Evaluation Policy

2007 – 2010

Contents

Evaluation Policy 2007 - 2010

"Not everything that counts can be counted.
And not everything that can be counted, counts"

Albert Einstein

1 Introduction

As from the strategic plan 2002-2006, Woord en Daad has included the need for evaluations in all its plans. Reasons mentioned were the need for transparancy and for instruments to gain insight in efficiency, effectivity and longer term impact of projects.

Woord en Daad has further developed as a professional organisation that wants work with an increasing quality level. Based on this focus on quality, it is a logical step to formulate and implement a systematic and coherent evaluation policy.
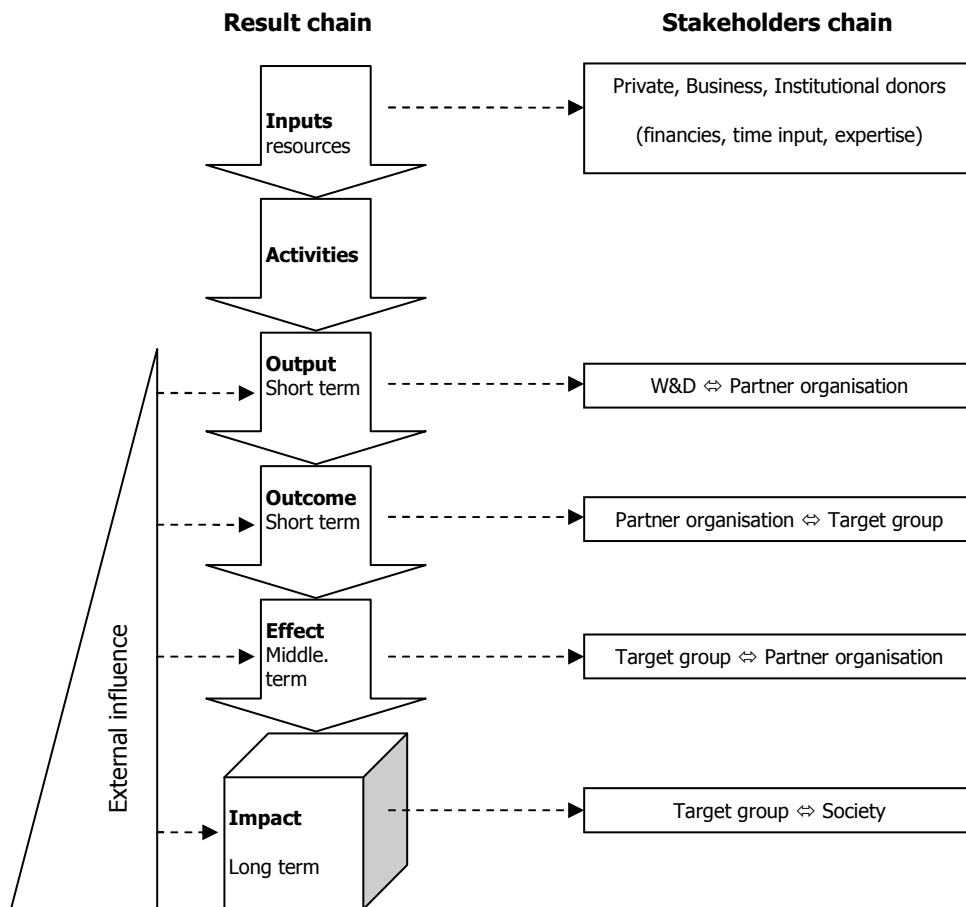
This document explains the definitions, models, objectives and principles that form the basis for this evaluation policy. Paragraph 3 uses the concept of a pyramid to elaborate the way monitoring and evaluation function, while paragraph 4 explains the way in which and the actors involved in evaluation. The final paragraph focuses on how evaluations are fed back into planning and policy making; how evaluations lead to learning.

