



Federal Ministry  
for Economic Cooperation  
and Development

EVALUATION REPORTS 044

# Evaluation in German Development Cooperation

A System's Review



# Contents

<b>Preface by the Federal Ministry for Economic Cooperation and Development</b>	<b>3</b>
<b>Preface by the Authors</b>	<b>4</b>
<b>Executive Summary</b>	<b>6</b>
1 Evaluation of development cooperation under changing conditions	6
2 Findings	7
3 Recommendations	14
<b>Comments of BMZ's Management</b>	<b>19</b>

# Preface by the Federal Ministry for Economic Cooperation and Development

Evaluations have been conducted by the German Federal Ministry for Economic Cooperation and Development (BMZ) since the 1960s and increasingly over time also by its implementing agencies and civil society organisations (CSOs) receiving funding from the BMZ. A first review of self-evaluations and independent evaluations of these organisations was published in 1999.

The new century brought profound changes in the way development cooperation is planned and implemented: Heads of states agreed on the Millennium Declaration with measurable goals to be attained by 2015 to reduce world poverty. The Paris Declaration on Aid Effectiveness of 2005 and later on the Accra Agenda for Action spelled out core principles for cooperation between donors and partner countries. Internally, the ministry has been pushing increasingly for more “joined-up” development cooperation aligning the work of its various implementing agencies to strategic goals negotiated between BMZ and partner country governments.

With a view to learn whether the German system of development evaluation performs adequately as measured against DAC evaluation norms and standards and new challenges emanating from changing framework conditions, the evaluation division of BMZ commissioned a second review of the evaluation function. The quality, independence, and utility of evaluation in BMZ itself, its implementing agencies and 12 CSOs receiving sizable funding of the ministry was reviewed and assessed as well from a system’s perspective. The study has been conducted by **Axel Borrmann** and **Reinhard Stockmann** (team leaders and authors of the synthesis report), **Paul Kevenhörster**, **Wolfgang Meyer**, **Katharina Michaelowa**, and

**Jörg Rech**. The team was assisted and advised by two peers of BMZ: **Dominique de Crombrughe**, Special Evaluator of Belgian Development Cooperation and deputy head of the DAC evaluation network and **Gerhard Siegfried**, head of Evaluation and Controlling of the Swiss Agency for Development and Cooperation (DEZA).

Reviewing the evaluation systems of 20 organisations turned out to be a daunting task – for the evaluators and for the reviewed agencies and BMZ. The study, which started with a kick-off meeting in December 2006, took almost two years to be completed. The report was finally published in two volumes in May 2009. An English translation will follow.

The report reflects the state of evaluation in October 2007. As usual, the opinions presented in this study are those of independent external experts and do not necessarily reflect the views of BMZ or of the other organisations assessed.

This summary report can be downloaded from BMZ’s website <http://www.bmz.de/en/service/in-fothek/evaluation/BMZEvaluationberichte/index.html>. Comments of BMZ’s management on the assessment are included at the end of this report.

It should be cited as: Borrmann, A; Stockmann, R. (2009): Evaluation in German Development Cooperation. *Evaluation Reports 44*. Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung.

*Evaluation and Audit Division  
Federal Ministry for Economic Cooperation and Development*

## Preface by the Authors

There is a long tradition of evaluation in German development cooperation. Nevertheless, the salience of evaluations differs greatly among individual development cooperation agencies. This study – which has been commissioned by the German Federal Ministry for Economic Cooperation and Development (BMZ) – has adopted a comprehensive, complex methodological approach to determine how individual agencies institutionalise, plan, execute and use evaluations, whether they are able to generate findings relating to the impact of their projects and programmes through evaluations and whether and how they pool their individual evaluation systems in a consistent overall system for bilateral German development cooperation.

