly needed. Ambitious expectations must be related to the priority accorded to these instruments and the resources devoted to them. Attention needs to be focused on the proper integration of monitoring and evaluation into the project cycle and hence, also, on the nature of project preparation. Priority needs to be given to the generation of information at project level. The debate will also need to focus on the demarcation of the two instruments' functions and on their relative importance and on their interlinking.

Another important issue is the role of recipient countries. Current policy gives a full and equal role to counterpart institutions and local evaluators, but the coincidence of donor and recipient interests that this presupposes proves, in practice, to be very difficult to realise. The limited interest of recipient countries in evaluation in its current form helps to explain the limited role played by local evaluators.

The following specific recommendations emanating from the study include:
1) ways of improving the quality of monitoring and evaluation and their application;
2) conditions for a more productive place for these instruments in the organisation of DGIS;
3) ways of enhancing participation by counterpart institutions and local experts in monitoring and evaluation.

7.2.1 *Towards improved monitoring and evaluation*

Greater diversity in the forms of evaluation would enable the need for information at both operational and policy level to be met more effectively. There are two basic forms, each meeting specific information needs in policy formulation and implementation.

The first, project review, is focused primarily on operational aspects of the project but may also take account of policy aspects. It checks on whether implementation is going according to plan both quantitatively and qualitatively. Project review has also to underpin decisions about the extension or modification of a project in the light of experience built up during its implementation. This is a legitimising function and reviews must, therefore, be carried out by independent experts.

The second, project evaluation, draws lessons from experience to underpin policy. It assesses project implementation and contributes to policy-making at project level and above. It should, therefore, focus mostly on discovering a project's impact on its intended beneficiaries and on its sustainability. Such evaluations should deal with policy, have a sound methodological foundation and should, at the preparation stage, use the available information on the project concerned. Given its costly and time-consuming nature, project evaluation must be employed selectively. An evaluation may cover several projects simultaneously and should also be used as an instrument to draw lessons from completed projects.

The precise meaning of the term monitoring, its proper function in project management, and the difference between monitoring and evaluation should be established.

It is considered worthwhile to employ the concepts 'internal' and 'external monitoring', as used in this study. External monitoring is part of the routine supervision of project implementation by embassies. Depending on its type and the specific circumstances of a project and on the capacity of the embassy to supervise it, external monitoring can be carried out by external experts. Internal monitoring should take place within the project and is the responsibility of project management. The configuration of internal monitoring should be decided at the time the project is formulated. Care should be taken to relate monitoring activities to the scale and complexity of the project and the type of information needed.

To improve monitoring and evaluation, more elaborate project formulation and design is needed. More appropriate planning techniques would clarify important assumptions, factors crucial to a project's success, links between different project

activities, the hierarchy of objectives, and standards against which project results can be judged. These techniques could include, among other things, the logical framework approach, cost/benefit analysis, environmental impact assessments and also gender-specific, organisational and institutional analyses.

Care should be taken to make a clear distinction between monitoring and the two forms of evaluation at project level. Although there is a certain functional overlap between external monitoring and project review, both focus on the operation of projects. Project review is essentially non-routine, is independent and assesses project design and implementation and the original project objectives.

The limited availability of project data has been a major hindrance to effective monitoring and evaluation. Generating baseline data during project formulation, studying effects during implementation, and planning field studies before evaluations would all improve the situation.

7.2.2 *An organisational framework for monitoring and evaluation*

An adequate procedural and substantive framework supporting those concerned with monitoring and evaluation should be established.

Procedural guidelines should be elaborated for:
1) evaluation, covering the various kinds and their specific applications;
2) the supervision of project implementation, including the function of external monitoring, possible links with monitoring systems in the developing countries and the conditions to be met when the use of external expertise is under consideration;
3) internal monitoring, and particularly the ways in which the DGIS can promote and supervise the establishment of internal monitoring systems.

These guidelines should show when and how projects may or must be monitored and evaluated. They should make clear what forms of monitoring and evaluation could be used in which circumstances and suggest how the quality of each kind of evaluation is to be established. Finally, they should contain procedures to be applied for feeding back monitoring and evaluation findings to the decision-making management levels in DGIS.

Care should be taken that the guidelines are sufficiently comprehensive for general application while leaving room to adjust monitoring and evaluation to the

diversity of projects and the varying circumstances and ways in which they are implemented.

The guidelines should provide support for all those responsible for the design and implementation of evaluations and of external and internal monitoring. They should list and describe the possible methods and means of data collection. They should also include requirements for reporting and the criteria to be used to assess completed evaluations.

Guidelines for external monitoring should set out the function of the monitors when it comes to project supervision and their general role within the project framework. They, too, should provide a summary of the methods and means of data collection.

The purpose of guidelines for internal monitoring is primarily to support those responsible for project formulation, appraisal and implementation, by making clear what forms monitoring may take.

An effective use of monitoring and evaluation presupposes familiarity with their possibilities and limitations among those involved in the design and application of these instruments. Where this is lacking, tailor made training programmes should be developed.