2 Definitions, objectives and principles

2.1 Definitions and models

- Evaluation is an assessment as systematic and objective as possible of projects that are concluded or in process: its design, implementation and results on longer and shorter term. The immediate objective is to know the relevance of the objectives and the effectivity, efficienty, impact and sustainability of. The most important categories of evaluation questions are briefly explained below:
  - o The relevance of the project: this relates to the logic of the interventions: the relation between objectives and the problem- and context analysis of the target group and the relation between objectives, activities, planned results and indicators of the intervention.
  - o The effectivity of the intervention: the extent to which objectives have been reached. Three subsequent levels of results are distinguished: output, outcome and effect.
  - o The impact of the intervention: positive or negative, primary or secondary, direct or indirect, intended or unintended results that are likely to be caused by the intervention on a longer term.
  - o The sustainability of the intervention: the extent to which it is likely that results continue after external inputs will have finished, and the extent to which results are risk proof.
  - o The efficiency of the intervention: the extent to which objectives have been realised without consuming more resources than necessary.
  - o Process indicators can as well be the main object of an evaluation: planning, monitoring, decision structure, learning and innovation, implementing capacity, participation of the target group, etcetera.
- Monitoring and evaluation are closely linked. The main differences are:
  - o Monitoring is often done at the project level, while evaluations are also done at program level or country level.
  - o Monitoring takes the given project objectives and planned results and indicators for granted, while in evaluations these are also object of the assessment.
  - o Monitoring focuses on intended results while evaluations included unintended results.
  - o Evaluations try to provide insight in causal relations and use several methods of data collection and are therefore instrument for feed back at a higher strategic level.
- Evaluability is the extent to which it is possible to evaluate a program or project in a reliable way. This depends on the extent to which proposed interventions are formulated with intended results and indicators that are quantifiable or can be assessed otherwise. The use of a logical framework as a tool (not as a philosophy) promotes the evaluability of interventions.

The result chain as presented below is used as a model. This chain can be linked to the stakeholders chain in the same figure

Result chain **Stakeholders chain**

Inputs
resources → Private, Business, Institutional donors

(financies, time input, expertise)

Activities

Output
Short term → W&D ⇔ Partner organisation

Outcome
Short term → Partner organisation ⇔ Target group

Effect
Middle. term → Target group ⇔ Partner organisation

Impact

Long term → Target group ⇔ Society
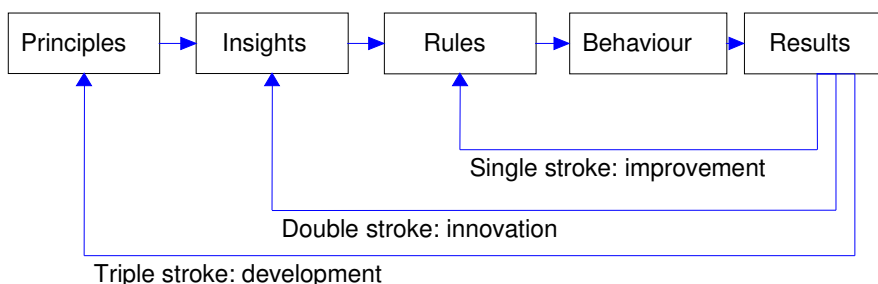
External influence

The following remarks apply to this figure:

- Woord en Daad uses this figure as an instrument and not as a basic philosophy of reality. The same holds true for related instruments like the Logical Framework and DRAM (Objectives - Results – Activities – Means). The figure suggests linear causal relations between input, activities and the results at all levels. In reality it is hardly possible to determine direct causality but rather plausibility. Even if causal relations are present, they are often not linear but complex and often results cannot be claimed by a single intervention or actor, because many actors and factors exert influence. Yet Woord en Daad chooses to use logic instruments while being aware of its limited applicability. For this reason, qualitative methods will be used as well as quantitative methods and especially with impact studies, special attention will be given to methods that are not based on these logic models.
- The terms output, outcome and effect are used as three levels of (intended) results. In general but not strictly, output corresponds to the results at the level of Woord en Daad, outcome to the results at the level of the partner organisation and effect to the results at the level of the target group.
- Apart from these three levels, impact is distinguished as a separate level. These are the effects on a longer term and (often) on the wider society than only the direct beneficiaries. As explained in the definitions above, impact also includes unintended results, both positive and negative. Objectives are formulated at the level of output, outcome and effect only.