The study covers twenty development cooperation organisations for which individual case studies had to be performed in order to produce the foundations upon which to assess the system as a whole. With such a large number of cases it was necessary to develop an evaluation strategy which has the appropriate level of complexity, but is at the same time practicable and allows horizontal comparisons. Our approach was based on pertinent rules applied to academic work to ensure the transparency of the approach and the results produced. It is for this reason that all the case studies in the second volume of this publication have been published without reservation.<sup>1</sup> This conforms with the requirement of the evaluation standards which apply to transparency. The study was based on the OECD-DAC Principles for Evaluation and the standards of the German Society for Evaluation (DeGEval).

The Hamburg Institute of International Economics (Hamburgisches WeltWirtschaftsInstitut), the Center for Evaluation (Centrum für Evaluation) and the two authors and directors of this study would like to express their gratitude to the German Federal Ministry for Economic Cooperation and Development for the special opportunity to systematically assess evaluation work in an entire field of policy – especially because these opportunities are so rare.

It is now up to the analysed organisations to take advantage of the findings and recommendations contained in this summarising volume and the documentation of case studies in their own work. We hope that our study will strengthen the assessed organisations' managements' belief or convince these managements that evaluation is first of all an important steering instrument with which to improve the quality of development cooperation and secondly, a means with which to persuade the public of the effectiveness and sustainability of projects and programmes which have been executed. We wish the Federal Ministry for Economic Cooperation and Development the energy and courage to not wait for the major organisational reform of development cooperation before making important decisions, but instead to make these decisions now – in order to put an end to the fragmentation of the German system for evaluating development cooperation and to harmonise and guide it. This study contains a number of fundamental strategy proposals especially regarding the demanding task of system development.

---

<sup>1</sup> Borrmann, A; Stockmann, R.(Ed.) (2009): Evaluation in der deutschen Entwicklungszusammenarbeit [*Evaluation in German Development Cooperation*], Vol. 2: Fallstudien [*Case Studies*], Münster; Waxmann (Vol. 2 in German only; Vol.1 forthcoming in English)

A study examining such a large number of organisations could only be carried out with a number of evaluators. This posed not only an organisational challenge – it also required careful quality assurance in the case studies, which were performed by a total of six evaluators. In addition to the two directors and authors of this report, the case studies were performed by Prof. Dr. Katharina Michaelowa, Prof. Dr. Paul Kevenhörster, Dr. Wolfgang Meyer and Jörg Rech. We would like to express a special word of thanks to these co-authors as well as the two peers, Dominique de Crombrughe and Gerhard Siegfried, who provided critical support in the study, and the director of BMZ's Evaluation and Audit Division, Michaela Zintl, who at the same time commissioned the study, mediated the work and acted as a referee while representing one of the organisations which was examined – the Federal Ministry for

*Axel Borrmann*

HWWI  
Hamburgisches WeltWirtschaftsInstitut gGmbH  
Heimhuderstr. 71 | 20148 Hamburg  
Tel: +49 (0) 40 / 34 05 76-0  
Fax: +49 (0) 40 / 34 05 76-776  
info@hwwi.org  
<http://www.hwwi.org>

Economic Cooperation and Development. We would like to thank the development cooperation agencies, those persons in charge of their evaluations and the numerous interview partners at various levels for their tremendous efforts and patience in satisfying our insatiable thirst for information and their ready willingness to discuss things with us.

We would furthermore like to thank all the members of our study team, at the top of the list Jörg Rech from CEval, who coordinated the complex assessment process in an outstanding manner. A special word of praise goes out to our research aids, Miram Grapp, Klaus Magdon, Jan Fendler, Hannah Becker, Julia Schneider, Ragnhild Barbu, Ruth Hoekstra, Sabina Ramonath, Judith Rucker and Angelika Nentwig, without their help this report would not have been possible.