Evaluation and monitoring need to be set in an appropriate institutional and organisational framework. Responsibility for them should rest with those sections of DGIS which are the main users of the information they generate. These units must be able to shape the planning and execution of evaluation and monitoring and must be given the necessary support for this purpose.

As the embassies are responsible for supervising project implementation, it follows that they are also responsible for external monitoring and its organisation. They must be properly equipped for these functions.

Since project reviews are concerned mainly with the operational aspects of implementation, it is recommended that responsibility for them be delegated to the embassies. This would have the additional benefit that counterpart institutions and local evaluators could play a greater part than hitherto in the preparation of reviews. Local evaluators could then also take part in final reporting. The immediate feedback of results to those making decisions about project implementation involving both embassies and counterpart institutions, presupposes a further delegation of decision-making powers to the embassies.

The DGIS's operational units should be responsible for commissioning evaluations. Since they are also the major users of the results, this would enhance the effectiveness of the feedback from these studies to policy-makers. However, the evaluation expertise of these units would then need to be enhanced and measures taken to ensure that they have the necessary organisational capacity.

Various forms of evaluation could be included in a rolling evaluation plan. This plan should be flexible and extend over several years. It would provide information about planned evaluations and could be used as a tool to set priorities. It would offer a means of coordinating evaluation within DGIS and of coordinating evaluations with other donors and recipient countries.

7.2.3 *Enhancing counterpart participation in evaluation and monitoring*

The joint involvement of donor and recipient in the design and conduct of monitoring and evaluation follows from their joint responsibility for bilateral development cooperation. Joint responsibility for monitoring and evaluation is a principle of policy that should be retained and the involvement of developing countries should be promoted wherever possible. However, this could be done in more flexible and creative ways than hitherto.

Greater efforts to strengthen the monitoring and evaluation capacity of counterpart institutions are needed. This could include supporting or establishing monitoring and evaluation units within government departments or private institutions. Where a developing country has established its own capacity, it should be exploited to the full.

Local evaluation capacity in the non-governmental sector should be used more intensively, by, for example, involving local research institutions in studies of the effects of projects. The scope for doing this varies from one country to another.

The features of local monitoring systems should be known before internal project monitoring systems are designed. Where possible monitoring systems set up within projects should be integrated with local systems to increase their sustainability.

While joint planning, design and conduct of evaluations should continue to be a major principle of policy, circumstances may arise in which one side or the other needs an evaluation for its own purposes. In that case it must also be possible for them to be instigated unilaterally, either by DGIS or by the recipient country.

Finally, there is room for greater flexibility in structuring joint evaluations. A proper division of tasks is desirable to prevent an overlap of responsibilities between expatriate and local evaluators. The composition of joint teams should be guided by the principle that members' contributions should complement one another.

Annexe 1.   The Operations Review Unit (IOV)

The Operations Review Unit, better known by its Dutch acronym IOV or Inspectie Ontwikkelingssamenwerking te Velde, was established in 1977. IOV is responsible for conducting evaluations of Dutch aid policy. Internal evaluations of projects are the responsibility of the operational units, i.e. the country or programme desks.

IOV is part of the Directorate General for International Cooperation (DGIS) of the Ministry of Foreign Affairs. It is a completely independent unit which directly reports to the Minister of Development Cooperation through the Director General. The Minister submits the reports to Parliament; they are discussed with the Permanent Committee on Foreign Affairs with respect to follow-up actions.

Initially, IOV placed the emphasis in its work on individual project evaluations. From 1977 up to the mid-1980s the reports were primarily intended for departmental management purposes. The status of these reports was confidential. During this period about 250 evaluation reports were produced. Gradually a need developed for more general evaluations based on project findings. In the mid-1980s a number of sector reports were prepared, such as those on drinking water, animal husbandry, women in agriculture and rural development, and primary health care.

Since then, emphasis has shifted from individual project evaluations to comprehensive thematic studies; they focus on policies and modalities of implementation and cover sectors, themes or programmes. They contain a review of relevant literature, and compare results with those of other donors concerning the same subject matter.

On average, the duration of these thematic evaluations is one to two years. The studies are carried out under the responsibility of IOV, with outside experts participating in various phases of the research. Field studies are undertaken by teams of independent external consultants. Increasingly, local institutions or experts are invited to participate in these field missions.

The synthesis report, based on the various field and desk studies, is written by IOV staff and published under its responsibility. Three to four such studies are published annually. Examples of recent evaluation studies published by IOV are: import support, export transactions relevant to development, food aid, the sector programme for rural development, project evaluation and monitoring in Netherlands bilateral aid, cooperation in higher education, environment and development cooperation, the Netherlands development programme with Tanzania 1970–92, the Netherlands development programme with Mali 1975–92, the Netherlands development programme with India 1980–92, and humanitarian aid to Somalia.

A reference group consisting of external experts and DGIS staff is appointed for every study. The reference group has three functions: to advise on methodology and approach, to counsel on relevant development theories, and to give feedback on evaluation results.