For learning in organisations, the model below is used:
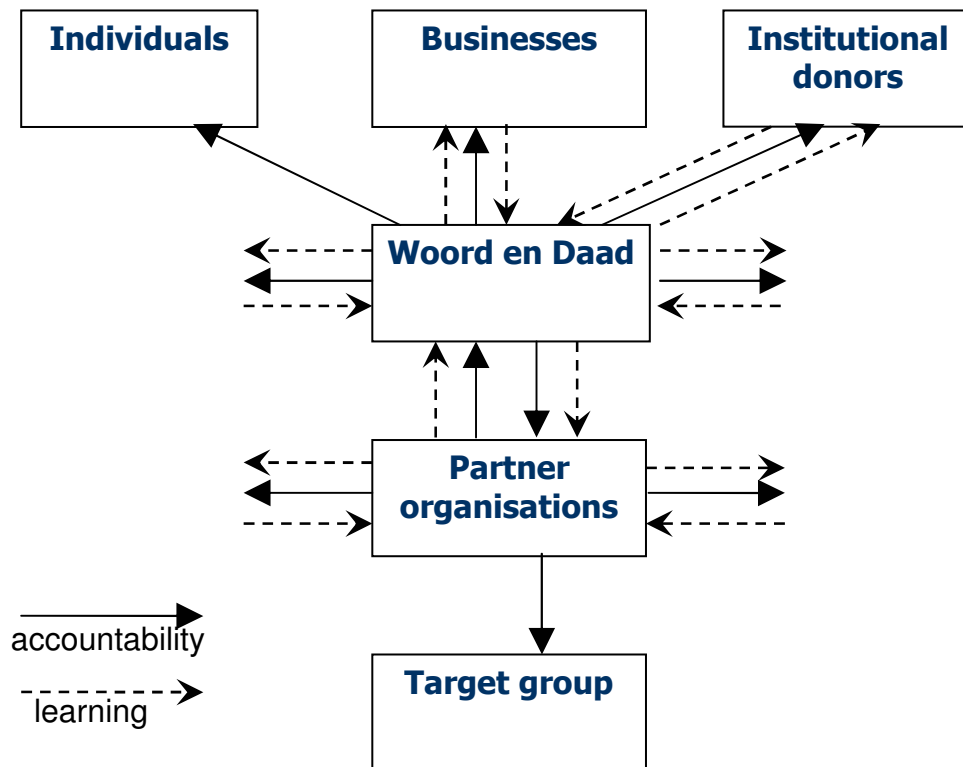
Levels of organisational learning

```
┌────────────┐    ┌──────────┐    ┌────────┐    ┌────────────┐    ┌──────────┐
│ Principles │───▶│ Insights │───▶│ Rules  │───▶│ Behaviour  │───▶│ Results  │
└────────────┘    └──────────┘    └────────┘    └────────────┘    └──────────┘
```

Single stroke: improvement

Double stroke: innovation

Triple stroke: development

	Examples	Examples of lessons learned
Rules – improvement	Project Manual; assessment forms	Adapting assessment criteria, developing MIS
Insights – innovation	Program formulation / chain concept	Set up Small and Medium Enterprise Development Department
Principles – development	Vision, mission and strategies	Adding lobby and advocacy to mission

In paragraph 5 the relation between this model and monitoring and evaluation is made.