*Reinhard Stockmann*

CEval  
Centrum für Evaluation | Universität des Saarlandes  
Postfach 151150 | 66041 Saarbrücken  
Tel: +49 (0) 681 / 302-3320  
Fax: +49 (0) 681 / 302-3899  
r.stockmann@mx.uni-saarland.de  
<http://www.ceval.de>

# Executive Summary

## 1 Evaluation of development cooperation under changing conditions

As a result of a new development policy agenda, the entire development cooperation system has been undergoing fundamental change since the turn of the millennium (through the Millennium Development Declaration and Goals (MDGs), the Monterrey Consensus, the Marrakech Declaration and the Paris Declaration). The pressure to achieve tangible success in poverty reduction requires an even more resolute and consistent steering of development cooperation based on results and evidence of impact.

Evaluation is thus being assigned an ever more important role. If existing evaluation systems are not modified, however, donors will no longer be able to adequately perform this task. German development cooperation is no exception here. What is needed is inter alia more complex evaluation designs and methods in order to be able to measure impact – including at a higher level of aggregation – a stronger national and international networking of institutions in charge of evaluation and more integration and responsibility (ownership) by partners in the planning and execution of evaluations.

Turning to German development cooperation, there is a two-fold need for changes to be instituted. The first task is to address those elements of the agenda emanating from the last system evaluation which are still relevant. Moreover, German development cooperation has to adjust to changed international conditions and undertake additional reforms.

Although German development cooperation agencies can be said to display tremendous adaptability, making concrete efforts and in

many instances achieving a high level of professionalism in their evaluation policy and actual practice, there are some fundamental problems to be found in the lack of a harmonised evaluation system. The evaluation system of German development cooperation is suffering from enormous institutional, conceptual, terminological and methodological heterogeneity. The DAC also criticised this situation in its last peer review. Reforms along these lines should thus be assigned top priority in the near future.

Because the lack of a harmonised evaluation system largely results from the general institutional fragmentation of German development cooperation, thus being exogenous, and the “joining up” of individual structures of German development co-operation which has been called for by so many actors is still floundering along without much progress to show, efforts to achieve endogenous reforms of the evaluation system may well be running up against their limits.

Nevertheless, development cooperation management and those parties responsible for evaluations should not wait for a sweeping reform of development cooperation, but instead take advantage of the numerous possibilities holding out the promise of success by pushing forward and systematising evaluations through endogenous reforms of the individual evaluation systems and general coordinating measures in order to be better equipped to meet future challenges.

To encourage a greater systematisation of evaluations, this study submits a number of specific proposals for action to be taken, adding a list of reforms which above all pursues the goal of significantly increasing the credibility and quality of evaluations. The recommendations are founded on a broad-based, multi-level analysis of the pe-

riod 2001 to October 2007. As was the case with the first System Analysis, the proposals are above all based on the DAC Principles for evaluations in the field of development cooperation. The results of the study are summarised in the following on the basis of these principles, with this summary being followed by a presentation of the most important recommendations.

## 2 Findings

### Independence and credibility

As set out by the DAC, the German development cooperation agencies have primarily developed an evaluation policy which indicates what goals are pursued with evaluations, what rules and methods are to guide actions and where domains of responsibility are located in institutions. The degree of coherence, systematisation and detail vary, however, and there is room for improvement at a number of development cooperation agencies.

Statutory provisions relating to budgets, international standards and the rules and guidelines of the BMZ and civil society organisations have not been able to prevent definitions, institutional arrangements, methods and procedures in German development cooperation varying to an extent which impedes the development of a common understanding of evaluations, comparable foundations and ultimately a homogeneous evaluation system. As a result, communication and cooperation are still not smooth and effective in the area of evaluations, nor are results particularly comparable.

It is now the declared aim of the BMZ to further develop its own evaluations and its implementing agencies into a complementary, consistent

German evaluation system marked by a division of labour and in addition to strengthen harmonisation with the organisations of civil society in the evaluation area. Specific steps include the establishment of a study group called the “Joined-up Evaluation Working Group”, the issue of new guidelines for technical and financial cooperation, guidance for action under preparation in response to the revised administrative regulations for the Federal Budget Code, the dialogue which has been commenced with the political foundations and churches as well as annual meetings held by the persons in charge of evaluations at German development cooperation agencies.