Annexe 2.   Evaluation Study Work Plan

---

1   Background to the study

Improving the quality and effectiveness of aid is a central theme of the Netherlands' policy of development cooperation. Learning from the past plays an important role and much of the necessary information is obtained by evaluation. Evaluations take two forms:
- those carried out by the Operations Review Unit (IOV) which focus on specific themes, sectors and programmes;
- those which are carried out under the responsibility of the operational units of the Directorate-General for International Cooperation (DGIS). These evaluations are limited to specific development activities (mainly projects) and are governed by specific guidelines (Evaluation Guidelines).

The quality and effectiveness of evaluations carried out by the operational units have been debated for some time. It was the lack of adequate information on how these evaluations worked that prompted this study.

2   Research questions

For the evaluation four principal questions were formulated:
1) What forms of evaluation can be distinguished in practice, and to what extent do they conform to the Evaluation Guidelines?
2) To what extent and in what way are the purposes of evaluation set out in the Guidelines achieved? What factors determine this? Are any other functions assigned to evaluation?
3) How far does current practice meet the wishes, expectations and capacities of the various parties involved (country desks, embassies, evaluators both in the Netherlands and in developing countries, counterpart organisations, executive agencies)?

4) How far does current practice within DGIS reflect the recommendations of the
    OECD Development Assistance Committee (DAC) and the views on evaluation
    current in the international donor community?

3   Design of the study

The study is concerned solely with the evaluations carried out by the operational
units of DGIS in the context of the bilateral aid programme. It covers projects con-
tracted out to consultants and those directly managed by DGIS. Project monitoring
is dealt with only in relation to evaluation.

The study analyzes the perceptions and opinions of those involved in evaluation,
including DGIS staff in The Hague and at embassies in developing countries,
evaluators, project staff and the personnel of counterpart organisations. Evaluation
and monitoring are studied with the Evaluation Guidelines in mind.

The investigation involved a combination of desk studies and field work. First, 180
evaluation reports from the period 1987–90 were appraised using criteria such as
coverage, clarity and completeness. This was followed by an assessment of the
quality of the contents of a random sample of 37 reports.

Field studies were carried out in 16 development projects in Burkina Faso, Egypt and
Pakistan. The aim of these studies was to obtain a detailed picture of the preparation
and implementation of evaluations and the use of evaluation results. The role of
monitoring was also studied. The three countries were chosen for their varying
capacities for local evaluation, for the differences in their supervisory structures
for project implementation and for the number of evaluations carried out in recent
years. The choice of projects was based on the quality of the evaluations, on the
supervisory structure of project implementation, and on the existence of internal
monitoring arrangements. Projects were chosen which would encompass the full
range of evaluation studies and monitoring systems. The desk and field studies were
supplemented with material obtained in interviews with experts covering general
aspects of evaluation and monitoring.

The study sought neither to assess the projects themselves nor to duplicate their
evaluation. It makes no comment on the accuracy of the information generated by
project evaluation, but focuses instead on the coverage and depth of the information
and on how it is used.

The study took place in the period from mid-1991 until mid-1993.

4   Implementation of the evaluation

4.1 *Desk research*

The desk study involving the analysis of 180 evaluation reports was carried out by Ms O.T.M. van der Kemp (Nedworc) and Ms A. Brederode (Consultants for Culture and Management) supported by Ms F.M.C. Groenewege (Nedworc). This study resulted in a working document 'Kenmerken van evaluatierapporten met betrekking tot bilaterale hulpprojecten 1987–1990' (Characteristics of evaluation reports of bilateral projects 1987–1990).

The assessment of the quality of the contents of a random sample of 37 reports. was carried out by W. Flikkema. This study resulted in the working document 'De kwaliteit van evaluatierapporten' (The quality of evaluation reports).

4.2 *Field studies*

The field studies were conducted by teams consisting of a Dutch consultant and a local consultant. The visits by the coordinating team of the IOV to the three countries had a dual purpose:
– to discuss monitoring and evaluation practices with government officials, donors, and consultants engaged in monitoring and evaluation of development activities;
– to discuss with the evaluation missions engaged in the field study the application of the research questions to individual projects.

The study in Burkina Faso was carried out by W. Stolz (private consultant) and Dr I. Alitou of AI Consultant in Ouagadougou. Fieldwork took place during November–December 1991 and involved the following projects:
– Sensibilisation et Formation des Paysans autour des Barrages
– Hydraulique Villageoise dans le Boucle du Mouhon
– Programme d'aménagement des zônes d'habitat spontané de la ville de Ouagadougou (Projet Urbanisme de Ouagadougou)
– Vallée du Kou
– Programmation et Exécution du Developpement Intégré Kaya (Pedi Kaya).

The study resulted in the working document 'Evaluatie en monitoring: deelstudie Burkina Faso' (Evaluation and monitoring: sub-study Burkina Faso).

