

Evaluation practices of German development aid agencies

**Monitoring of the implementation of the last system review's
recommendations**

Summary

Published by

German Institute for
Development Evaluation (DEval)
Fritz-Schäffer-Straße 26
53113 Bonn, Germany
Tel: +49 (0)228 33 69 07-0
E-mail: info@DEval.org
www.deval.org

Authors

Dr. Kim Lücking
Simon Freund
Simon Bettighofer

Bibliographical reference

DEval (2015a), *Evaluierungspraxis in der deutschen Entwicklungszusammenarbeit. Umsetzungsmonitoring der letzten Systemprüfung und Charakterisierung wesentlicher Elemente*, DEval, Bonn.

© German Institute for Development Evaluation (DEval)

Introduction

The German Institute for Development Evaluation (DEval) began its work in summer 2012. One of the first projects to be undertaken was the "Baseline study of DEval and its environment".

The project covers two mutually complementary areas of enquiry:

- The first part is a survey of the situation from which DEval started and establishes a baseline for the objectives the institute has set itself.
- The second part is based on a previous study of the evaluation system used in German development cooperation. Conducted in 2007/2008 at the request of the Federal Ministry for Economic Cooperation and Development (BMZ), this "system review" was intended to determine the status of German development cooperation evaluation and, on this basis, make recommendations for action to further advance evaluation practice (Borrmann and Stockmann, 2009a). This part of the study includes a monitoring exercise to review implementation of these recommendations and draws on further findings on evaluation practices in German development cooperation.

Whereas the data on DEval's baseline situation are primarily of interest to the institute itself, the target audience for the second area of enquiry is more diverse. To do justice to the interests of these target groups and to reflect the scope of the results, two reports have been compiled (on the second part of the study cf. DEval, 2015).

This report addresses the implementation monitoring, a task mandated by the Federal Ministry for Economic Cooperation and Development (BMZ), as well as reviews the current status of evaluation practice in German development cooperation.

Implementation monitoring and current evaluation practice

The BMZ initiated the first ever systematic study of evaluations in the field of German development cooperation in 1997/1998 (Borrmann et al., 1999). The study in 2007/2008 was, therefore, the second of this kind and will be referred to below as the "system review".

This "system review" provided recommendations for reforming the evaluation system as a whole. One of the main recommendations was to "establish an independent evaluation agency which advises and supports the BMZ on system development and the development cooperation organisations in the development of instrument and in improving quality. Such an agency should conduct evaluations independently with a high level of competence and credibility, particularly with regard to effectiveness, efficiency and sustainability of development cooperation" (Borrmann and Stockmann, 2009a: 5¹). With the founding of DEval, the BMZ has indeed fulfilled this system-changing recommendation.

Further recommendations at the system level were based on 20 case studies. They were designed to answer the questions of "how the individual development cooperation organisations institutionalise, plan, implement and use evaluations, whether evaluation is enabling them to learn about the impacts of their projects and programmes, and whether and how their individual evaluation systems fit together to form a consistent overall system for bilateral German development cooperation"

¹ author's translation

(Borrmann and Stockmann, 2009a: 5²). Looking at the organisations surveyed in the case studies, the "system review" arrived at specific recommendations for improving evaluation practices in each one.

The BMZ subsequently entered into agreements with the respective organisations on the implementation of the recommendations they had mutually accepted. As part of the "Baseline study of DEval and its environment", DEval was requested by the BMZ to conduct an "implementation monitoring" exercise to determine whether and to what extent these recommendations have been put into practice. Given that not enough time had elapsed to justify a new "system review", DEval's study did not appraise the organisations individually nor make individual recommendations outside of drawing general conclusions. Thus, for this part of the enquiry, the data was aggregated at the level of the various groups of actors: implementing organisations, political foundations and other civil society organisations.

The analysis was complemented by in-depth enquiries into the evaluation practices of various German development cooperation actors in order to identify trends, as well as to determine the status quo. The insights gained contribute to a better understanding of the environment in which DEval must position itself in order to operate effectively.

Method used

On the basis of a rapid evidence assessment of current studies, six dimensions of evaluation practice were distinguished. These individual dimensions are: *evaluation objectives, organisational context, structures and processes, human resources, financial resources and technical resources*.

The methods employed to gather the data included an online questionnaire survey of evaluation practice in German development cooperation, in which 34 organisations took part. Additional questions regarding the implementation of the recommendations of the last "system review" were also posed to the 17 organisations that had previously participated in this review. Moreover, to achieve a better understanding of their implementation record and deeper insight into evaluation practice, DEval conducted extended interviews over one or two days with these organisations, speaking at length with managers, staff with evaluation responsibilities and staff in operational roles. To gain an impression of the designs and methods being applied in German development cooperation, a meta-evaluation was made of evaluation reports produced by development cooperation organisations. This survey was complemented by assessments drawn from national and international experts in development cooperation.

At a wrap-up workshop, bringing together representatives of the organisations participating in the survey, there was a concluding discussion of the results obtained and conclusions to be drawn.