Efforts have in the meantime had an impact on the principles of evaluation of the development cooperation agencies under review. The majority of governmental as well as non-governmental development cooperation agencies make explicit reference to the DAC principles as well as the DeGEval evaluation standards. However, some actors hold certain reservations with regard to some of these standards. The political foundations and the DEG (Deutsche Investitions- und Entwicklungsgesellschaft), a development finance institution, for instance, consider the DAC Principles to only be suited to a limited extent for their specific activities, among other things due to the demand for transparency. Instead, the DEG feels obligated to uphold the MDB-ECG standards.<sup>2</sup> The problems which development cooperation agencies have in implementing the DAC and DeGEval principles in a concrete manner in their actual evaluations cannot be overlooked, however. Numerous organisations still have difficulties in this respect. The same also applies e.g. to independence.

Although independence and impartiality have been considerably strengthened in the evaluation systems of the development cooperation agencies examined in the study since the last System Analy-

2 “Multilateral Development Banks – Evaluation Cooperation Group – Good Practice Standards for Evaluation”

sis, deficits remain as a result of (1) lack of institutionalisation of the evaluation function, (2) limits to the evaluation powers and authorities of those persons in charge of the evaluations and (3) insufficient or inadequate efforts to assure the quality of impartial, external expertise.

The institutionalisation of the evaluation function has made further progress since the last System Analysis. Two-thirds of the organisations have an evaluation unit or at least an evaluation officer who is organisationally separate from the operational areas in accordance with DAC requirements. In nine cases they have central positions which provide them with direct access to the management, as is expected by the DAC. These organisations have as a result met a key precondition for credible evaluations. In no other German development cooperation agency the institutional independence of evaluations reached a degree comparable to that of the KfW Entwicklungsbank. One-third of development cooperation agencies, on the other hand, have not achieved any basic institutionalisation of the evaluation function. Here there is a real danger of a mere pretence of independence, insufficient evaluation powers and authorities, limited quality assurance and evaluation free zones.

In eight out of 13 organisations the evaluation units or officers are limited in their planning, execution and quality assurance powers, in some cases severely. These are serious weak points in the evaluation systems under review. Especially in highly decentralised organisations, the evaluation units and officers are de facto unable to ensure quality assurance in evaluations which are supervised by operational units, or are only able to do so insufficiently.

The organisations do not differ in their perception of the necessity and usefulness of independent evaluations in principal, but rather in terms of the balance between the two forms. A number of organisations attach far too little importance

to independent evaluations. Here the potential for innovative ideas from neutral third parties is not taken advantage of enough, thereby nourishing fundamental doubts regarding the credibility of evaluation work. On the whole, however, the percentage of independent evaluations has grown significantly since the last System Analysis.

Doubt was once again expressed regarding the independence and impartiality of the evaluators used by the German development cooperation agencies. Indeed, there is a strong tendency especially on the part of decentralised German development cooperation agencies to shift responsibility for evaluations to the operational level. This accordingly creates a latitude which allows the selection of evaluators to be guided by the interests of the respective units and fails to preclude attempts at exercising influence on the evaluation report. There are gaps in quality assurance, although it would be going too far to forward a blanket accusation here. The independence of evaluators from their commissioning agencies is generally assured by granting the evaluators unrestricted responsibility for the report in accordance with the DAC recommendation. The KfW constitutes a notable exception here, however.