The study in Pakistan was carried out by Dr R.A. van de Putte of the Management for Development Foundation and Dr J. Tirmizi of Social and Economic Enhancement Research in Lahore.   Fieldwork took place during January–February 1992 and involved the following projects:
- Animal Husbandry In-Service Training Institute
- Development and Utilization of Human Resources Project
- Pak-Holland Metal Project
- Quetta Sewerage and Sanitation Project
- Matric Education Project
- Pata Integrated Agricultural Development Project.

The study resulted in the working document 'The monitoring and evaluation of Dutch aided projects in Pakistan'.

The study in Egypt was carried out by B. Evers of the Development Research Institute, Tilburg University and Prof. S. Youssef of the Department of Public Administration, American University in Cairo. Fieldwork took place during February–March 1992 and involved the following projects:
- Horticultural Development Project Fayum
- Re-Use of Drainage Water Project
- Productive Families Project
- Damietta Primary Health Care Project
- National Potato Cultivation Project.

The study resulted in the working document 'Monitoring and evaluation in Dutch–Egyptian bilateral cooperation projects'.

4.3   *Coordination*

The study was coordinated by Ted Kliest of IOV, Willem Koot of Consultants for Culture and Management and Bert van de Putte of the Management for Development Foundation. The coordinators were responsible for designing the evaluation study, drafting the terms of reference, the supervision of the desk research and field studies, and for writing the final report.

The study was supported by an advisory group, consisting of:

Ms H.I. von Metzsch	–	DGIS/IOV (Chairperson)
Prof. H. Thomas	–	Institute of Social Studies, The Hague
Dr T. Dietz	–	Department of Human Geography, University of Amsterdam
R. van de Geer	–	DGIS/Southern African Countries Section
P.J.Th. Marres	–	DGIS/West African Countries Section
A.H. Pieper	–	DGIS/IOV
K. Broersma	–	DGIS/Technical Advice Section.

Comments by the Minister for Development Cooperation, sent to Parliament on 9 September 1993

Introduction

Attention has been focused on the quality of project evaluation, as undertaken by the operational units of the Directorate-General for International Cooperation (DGIS), in the Explanatory Memoranda accompanying the Development Cooperation Budget and in the policy statement *A World of Difference*. In its report on bilateral cooperation with Indonesia, the Court of Audit called for improvements in the planning, frequency and quality of evaluation, while the Lower House of Parliament has stressed the importance of evaluation and the need for effective feedback so as to optimise future development activities. In view of the Court of Audit report, the House asked for information on ways of making structural improvements in project evaluation. In its response the Government drew attention among other things to the intensification of the role of the Operations Review Unit (IOV) in coordinating and assuring the quality of project evaluation, to the more rigorous requirements set for evaluation studies and to ways of strengthening the role of developing countries in evaluation.

Since 1991 an IOV officer has had the job of assisting the operational units of the DGIS in the design and conduct of project and programme evaluations. The lack of a clear overall picture of current monitoring and evaluation was one of the factors necessitating a study of these instruments in the processes of policy formulation and implementation.

Research approach

The study focused on the role and function of decentrally organised evaluations in the process of policy preparation and implementation within DGIS. It also examined the different forms of project monitoring in relation to evaluation.

The evaluation and monitoring of Dutch bilateral aid projects were examined in the light of formal policy regarding these instruments, as laid down in the Evaluation and Reporting Guidelines. The views of the parties involved in evaluation and monitoring were also analysed.

Four principal questions were addressed:
1) What forms of evaluation can be distinguished in practice, and how far do they conform to the Evaluation Guidelines?
2) How, and how far, are the functions of evaluation distinguished in the Guidelines realised, and are any other functions assigned to evaluation in practice?
3) How far does current practice meet the wishes, expectations and capacities of the various parties (country desks, embassies, evaluators both in the Netherlands and in developing countries, counterpart institutions, agencies carrying out projects)?
4) How far does current practice within DGIS reflect the starting points for evaluation of the OECD's DAC and the views current in the international donor community?

The investigation combined a desk study of the quality of evaluation reports with field studies and interviews:
– the desk study assessed the quality and content of 180 evaluation reports in the light of the requirements laid down in the Evaluation Guidelines and of the principles adopted by the DAC;
– field studies were carried out in 16 projects in Burkina Faso, Egypt and Pakistan to obtain a clear picture of the preparation and conduct of evaluations and information feedback and of the role of monitoring in project management;
– the desk and field studies were supplemented by interviews with DGIS staff, embassy staff, officials of counterpart institutions, donor offices and local advisory bodies in all three countries.

The study sought neither to reach judgements on the projects themselves nor to duplicate their evaluation. No comment, therefore, is made on the accuracy of the information generated by project evaluation, rather the breadth and depth of this information and the uses made of it are the issues addressed.

The study was supervised by a committee composed of external experts and members of the DGIS staff.

The functions of project evaluation

The study's main finding was that there were shortcomings in the conduct of evaluation and monitoring which limited their effectiveness. These shortcomings were not unique to the DGIS: evaluations conducted by other donors – USAID, CIDA, NORAD, EU – had also been found not to meet the relevant requirements.