2.2 Objectives

Woord en Daad collaborates with partner organisations at one hand and with private, business and institutional donors at the other hand. Apart from this, Woord en Daad has relations with wider civil society, including research institutions and other organisations.

These relations are summarised in the figure below with intended lines of accountability and learning between these actors.

```
Individuals        Businesses        Institutional
                                        donors


                    Woord en Daad


                    Partner
                    organisations


accountability
                    Target group
learning
```

Accountability

Woord en Daad regards it as very important to give account of its stewardship to all stakeholders involved. It is essential that in the chain from donor, through Woord en Daad and partner organisations to the target group not only upward accountability occurs (from partner organisation to Woord en Daad and from Woord en Daad to donors), but also downward accountability (including account-ability toward the target group) and lateral accountability to the wider society in the context of Woord en Daad as well as the wider society in the context of partner organisations.

Learning and improvement

Systematic evaluations also serve as an instrument to be a self learning organisation at all levels in the chain. The objective is that such learning would occur together with the various actors in the chain, from donor to partner organisations (and in a sense also the target group). Evaluations are meant to achieve continuous improvement and if needed adjustment of projects and programs, strategies and methods. Through comparative evaluation studies, best practices may be identified that are applicable in more than one context.

For Woord en Daad, the objective 'learning' precedes the objective of 'accountability'. This is based on the conviction that from an open learning attitude, accountability will probably follow and that real accountability will only occur in a non-threatening environment that focuses on learning.

2.3 Principles

- **Ownership** of evaluations is in the first place with organisations that carry out projects. This principle does not exclude the need for objectivity and external elements in evaluation studies. It is also important that evaluations carried out by partner organisations are done in such a way that the needs of the (sometimes various) donors are also met, so that duplication of evaluations on the same projects can be avoided as much as possible.
  This principle of ownership holds even when Woord en Daad proposes, initiates or coordinates evaluations. In such cases there will rather be shared ownership or at the very least consultation of partner organisations at a very early stage.

Apart from these, program-evaluations at the level of Woord en Daad can be carried out. Ownership of such evaluations is with Woord en Daad. This is further elaborated in paragraph 3.
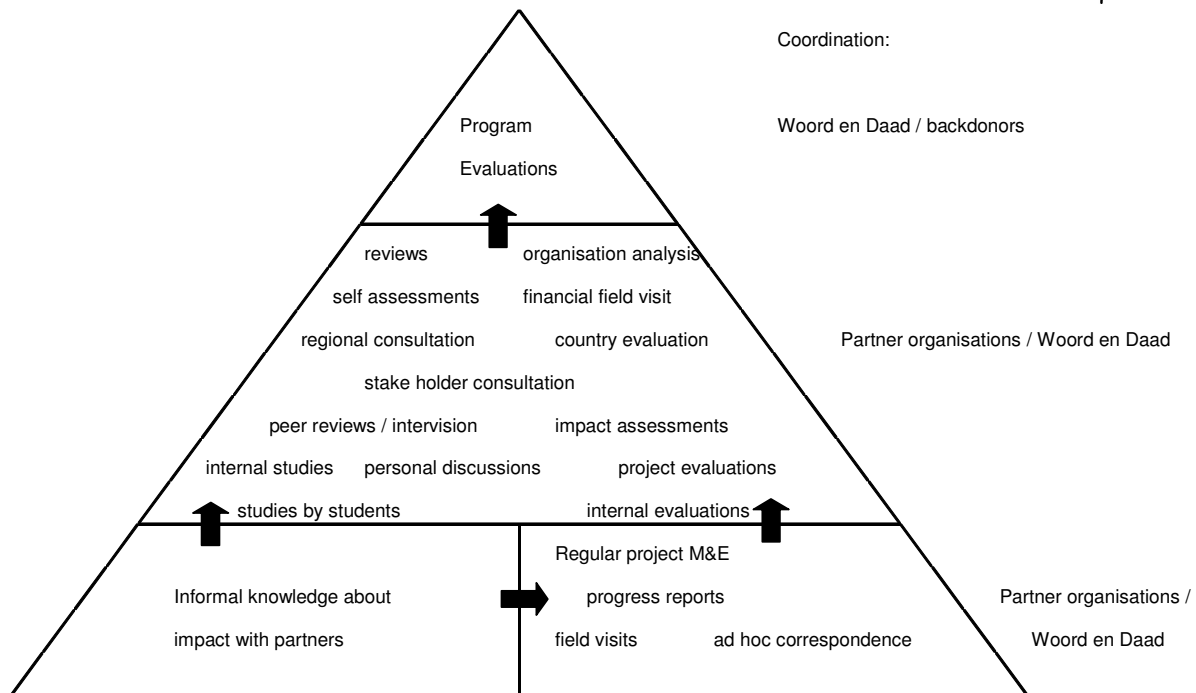
- **Usefulness**. Evaluations should be set up in such a way that the outcomes can be used as input for policy development and adjustment, for partner organisations as well as for Woord en Daad. This principle will serve as a criterion for accepting and selecting evaluation studies to be carried out.