In general, the organisations in the study use a broad spectrum of evaluation forms. Naturally enough they are strongly oriented towards the project cycle for the purpose of project steering. At least half of the development cooperation agencies do not make full use of the potential offered to them by other forms of evaluation, however. Although interest in empirical comparative, sectoral, topical and instrument-related evaluations has grown among governmental implementing organisations and civil society organisations, such evaluations have – with a few exceptions – yet to take on any major importance. The fact that final and in particular ex post evaluations have hardly been able to gain currency in Germany so far is not primarily due to financial factors. It is a result, rather, of views on its useful-



ness. Because the German development cooperation evaluation system is focused above all on instrumental learning, forms of evaluation which tend to serve the purpose of conceptual learning have not received sufficient acceptance to date.

Development cooperation agencies, on the other hand, point out the problems they have funding such evaluations. Ensuring sufficient resources is indeed a key determinant for the independence and credibility of evaluation systems. There is a need for action here. A number of evaluation units and officers are not personally endowed with the resources and skills they need to ensure the necessary improvement in quality assurance in evaluation work, nor are they equipped to perform the more frequent impact-oriented, multi-project, ex post evaluations or make increased use of external evaluators.

On the whole, most evaluation systems are underfunded. Even if there is no hard, reliable data or standards, there is evidence that a benchmark of one per cent of annual expenditures on projects and programmes for independent evaluations is adequate. The level attained by most development cooperation agencies is, however, in the realm of lower tenths of one per cent.

If one considers evaluation systems to above all attain credibility when they have a clear evaluation policy while guaranteeing independence and impartiality, when they are based on standards, perform sufficiently frequent, diversified evaluations and are outfitted with adequate resources, then one can also attest to a considerable improvement on the part of the German development cooperation agencies in the study since the last System Analysis, especially as a result of the institutional strengthening of independence and the resolute orientation towards standards. The limited powers and authorities of evaluation units and officers, the failure to take advantage of various forms of evaluations and the insufficient resources available provide ample room for a fur-

ther improvement in credibility, however. If one takes into account that the two largest German development cooperation agencies, the KfW and the GTZ (Deutsche Gesellschaft für Technische Zusammenarbeit), alone account for over 50 per cent of the BMZ's bilateral development cooperation and the fact that their evaluation systems can be assessed as good and trustworthy in considering the overall balance of strengths and weaknesses, then German development cooperation agencies would appear from this perspective to offer a positive picture.

### **Usefulness**

Evaluations are especially used for direct project/programme management in German development cooperation. Other evaluation objectives such as accountability (legitimation and presentation to the outside world) and controls are clearly of secondary importance and have steadily lost significance over time – even at the BMZ, the ministry in charge of policy-making in this area. “Learning” is considered to be the dominant motive underlying evaluations by all of the development cooperation agencies studied. The results of evaluations are for the most part used only by the direct stakeholders, and frequently even by those responsible for projects/programmes. Convincing implementation mechanisms are lacking at many development cooperation agencies. Scarcely any organisation has an evaluation process which ends in a “management response” in which there are arrangements spelling out which evaluation recommendations are to be implemented over what periods of time and which ones are to be rejected and for what reasons. A formalised implementation monitoring system which is used to review the implementation process can scarcely be said to exist. In those cases where it has been implemented (for instance by the BMZ), the resources needed for determined assessment and sanctions in the case of non-implementation are lacking.

Even though institutional learning is deemed to be quite important by many of the evaluated development cooperation agencies, they do not appear to be sufficiently successful in implementing this – in spite of elaborate knowledge-management systems (for example, at the GTZ). A particularly interesting approach is the “delegates system” of the KfW. Staff is assigned to evaluations in order to generate broad-ranged knowledge which not only expands awareness of the benefits offered by the results of evaluations from “other” programmes and projects – the results of evaluations can also continue to be used by these organisations themselves.

An important underlying factor in the low level of use made of evaluations is also the form of evaluation. When almost all evaluations (except for those of the BMZ) recur in specific situational context conditions, it is scarcely surprising when there is doubt as to how the findings generated by such evaluations can once again be used for an organisation’s “own” projects implemented under very specific conditions.