Project evaluation did not adequately perform all the functions assigned to it by official policy. This form of evaluation has a strongly operational emphasis and is oriented towards the future; it thus helps to underpin project implementation.

In the ten years that have passed since the Evaluation Guidelines were issued, the emphasis in project evaluation has come to lie on improving implementation and on possible adjustments to the evaluated projects themselves. I believe that this is indeed its main function; its role in underpinning general policy derives from it. Evaluation reports are read and used by – naturally – the country desks, the embassies and the sectoral units, notably the Technical Advice Unit; they provide the basis on which decisions to modify or extend projects or (and this does frequently happen) to terminate them are reached. The aggregate results of individual evaluations are reflected in annual plans and in country and regional policy plans. The sectoral units consider evaluation reports from a sectoral viewpoint and the lessons learnt find their way into sectoral memoranda and policy plans and into manuals of practice. Some feedback does therefore occur, but not on a systematic basis. This means that general adjustments to policy for a country, sector or theme are indeed based partly, but not wholly, on evaluations of separate project interventions.

It is questionable whether individual project evaluations are in all cases an appropriate source of information for feedback to policy level. Leaving aside policy decisions at project level, a collection of project evaluations will always be needed if conclusions are to be drawn for a higher (macro or sectoral) level; the micro mid-term review, something which is undertaken frequently, would appear to be a less appropriate input into policy formation than sectoral and thematic evaluations.

As I see it, policy support is efficiently provided by the work of the IOV itself. In recent years the Unit has carried out sectoral and thematic investigations, making use where possible of evaluation reports. Since IOV reports are intensively studied by the staff of DGIS, this constitutes another form of feedback. In addition, greater emphasis is now placed on research aimed directly at underpinning policy, notably in the form of baseline surveys; this is partly the result of the introduction, in June

1992, of the Development Test, which requires a good description of the position prior to the start of the project. Examples of such studies include profiles of the environmental situation, poverty and the position of women in a country, region or district. Such profiles can then be used to derive indicators for use in the monitoring and evaluation of interventions.

As stated in the recently issued paper, *Research and Development Cooperation*, the aims of general research policy also include the identification, by means of the thorough analysis of processes in developing countries, of the conditions under which interventions are successful.

The IOV observes that feedback from evaluation is impeded by a number of institutional and organisational factors:
– there is no repository or 'institutional memory' for the findings, recommendations and lessons set out in evaluation reports;
– the role which the embassies and country desks play in project management puts a strong emphasis on administrative procedures, with the result that full use is not made of the development-related knowledge and experience possessed by many staff members;
– the build-up of knowledge about projects and development cooperation in a wider sense is hampered by the regular rotation of staff and heavy pressure of work.

I endorse these observations. There is no central structure which processes evaluation findings; decentralised arrangements do exist, as noted above, but a more systematic approach is indeed desirable. When it comes to project evaluation, the IOV itself has a supervisory function, but this is of recent date; indeed, this study is an early exercise of that function.

The observation that project management strongly emphasises administrative procedures is also justified. On this point my expectation is that once the more rigorous procedures have become sufficiently established, capacity will be released which will make it possible once again to focus greater attention on substantive matters.

The complaint that the rapid rotation of staff and the associated frequent redistribution of project portfolios lead to a loss of knowledge both of individual projects and of development cooperation in a broad sense has also come to me from other quarters. While this observation may be correct, it should be borne in mind that the integration of home and embassy-based foreign ministry staff has also brought benefits. Measures have in any event been taken to stem the loss of knowledge, for example, through the expansion of the sectoral specialists programme and

of the Spearhead Programmes Coordination and Technical Advice Department (DST).

Methodology of evaluation

The IOV study found that evaluations carried out at project level did not meet the relevant methodological requirements. In the absence of a proper methodological framework the information generated on project implementation and results was not properly underpinned, tending rather to be rooted in evaluators' impressions. This limited the utility of project evaluations.

According to the IOV report the methodological deficiencies are due to the following factors:
– the country desks, embassies and counterpart institutions are unable to put the necessary care and effort into planning, organising, preparing and implementing project evaluations;
– the choice of Dutch evaluators is based mainly on their technical expertise and familiarity with the region concerned, but knowledge and experience of evaluation methods are not seen as very important;
– the time set aside for evaluations is very limited;
– effective evaluation is made difficult by deficiencies in project formulation;
– evaluation is hampered by a shortage of information on outputs achieved and effects on the target group;
– the dominant position of DGIS in planning and organisation and in the utilisation of results, and that of the Dutch mission members in the conduct of evaluations, means that the emphasis in Dutch development policy on participation by counterpart institutions and evaluators from developing countries in the planning, design, implementation and use of evaluation has, in practice, had only a modest impact.