- **Transparency**. This principle is included in one of the objectives of evaluations; to provide accountability. This is why it is important that evaluations themselves are done in a transparent way and that the outcomes are transparent. Outcomes will always be made available to all relevant stakeholders: Woord en Daad personnel, partner organisations, target group, institutional donors, business and private donors. Special attention will be given to providing accountability toward the target group.

- **Quality**. Quality is important in all aspects of the work of Woord en Daad, including evaluations. For all instruments used for evaluations, specific quality criteria apply. Validity and reliability are guaranteed as much as possible. In the description of the evaluation pyramid the quality requirements for each level of evaluations is worked out.

- **Qualitative and qantitative**. Woord en Daad wants to take position against the (neopositivist) idea that reality can be reduced to a set of data and thereby can be planned and constructed. Because of this, qualitative methods will be used as well as quantitative methods and evaluation studies will expressly include a focus on behaviour change, character and mentality formation and power relations between and within stakeholders.
Using a diversity of evaluation instruments and methods also includes that Woord en Daad wants to use and experiment with alternative and innovative methods like the Most Significant Change Method and Outcome Mapping. By using a diversity of methods, a light version of triangulation can be obtained.

- **Plausibility vs. Causality**. The further the evaluation reaches into the results chain, the more the link with the development intervention will be mentioned in terms of plausibility rather than causability, because of the increasing influence of external actors and factors.

3 Monitoring and Evaluation – the pyramid

Implementation of the evaluation and monitoring policy can be summarized with the concept of a pyramid. This concept wants to express that a variety of instruments is made use of, while studies at a higher level make use of lower levels and where the quality of the basis and middle layer of the pyramid are essential for the whole.

The following considerations and principles underlie this concept:

- Studies show that more knowledge about impact is present within partner organisations than formal monitoring lines show. At the same time, short external evaluations are only to a limited extent able to gather new information about impact.

- Carrying out evaluation studies is essential, but at the same time Woord en Daad strives to employ resources as optimal as possible. Therefore it is logical to make as much as possible use of information that is already present.

- The principle that ownership of evaluations should remain as low as possible in the chain, so that the effect on learning remains as large as possible. Because of this, maximum use should be made of information that is already present with partners, through a diversity of instruments that should be adapted to needs of partner organisations.

```
                                                    `                                    Woord
                                                                                         + Daad

                                                                   Coordination:

                                  Program               Woord en Daad / backdonors

                                  Evaluations
                                         ▲

                         reviews          organisation analysis

                    self assessments      financial field visit

                  regional consultation       country evaluation        Partner organisations / Woord en Daad

                          stake holder consultation

                 peer reviews / intervision       impact assessments

              internal studies   personal discussions       project evaluations

                  ▲  studies by students            internal evaluations  ▲

                                            Regular project M&E

                     Informal knowledge about       progress reports                Partner organisations /

                     impact with partners   ▶   field visits     ad hoc correspondence    Woord en Daad
```

The basis

Regular project monitoring is described in paragraph 4.10 of the project manual (Organisational handbook, par. 7.3). The partner organisation reports about the activities, intended results with indicators that have been defined in project agreements with a fixed frequency. In most cases, this means half yearly narrative reporting and quarterly financial reporting.