The external use of evaluation results is above all limited by the lack of transparency. Only the evaluation reports of the BMZ have been made publicly available in unabridged form for several years. This not only has a negative effect on the integrity of development cooperation work – the quality of evaluation reports also suffer, as they are insulated from any and all external criticism and expert discussion whatsoever as a result.

As a result of the overall low level of involvement of partner organisations in the development process (except for the church organisations), it is the learning of donor organisations which stands at the forefront. Partner organisations which are moreover frequently not involved in the planning of evaluations, the formulation of terms of reference or the selection of evaluators, etc., and in many cases do not even receive the complete evaluation reports for them to read and are

instead only informed about the initial (and of course provisional) results of an evaluation at a debriefing after the evaluation is completed and receive a brief summary of the findings at a later point in time will probably be unable to profit much from the evaluations which have been performed.

### Quality

Most German development cooperation agencies make considerable efforts to improve the quality of their evaluation systems. They take the following factors into account in differing degrees, however:

In addition to the DAC Principles and Criteria and the DAC Quality Standards many of the development cooperation agencies in the study also attempt to use the DeGEval evaluation standards.

The orientation towards results, which is supported by all of the organisations in the study, is one of the biggest changes which have taken place since the last System Analysis. For this reason, development policy as a policy field is playing a leadership role for all others which continue to practice an input or output approach. Among the governmental organisations, the GTZ has executed this reorientation most resolutely. The Welthungerhilfe (DWHH) and Misereor stand out in the non-governmental area. Some political foundations (with the exception of the Konrad-Adenauer-Foundation and the Heinrich Böll Foundation) are among the organisations which have least addressed the question of what an orientation toward impact means for the evaluation of an organisation, citing their special role as the reason.

While on the one hand the orientation of development cooperation towards impact and its evaluation must be characterised as path-breaking, it can be criticised on the other hand that the designs and methods used to measure impact and

the imputation of the causal factors at work are still far from meeting the professional standards needed for reasonable quality. With few exceptions (above all the BMZ), most of the development cooperation agencies in the study are conducting evaluations along the same lines as they have in the past. The selection of methods, the time frames and the amount of funding are generally not adequate to the demanding task. The search for rigorous methods which can be used in regular operations has only begun.

In comparison to the last System Analysis, it would appear that only two of the development cooperation agencies in the study have expanded their evaluation units on a significant scale (the GTZ and the KfW). In spite of the immense growth in the number of evaluation tasks, the number of staff working at the division in charge at the BMZ has stagnated. The ministry ranges towards the very bottom in international comparison. As a result, it is running the danger of no longer being able to perform key evaluation tasks and its steering function in the evaluation area. Measured in terms of the budgets of governmental as well as non-governmental development cooperation agencies, there appears to be a large degree of heterogeneity in the human resources assigned to evaluation tasks. The amount of staff in evaluation departments does not correlate directly with the size of an organisation or its development cooperation budget. It is apparently more the importance which an organisation attaches to evaluations which matters. Many organisations – with the BMZ heading the list – evidently need to boost the number of staff considerably.

With regard to the competence of the human resources used, it can be criticised that scarcely any organisation hires experienced or even trained evaluators. It is experience in development cooperation, rather, which counts. Because most of the organisations rotate staff, the level of competence which is present in a department is subject to constant change. Staff that have become

“veterans” in evaluations over time and are thus familiar with the requirements and tasks of evaluations are constantly replaced by newcomers. Because there is scarcely time for systematic additional training prior to taking up new positions and all too often there is hardly any time available afterwards, either, it may take a long time before a high level of competence is attained. National and international programmes of vocational education and advanced training are not taken advantage of on a sufficient scale by many development cooperation agencies.

Most of the development cooperation agencies also use external experts in their evaluations. Importance is attached to a large number of qualifications in the recruitment of staff. Profound knowledge of methods and evaluations are only explicitly required by very few development cooperation agencies, however. Instead, value is attached to experience in development cooperation and evaluations.