These findings, relating to project evaluations carried out over the past decade, have prompted a number of measures aimed at remedying shortcomings. I realise, however, that many changes have only been put into effect since 1990, so that their impact may not yet have been visible in the period covered by the IOV study. Mention has already been made of the expansion of the Spearhead Programmes Coordination and Technical Advice Department and the sector specialists programme. Internal advisors and the Bilateral Projects Committee now place considerable emphasis on measurement and measurable indicators; the new procedural document on implementation (no. 1.6, 11 May 1992) also requires that this theme be

stressed in implementation schedules. It was partly to help bring out the impact of projects at the identification and appraisal stages that the Development Test was introduced, in June 1992; this requires that the expected effects on poverty, environment and women in development be indicated and that expectations of sustainability and of feasibility be formulated. Further improvements are however needed (see below).

So far as the limited role of counterpart institutions is concerned, I would make the following point. It is Netherlands' policy to maximise recipient-country input into evaluations, both for the sake of efficiency and in order to promote the recipient's sense of ownership. This applies both to the official counterpart organisation (the government) and to local expertise. I regard an increased local input as a development objective in its own right and am prepared to accept some loss of efficiency – should this prove unavoidable – as a price worth paying.

The IOV observation that the mobilisation of counterpart institutions has been less effective than was hoped is an indication of the intractable nature of the problem. Local institutional capacity is often limited and local expertise still scarce, certainly in the area of evaluation; in addition, potentially suitable candidates are sometimes too closely involved and are therefore disqualified. As a rule debriefing takes place in the developing country (in accordance with the Guidelines), thus ensuring that an input from local experts is at least possible. On the matter of local input there are wide differences between countries which can reflect the level of development; in general it must be remembered that African countries in particular are going through a difficult period in the administrative field.

Monitoring

The IOV report distinguishes between monitoring which is internal to projects and monitoring by outsiders. It notes that initiatives have been taken in the area of internal monitoring but that they tend not to be sustainable and have yet to have any systematic and general impact. External monitoring by means of regular progress reporting, contacts between project teams and embassies and where necessary the use of monitoring missions is an effective tool of project supervision in the broad sense. Monitoring is however very much a donor-oriented activity and is not generally rooted in the administrative structures of the developing countries.

I agree with the IOV that all projects should have a clear internal monitoring system in keeping with their complexity. Monitoring is of course not an end in itself but

a means of identifying any adjustments that may be needed. The surveillance of project implementation is the job of the embassies, and the appointment of sectoral specialists was intended to equip them with the necessary technical capacity in this area. Monitoring capacity is to be strengthened in the context of the planned further delegation of powers to the field.

Conclusions

Feedback to policy level needs to be made systematic and the roles of the embassies and country desks in the respective functions of monitoring and evaluation must be clarified. It is also time the Evaluation Guidelines were revised: here I plan to follow the IOV recommendation that project review be distinguished from evaluation proper (a change reflecting what is anyway current practice), with the former focusing on implementation and the latter on its effects. Evaluators must also be able to reexamine project objectives. Strong emphasis must be placed on adherence to the revised procedures for implementation (Procedure 1.6) and project termination (1.7), under which the country desks must make a *ex post* assessment of all terminated projects on the basis of project completion reports drawn up by the embassies. Embassies will also have greater scope for investigating whether and to what extent projects are being implemented in accordance with the agreed design.

The Spearhead Programmes Coordination and Technical Advice Department will have a greater role in the assessment of terms of reference for evaluation missions, the selection of evaluators and the assessment of the quality of the work done. I am also considering the establishment of a systematic database of experts aimed at improving the selection of evaluators. With regard to the basic data (variables and indicators) needed to make evaluation (both mid-term and *ex post*) possible, improvements can be expected from the new procedure (1.6.1) requiring that implementation schedules include descriptions of the initial situation and the expected final situation. I also expect improvements in this area as a side-effect of the introduction of the Development Test.

It would however be inefficient for all individual donors to gather their own basic data, and I take the view that such data collection is properly the job of the recipient countries. To assist them in this task a project has been launched together with the World Bank and UNDP aimed at the coordinated collection of data on poverty.

With regard to the rotation of country-section staff steps already taken ensure that staff on scale 12 or above remain in post for an average of four years; this position