After each project agreement, department Projects and Programs writes an 'end memo' (or a 'progress memo' in case a project continues). Annual field visits are an integral part of regular monitoring and also ad-hoc correspondence and personal discussions between project officers and partner organisations take place.

Knowledge about impact that partner organisations have is often informal knowledge, which is formed by daily interactions between field staff of the organisations and the target group. These field staff are often close to the target group with respect to the physical and socio-economic context. Personal discussions play an important role in making available such information for the information channel of the regular monitoring. Apart from this, in discussion with partner organisations ways should be found to make this knowledge available for further use in the chain. This knowledge should be assessed through various means in the middle layer of the pyramid.

The middle layer

The middle layer uses the basis of the pyramid as an input. It makes use of a variety of instruments. Some examples are:

- Studies where local students stay in the project area for several months and report all changes they perceive or that people inform them about.
- Impact assessments that make use of alternative methods, like the Most Significant Change method and Outcome Mapping.
- Internal or external impact studies.
- Reviews, where internal or external desk studies are carried out, based on information that is already present.
- External project evaluations.
- Personal discussions with partner organisations and people from the target group.
- Various internal or external evaluations of projects, processes, themes, organisations or country programs. These include (facilitated) self-assessments.

- Regional partner consultations where experiences of projects and programs is shared between partner organisations and with Woord en Daad.
- Stakeholder consultations, that include the target group
- Financial field visits. These are paid to partner organisations every five years and are described in the Project Manual, par. 8.6.
- Intervision processes and peer reviews. This may be done at project level, but also at program level. It could be done between partner organisations or with other local organisations. It may also be done between Dutch organisations.

The intention is that this variety of instruments will be used in such a way that for all program-evaluations (the top of the pyramid) relevant and useful building blocks are present. The combination of this variety of methods can be regarded as (a simple form of) triangulation: when results of different instruments tend to confirm each other, the probability increases that such results correspond to reality.

The top

Program-evaluations are evaluations of programs of Woord en Daad (Education, Job and Income, Basic Needs and Emergency Relief). Such evaluations may focus on the whole program or on a specific theme. Use will always be made of the building blocks from the first two layers of the pyramid. After an extensive desk study that summarizes these input, a field study can be done in order to validate the desk study. The evaluations that serve as inputs may be sampled for validation.
The main difference between these evaluations and the levels below is that ownership is fully with Woord en Daad in this case. As much as possible, such evaluations will be discussed with back donors in order to avoid duplication of evaluations. The initiative for these evaluations may also be with institutional donors. Another difference with evaluations of the middle layer is that these are always more formal evaluations with higher quality requirements. Each program of Woord en Daad will be evaluated once in four years.