General findings

The surveys show that the "system review" provided an important stimulus for advancing the evaluation practice of German development cooperation. The changes it initiated can be described – in view of their scope and depth – as system-forming.

² author's translation

The implementation monitoring included an analysis of the recommendations of the "system review". It was found that, of the 181³ recommendations, 38 per cent were directed at six political foundations, 35 per cent at the five governmental implementation organisations, and 27 per cent at six other civil society organisations. Regarding the implementation plans in order to measure conformity with the recommendations the survey found that the organisations agreed in full with 59 per cent of the recommendations and with a further 33 per cent at least in part. This majority support for the recommendations is further reflected in the finding that most of them have been implemented in full or at least in part. The data drawn shows that 54 per cent of recommendations were implemented in full, 29 per cent in part and just 12 per cent not at all. No data on implementation is available for 5 per cent of the recommendations.

By far the most important factor driving the implementation of recommendations seems to be agreement with them on the part of the leadership or management of an organisation. Financial resources are, on the other hand, often found to have had a constraining influence on implementation. This applies in particular to those recommendations that required substantial additional expenditure (e.g. to set up a new staff unit, to conduct more ex post evaluations).

Results by dimension

The analysis of the individual dimensions of evaluation practice show that the main focus in the recommendations was on the dimension of "structures and processes". Next came a high proportion of recommendations concerned with "evaluation objectives" and "human resources". Together those three categories accounted for nearly four fifths (79 per cent) of all the recommendations made. Significantly fewer recommendations of the "system review" concerned the dimensions of "organisational context", "financial resources" and "technical resources".

Evaluation objectives

The recommendations in this dimension were intended to extend the spectrum of evaluations by development cooperation organisations with regard to their own goals and thus increase

- the diversity and benefits of evaluations conducted (e.g. ex post, impact, joint and cross-sectional evaluations) and
- the use and transparency of evaluation results.

The proportion of the 49 recommendations in this dimension that met with full agreement comes to 47 per cent, which is lower than in any of the other dimensions. Moreover, the proportion of these recommendations that were not or only partially implemented is relatively high, standing at 51 per cent (n=25). One explanation for this finding is that many of the recommendations call for more forms of evaluation to be used. Although more cross-sectional evaluations are now being conducted, development cooperation organisation respondents expressed doubts about the usefulness of ex post and impact evaluations for their own work. Some development cooperation organisations also said they do not have the necessary funds to perform such evaluations. Consequently, ex post and impact evaluations, which are more resource intensive, tend to be carried out by government implementing organisations rather than civil society organisations.

The organisations surveyed reported that evaluations are used above all for the improvement of projects and programmes or for instrumental learning. This happens primarily through project participants for programme/project management purposes. Many organisations lack the standardised

³ In the system review, 128 recommendations were made at the organisational level. Since some of these recommendations were very wide-ranging, so they have been subdivided according to various content aspects.

processes needed in order to use cross-project evaluations for conceptual learning within the organisation as a whole. The commitment to accountability indeed ties resources to documenting project progress – resources which some organisations would like to use for conceptual learning from cross-project evaluations.

With regard to the use and transparency of evaluation findings, the survey found very few cases of organisations publishing full reports as a matter of principle. The practice now commonly found in most development cooperation organisations is the publication and dissemination of summary reports. In some cases, the full reports are made available on request. Factors which can prevent individual organisations from wanting to publish evaluation results include legitimate interests in protecting partners as well as concerns of some with regard to the level of donations. There is also a worry in some organisations that complete transparency might contradict the aim of learning from evaluation, because critical results might, in certain circumstances, be less addressed, if full reports are to be published without comment.

Organisational context

The aim of the recommendations in the dimension of "organisational context" was to

- increase networking among the development cooperation organisations and with the research community specifically on matters of evaluation and
- foster understanding and acceptance of evaluation among staff.

The majority of the 16 recommendations in this dimension were shared by the development cooperation organisations (69 per cent), with 62 per cent (n=10) of them being implemented in full. On the whole, the development cooperation organisations were in agreement with the idea that sharing information and experience through evaluation networks adds value and can help to improve the quality of evaluations. The degree of networking among the German development cooperation organisations in the evaluation field has increased significantly in recent years. The same can be said of collaboration with the research community, although to a lesser extent. Staff members of the development cooperation organisations are open-minded and receptive with regard to evaluation.

Structures and processes

The aim of the recommendations in the "structures and processes" dimension was to strengthen

- the role of an evaluation unit within the organisation,
- the institutionalised foundations of evaluation,
- the defining of required evaluation processes, and
- the inclusion of partners in the evaluation process.

In terms of absolute values, this is the dimension in which the most recommendations (n=56) were made and implemented (31 in full, 19 at least in part). In relative terms, however, the number of fully implemented recommendations is at an average, at 55 per cent. Agreement with the recommendations in this dimension was calculated at 64 per cent.

Within the four content-defined clusters of recommendations, major differences were found in their implementation. Although there is a high level of willingness to establish clearer evaluation structures and processes, there is a comparatively high proportion of only partly implemented recommendations regarding the institutional role of evaluation units and the inclusion of partners in the evaluation process.