Almost all organisations make use of the expertise of external experts, although without always assigning them complete responsibility for reports (this is the case, for example, with the KfW). Frequently no quality assurance of evaluators' work is carried out by evaluation units, particularly at those organisations which are highly decentralised. Because most of the evaluation contracts only involve limited fees, it is possible to freely award them to individual evaluators without being subject to any major accountability requirements. It would be beneficial to the quality of the system, however, if tenders were competitive. The GTZ and at times the BMZ are increasingly opting for this approach.

The quality of processes, which is to say the quality of planning and execution processes, constitutes an additional important feature in a high-quality evaluation system. The study has also produced a pretty heterogeneous picture in this regard as well. The majority of the evaluations per-

formed by the organisations evaluated are based on project progress and its needs, and are frequently initiated by operational units. With the exception of the KfW, multi-project and ex post evaluations are generally subject to discretionary decisions and a selective “picking and choosing” – hence, they are not representative. One positive aspect which should be noted is that almost all the organisations have instruction manual guides, working aids and forms to help structure the planning and procedural process. The fact that quality assurance is very splintered in many organisations (and this also applies to the review of adherence to rules) has a negative impact, especially when the evaluation is decentrally organised, which is usually the case. Only at the BMZ the evaluation division is in charge of execution and quality assurance of all evaluations (this applies to a large extent to the KfW and in part to the GTZ as well). In the case of non-governmental organisations, however, the operational units themselves are responsible for the quality assurance of evaluations. This however makes it impossible for quality assurance to be carried out on a uniform basis.

The development cooperation agencies hardly have any hard data on the frequency of evaluations at their disposal. This is in part due to the decentralised nature of evaluations at many organisations. The majority of all evaluations carried out (with the exception of the BMZ) relate to project progress and serve the immediate purpose of steering the project. Final and in particular ex post evaluations have not gained currency in German development cooperation to date (with the exception of the KfW and the GTZ). Because the German development cooperation evaluation system is above all focused on the instrumental learning, evaluation forms which are more for the purpose of conceptual learning have to date not met with sufficient acceptance.

In general the German development cooperation evaluation system suffers from a lack of comparable empirical evaluations both within individual

sectors or countries as well as between development cooperation agencies. These forms of evaluations – which are repeatedly called for by development cooperation experts – would generate considerable useful know-how, but they are difficult and time-consuming to organise.

### Partnership

The implementation of evaluation results depends inter alia on whether and to what extent the most important stakeholders are involved in the planning and reporting phases. This study presents a sobering picture in this respect: no significant changes are apparent in comparison with previous system studies. Merely the church agencies MISEOER, Church Development Service (EED), Caritas International as well as Welthungerhilfe attach value to the tight involvement of partner organisations. The KfW and most of the political foundations represent the other side of the spectrum. They scarcely involve their partner organisations, arguing that they are the subject of the evaluation. This does not explain, however, why as a rule no independent domestic (local) evaluators are used, as is usually the case at most of the other governmental and non-governmental development cooperation agencies.

Except for the church and a few other non-governmental organisations, evaluations in the field of development cooperation are donor-centred activities which above all serve the needs and interest of the German development cooperation agencies. As a result, partner organisations are often robbed of the possibility of becoming actively involved in evaluations and using the results of evaluations in their own work. For this reason it is scarcely surprising that only those development cooperation agencies which involve their partner organisations in evaluations report that these are interested in evaluations. The organisations which do not involve their partners stated in the study that their partners were not very interested in evaluations.

To arouse an interest it is not only necessary to inform and make partners aware of the usefulness of evaluations – it is also important to actively involve them in the planning and execution of evaluations and to involve local evaluators. This is hampered by the lack of competence in many countries down to the present, however. Internal and external domestic evaluation staff not only lacks experience – they also lack appropriate qualifications. In spite of this obvious deficiency, only very few development cooperation agencies support their partner organisations (as for example Capacity Building International (InWEnt) does) through systematic programmes in the area of vocational education and training.