will be maintained. In addition, project review and evaluation is to be included in the internal training programme.
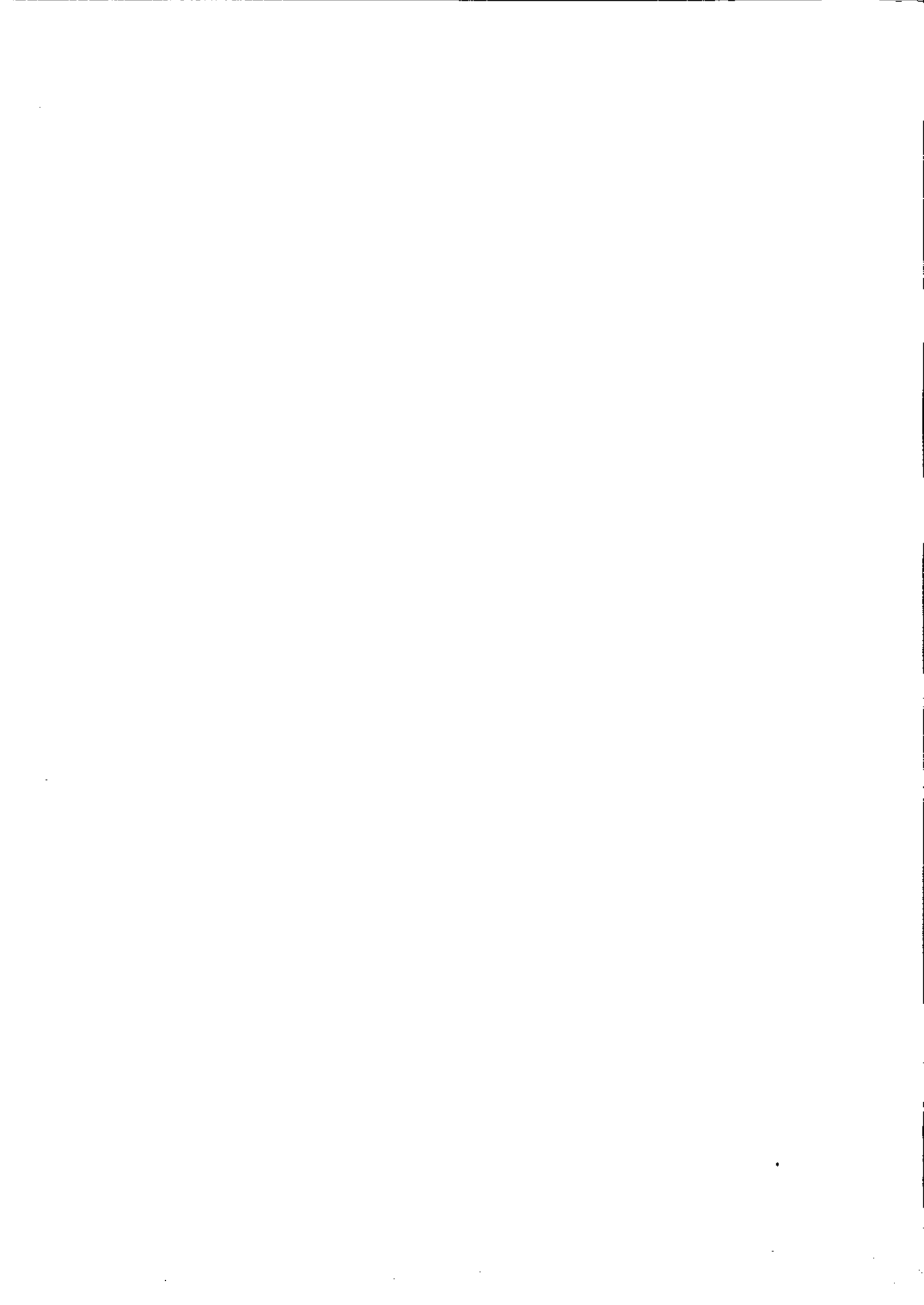
The attachment to the Spearhead Programmes Coordination and Technical Advice Department of non-transferable development experts helps both to counter the loss of knowledge about development cooperation and to establish some measure of institutional memory. With regard to knowledge of projects in region and programme countries, much of that memory will be represented by the new management information system (MIDAS). Its further development will be given a high priority in the coming period.

Bibliography

Al-Jalaly, S.Z. *The project appraisal management and control system in Pakistan, a critical study.* Peshawar, 1991.

Bamberger, M. The monitoring and evaluation of public sector programs in Asia: Why are development programs monitored but not evaluated? *Evaluation Review,* Vol. 13, No. 3, 1989, pp. 223–242.

Berlage, L. & Stokke, O. *Evaluating development assistance: approaches and methods.* London, 1992.

Binnendijk, A.L. Donor agency experience with the monitoring and evaluation of development projects. *Evaluation Review,* Vol. 13, No. 3, 1989, pp. 206–222.

Casley, D.J. & Kumar, K. *Project monitoring and evaluation in agriculture.* Baltimore, 1987.

CIDA. *Standards for bilateral project evaluations.* 1990.

CIDA. *Management of bilateral project evaluation.* 1991a.

CIDA. *Guidelines for developing abstracts of bilateral evaluation reports.* 1991b.

CIDA. *A study of the quality of CIDA's bilateral evaluations 1988–89 and 1989-90,* 1991c.

Clayton, E.S. & Petry, F. *Monitoring systems for agricultural and rural development projects.* Economic and Social Development Paper No. 12. Rome, Food and Agriculture Organization, 1981.

Cracknell, B.E. Evaluation: a code of good practice. *Development Journal,* launch issue, 1992, pp. 8–10.

Cracknell, B.E. *EEC Commission and member states review of effectiveness of feedback mechanisms,* final synthesis report. 1991.

Deutsche Gesellschaft fur Technische Zusammenarbeit (GTZ). *Methods and Instruments for Project Planning and Implementation (outlines).* Eschborn, 1991.