Evaluation units and evaluator posts have, in many cases, been newly created or strengthened since the "system review". Most personnel in charge of evaluation regard themselves as service providers

for their organisation, and less as providers of independent scrutiny. In discussing their role, they raised the question of how far an internal unit in an organisation can be called independent or achieve independence. Moreover, concerns were also voiced that greater independence of their work might only be achieved at the expense of the relevance and acceptance of evaluation results.

Taking account of OECD DAC criteria has been embedded in the evaluation guidelines for German development cooperation organisations. These criteria do not, however, appear to have been consistently applied in the evaluation reports.

On the question of institutionalised foundations of evaluation, the survey finds that the German development cooperation organisations have succeeded in formalising and systematising their evaluation practices. Evaluations now largely follow evaluation guidelines and plans. Nevertheless the degree of formalisation varies according to the phase in process of evaluation. The phase of transferring and applying evaluation results is the least formalised, even though some organisations do have standardised procedures for appraising and implementing findings.

As for the objective of including partner organisations from recipient countries in the evaluation process, it is found that the question of whether and to what extent partner organisations are integrated strongly depends on the type of cooperation. The overall finding was that systematic and equitable inclusion of partners in the evaluation process is not the rule.

Human resources

The aim of the recommendations in the "human resources" dimension was to advance

- staff numbers working in the evaluation unit,
- the monitoring and evaluation expertise of staff,
- the monitoring and evaluation expertise of partners,
- the use of independent experts, and
- the methods being applied.

Of the 38 recommendations, the organisations were found to have agreed with 66 per cent, while 61 per cent were fully implemented. Since the "system review", more human resources have been deployed on evaluation. In some organisations, however, the upscaling and upgrading process was not yet completed at the time of the survey. The methodological standards expected by organisations have also risen in many cases, which means evaluations require more attention and have become more labour-intensive.

The level of specialist expertise found in evaluation units has risen. However, the quality of an evaluation depends not solely on the efforts of the evaluation team, but also on the expertise of expert consultants who are brought in and on the expertise of operational staff and partners who maintain managerial roles in the evaluation. It is not possible for evaluation units to provide systematic quality assurance in all cases, especially in view of the large number of evaluations undertaken in some organisations.

With regard to the commissioning of external appraisers, nearly every organisation covered by the survey was found to have increased the use of such consultants, some even giving them sole responsibility for reporting. Referring to the repertoire of methods used by external appraisers, some organisations note that there is room for improvement. More weight tends to be given to country-specific and organisation-specific knowledge when selecting consultants.

Many German development cooperation organisations do not regard building evaluation capacities in partner organisations as one of their own tasks, even though some organisations are very active in this area.

It was frequently found that evaluation designs are not based on verifiable impact hypotheses. There is room for improvement in the spectrum of evaluation methods applied and in the use of an appropriate mix of methods.

Financial resources

The aim of recommendations made in the "financial resources" dimension was to strengthen budget allocations for evaluations and evaluation systems.

Twelve recommendations referred to aspects of the financial framework. Overall, 58 per cent of the recommendations were followed and 42 per cent fully implemented. The fully implemented recommendations include the need to earmark sufficient funds for ex post evaluations and for more sophisticated survey designs. As for the two recommendations that were not yet implemented, these concern a general increase in funding and a binding commitment to set an evaluation budget.

Evaluation practice depends in many respects on the financial resourcing of the evaluation teams within an organisation. Since the "system review", there has been an increase in the funds earmarked for evaluation within the German development cooperation organisations. Most respondents stated that they consider the financial resources provided to be sufficient. This view does not, however, apply to allocations for conducting methodologically sophisticated or cross-project evaluations.

Technical resources

The aim of recommendations made in the "technical resources" dimension was to facilitate or improve evaluation workby, for instance, enabling access to statistical or survey software, evaluation literature, or to an internal knowledge management system.

The majority of the 10 recommendations were both shared and fully implemented by respondents (each scoring 60 per cent; n=6). It was found that staff in the German development cooperation organisations have good access to specialist literature and evaluation manuals. Moreover, the majority of organisations surveyed make use of knowledge management systems, although predominantly for retrieving information rather than for institutional learning purposes. On this point, representatives of development cooperation organisations stated that specific technical resources do not necessarily need to be made available to themselves as commissioners, but rather to the external consultants.

Conclusion

The implementation monitoring of the "system review" recommendations shows that there has been remarkable progress in evaluation practice in German development cooperation. For example, it was shown that in the organisations surveyed, these specific recommendations contributed to the creation of dedicated evaluation units, to improved staffing levels for evaluation and to larger evaluation budgets. Moreover, there has been a wider introduction of evaluation guidelines. On the other hand, just as in previous system reviews, it was again found that there are areas in which the requirements of the "system review" and the OECD DAC principles do not entirely coincide with the interests of the development cooperation organisations surveyed. Examples of this particularly pertain to the demand for ex post evaluations in order to assess the impact and sustainability of projects and programmes, or for increasing the transparent use of evaluation results.

On the whole, the study shows that the "system review" was a key driver for the practice of evaluation in German development cooperation, as well as promoted or initiated many developments.