4 How and by whom

4.1 *Quality requirements*

- Basis: the quality requirements for regular monitoring of projects is laid down in the Project manual (paragraphs 4.10, 8 and 9). For informal knowledge about impact no quality requirements can be formulated because of its informality. Instruments of the middle layer that attempt to assess this information are subject to quality requirements.
- Middle layer: the quality requirements and the procedure for assessing evaluation proposals is laid down in paragraph 7.7 of the organisational handbook.
  - o For each evaluation, a Terms of Reference has to be defined in consultation with relevant stakeholders. A standard format is developed that defines the minimal criteria for the information provided in the Terms of Reference (Handbook 7.7.2.1).
    - Before the evaluation is carried out, assessment takes place. A standard form is developed (handbook 7.7.2.3). Assessment criteria include relevance, usefulness, clarity and reliability of the methodology to be used, relation to Woord en Daad program and verifiability (so that the evaluation can be used as input for a program evaluation).
    - After each evaluation project, an end memo is written, which is used for feed back of conclusions and recommendations. See paragraph 5 of this document.
  - o Assessment of these quality criteria is guaranteed because in internal and external audits, the procedure for evaluations is included.
- Top: quality requirements applicable to program evaluations at the top of the pyramid, are::
  - o The same criteria as appy to evaluations of the middle layer of the pyramid.
  - o An external committe or researcher is always involved with the complete evaluation process.
  - o Additional specific criteria for validity, reliability and usefulness could apply after consultation with institutional donors.

4.2 By whom

The evaluation policy and discussion about it with partner organisations is coordinated by the department Advice and Research. Evaluation plans are made by partner organisations in the first place. The intention is that evaluation plans are an integral part of multi-annual plans of partner organisations and that assessment and monitoring of these is also an integral part of the work of the department Projects and Programs. Until this situation is realised, the department Advice and Research is also involved in direct monitoring and assessment of evaluations, especially when these include more than one project. Program-evaluations are fully the responsibility of Advice and Research.

Partner organisations make as much as possible evaluation plans, integrated in multi-annual plans. For these evaluations, an amount of 2% of the costs of the project portfolio is reserved (1% of costs of investments for infrastructure). Evaluation plans have to cover the complete project portfolio, but no rigid and fixed rules are set that each individual project should be evaluated with a fixed frequency. To guarantee that these evaluations can serve as inputs for program evaluations, it is important that each evaluation is done in consultation with Woord en Daad.

For program evaluations an external consultant or committee will always be involved during the whole project in order to safeguard independence. This also holds for the phases that will be executed internally. The following phases can be distinguished for program evaluations:

- Planning.
  - o Designing Terms of Reference, including evaluation team.
  - o Responsibility: Advice and Research.
  - o Execution: Advice and Research.
- Orientation.
  - o Desk study, in which the results of the building blocks from the lower two layers of the pyramid are summarized.
  - o Execution: in most cases Advice and Research, possibly in conjunction with partner organisations. Also possible by external consultant.
- Analysis of these data.
  - o The results of the orientation study are analysed and the main conclusions summarized in a report. Based on this, a specific ToR can be designed for the in depth study.
  - o Responsibility: Advice and Research.
  - o Execution: in most cases Advice and Research, possibly external.
- In depth study
  - o In this phase, samples of the used building blocks of the first two layers of the pyramid can be externally validated. Also the orientation study can be validated. In many cases a field study will be part of the in depth study.
  - o Responsibility: Advice and Research designs a specific ToR for the in depth study (at least if a field research is part of this phase)
  - o Execution: external.
- Analysis, conclusions and reporting
  - o Responsibility and execution: evaluation team.
- Recommendations and feed back into policy
  - o Responsibility: the external evaluation team will include recommendations, based on the conclusions. Independent of this, Woord en Daad (and possibly partner organisations) will formulate their own recommendations based on the conclusions.

5 Feed back and learning

Feed back to stakeholders

Feed back to all relevant stakeholders needs to be included sufficiently in all evaluations. Fee back to the target group requires special attention and in most cases special activities, because just making the reports available (upon request) will in most cases not lead to real downward accountability.

Feed back into policy making – how evaluations lead to learning

Because learning and innovation are a main objective of this evaluation policy, the feed back of evaluation results is as important as the evaluation itself.