DGIS. *Instructie Evaluatie* (Evaluation Guidelines). 1982a.

DGIS. *Instructie Rapportage* (Reporting Guidelines). 1982b.

DGIS. *Contracting out: Bane or blessing?* The Hague: Ministry of Foreign Affairs, 1989.

DGIS. *A world of difference: A new framework for development cooperation in the 1990s.* The Hague, 1991.

DGIS. *Sector programme for rural development.* The Hague, 1992.

FAO. *Guiding principles for the design and use of monitoring and evaluation in rural development projects and programmes.* ACC task force on rural development, Rome, 1984.

Finnida. *Review of evaluation reports on the Finnish development projects in the 1980's.* Helsinki, 1991.

Frerks, G.E., Thomas, H. & Tomesen, L.B.M. *Effect monitoring and impact evaluation: report of a workshop.* Islamabad 1990.

Hoole, F.W. *Evaluation research and development activities.* Beverly Hills/London. 1978.

Hussain, H.C.M. *Project appraisal, monitoring and evaluation processes, with special reference to Pakistan* (2nd. revised edition). Karachi, 1991.

IFAD. *Monitoring and evaluation concepts and terminology in the United Nations systems: a glossary*. Rome, 1984.

Inter-American Development Bank. *Annual report on operations evaluation during 1989*. Washington D.C., 1990.

Kuyvenhoven, A. & Pelt, M. van (eds.). *Evaluating project performance methodology and development of capabilities*. Den Haag, 1988.

Lawrence, J.E.S. Engaging recipients in development evaluation: The 'stakeholder approach'. *Evaluation Review*, Vol. 13, No. 3, 1989, pp. 243–256.

Leviton, L.C. & Hughes, E.F.X. Research on the utilization of evaluations; a review and synthesis. *Evaluation Review*, Vol. 5, No. 4, 1981, pp. 525–548.

Marsden, D. & Oakley, P. Future issues and perspectives in the evaluation of social development. *Community Development Journal*, Vol. 26, No. 4, 1991.

OECD. *Methods and procedures in aid evaluation. A compendium of donor practice and experience*. Paris, 1986.

OECD. *Evaluation in developing countries. A step in a dialogue*. Paris, 1988.

OECD. *A review of donor's systems for feedback from aid evaluation*. Selected Issues in Aid Evaluation, No. 4. Paris, 1990.

OECD. *Principles for evaluation of development assistance*. Paris, 1991.

OECD. *DAC principles for effective aid*. Paris, 1992.

Palmer, M., Leila, A. & Yassin, E.S. *The Egyptian bureaucracy*. New York, 1988.

Patton, M.Q. *Utilization focused evaluation* (second edition). London, 1990.

Putte, R.A. van de. *Monitoring for development control, the design of a model*. Enschede, 1991.

Rist, R.C. (ed.). *Policy and program evaluation: Perspectives on design and utilization*. International Institute of Administrative Sciences. Brussels, 1990.

Rossi, P.H. & Freeman, H.E. *Evaluation, a systematic approach*. London, 1985.

Sant, J. van. Qualitative analysis in development evaluations. *Evaluation Review*, Vol 13, No. 3, 1989, pp. 257–272.

Scanteam international. *Learning from experience: A study of the feedback from evaluations and reviews in Norwegian bilateral aid*. Oslo, January 1993.

SIDA. *Baseline study handbook, focus on the field*. Stockholm, 1992.

Stokke, O. *Evaluating development assistance: Policies and performance*. London, 1991.

Tomesen, L. & Thomas, H. *Small for development: An account of 10 years Pakistan-Netherlands industrial development cooperation*. Islamabad, 1992.

UNDP. *UNDP policy and procedures manual*. New York, 1987.

USAID. *A.I.D. evaluation handbook*. Washington, D.C., 1987a.

USAID. *Designing monitoring and evaluation systems: Issues and opportunities*. Washington, D.C., 1987b.

USAID. *Methodologies for assessing the impact of agricultural and rural development projects: A dialogue*. A.I.D. Program Design and Evaluation Methodology Report No. 11. Washington, D.C., 1989.

USAID. Review of the 'state of evaluation' in AID. *AID. Evaluation News*, 1992, Vol. 4, No. 1.

Von Oven, R. Common problems in monitoring and evaluation of agricultural development projects. *Zeitschrift für auslandische Landwirtschaft*, 1979, Vol. 18, No. 3.

Weiss, C.H. Evaluation for decisions; Is anybody there? Does anybody care? *Evaluation Practice*, feb. 1988, pp. 5–25.

World Bank. *Built-in project monitoring and evaluation: An overview.* Washington, 1987.

World Bank. *Annual report on operations evaluation, FY90.* Washington, 1991.

World Bank. *Bank experience in project supervision, 1992.* Washington, 1992.

Youssef, S.M. *System of management in Egyptian public enterprises.* Cairo, 1983.

EVALUATION AND MONITORING

1995