The model for learning in organisations can be linked to the concept of the pyramid in the following way:

Level of learning		Level in the pyramid		Level of feed back
Single stroke learning	Improvement	Basis	Regular monitoring	Staff
Double stroke learning	Innovation	Middle layer	Several evaluation studies	Management Team
Triple stroke learning	Development	Top	Program evaluations	Board

The third level of learning, development, does not only happen through individual program evaluations, but rather by comparing and analysing conclusions from different program evaluations. In this way not only improvements and innovations within programs will be realised but learning also occurs about the positioning of the different programs.
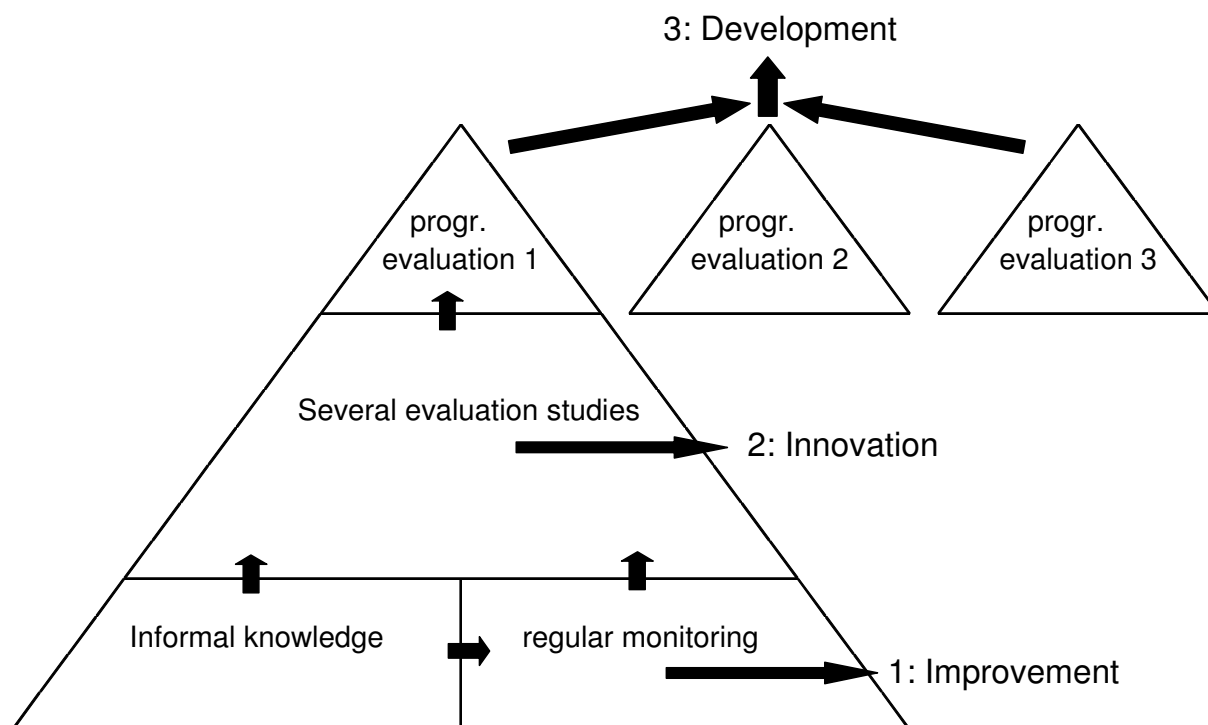
For each evaluation an end memo is made in which briefly the main conclusions (and possibly recommendations) are listed. These end memo's (as mini summaries of the underlying reports) are always input for the management team when evaluations from the middle layer are concerned. For evaluations of the top of the pyramid the end memo's are also input for the board of Woord en Daad. In this way, learning at all levels in the organisation is encouraged.

When writing the annual plan, the department Advice and Research reserves specific time to gather and analyse lessons learned in the past year, among others from evaluations done, so that such lessons will be included in making the new year plan.

Partner organisations are also encouraged to build in mechanisms into their organisations to systematically make use of lessons learned. The objective of this is to become self learning organisations throughout the chain. At each regular field visit, evaluations of the past period and their implications will be discussed. Conclusions from wider evaluation studies can be discussed during regional or international partner consultations.