### **Networking, system formation and change**

Even though a significant intensification of the level of networking is apparent in comparison to the last System Analysis, above all due to various activities of the BMZ, many of the development cooperation agencies in the study are only insufficiently linked with one another in the area of evaluation. This is not only related to the size of an organisation. While the BMZ and the GTZ have a high level of international and national networking, respectively, and the GTZ virtually plays the role of a “lead agent” at many other development cooperation agencies, the same does not apply, for instance, to the KfW. Moreover, in contrast to the BMZ and the GTZ, the KfW had scarcely worked with universities or scholarly research institutes down to October 2007. This applies to an even greater extent to all the other development cooperation agencies in the study, which at best only make use of research in individual cases. In this context, it is worth noting the practice at the KfW, however, where the directors of the evaluation unit are highly acclaimed persons from the field of scholarly research. This

also helps promote a transfer from research to the field of practice.

The most important impetus for communication over the last few years has come from the BMZ and the “Evaluation of Development Policy” study group at DeGEval. The aim of the BMZ is to develop a master plan for the overall structure of evaluation in development cooperation specifying the role of the individual actors. The BMZ has initiated an annual meeting of evaluation units which, however, merely promotes a comparison of experience and does not produce any concrete working results. The latter is the task of the “Joined-up Evaluation” working group, which was set up by the BMZ and whose members include the KfW, the GTZ, InWEnt and the German Development Service (DED) (but not the governmental implementing organisations DEG<sup>3</sup>, PTB (National Metrology Institute) and BGR (Federal Institute for Geosciences and Natural Resources)). The aim is to develop a “common language”, a mode of procedure, quality criteria and to issue a joint evaluation report of the implementing organisations. Remarkable results have already been produced along these lines, but the process has proven to be extremely time-consuming and tedious. In a second thrust, the BMZ has been instituting a process aimed at harmonising performance assessments and evaluations in the area of civil society since 2006.

The establishment of a Study Group for Evaluations at the DeGEval, which goes back to a private initiative, is the only platform where contractors, governmental and non-governmental commissioning agencies, researchers and interested experts meet on a regular basis to discuss current topics. Moreover, it is also a voluntary amalgamation of actors who try to come up with answers to questions relating to evaluations in working

<sup>3</sup> The DEG is a market-oriented development financing institute for the private sector engaged in the self-funding business and not an implementation organisation. Hence, any “development cooperation strategy or harmonised evaluation” is not binding on the DEG.

groups (e.g. working groups for “transparency, information flow and follow-up in evaluation processes”, “learning from evaluations”, “evaluation to foster and promote democracy” and “evaluation of impact”).

Even though the networking of development cooperation agencies has made progress as a result of this initiative, it is nevertheless still a far cry from constituting a consistent German evaluation system for development cooperation. The study supplies abundant evidence demonstrating this.

Although the working committees and study groups initiated by the BMZ, which have been constituted on a voluntary basis, are necessary pioneers in a growing network or even in a future harmonised evaluation system, this is by no means sufficient – as developments over the past few years show. The particularistic interests of the individual development cooperation actors are too strong for this.

The BMZ’s determined and energetic attempt to provide the German development cooperation evaluation system a direction and lend it a voice in the international arena should not obscure the fact that the overall system has developed in a highly skewed direction. The roots of this problem are not to be found in the evaluation area, however, but are rather a result of the considerable expansion of staff resources by the implementing agencies while the number of staff working at the BMZ has stagnated. This has led to a general steering deficit, and must be criticised as one of the fundamental problems facing German development cooperation. This development, which can be witnessed through the entire development cooperation area, is reinforced, however, by the fact that in particular the GTZ and the KfW have successfully perceived their evaluation tasks, as they have been repeatedly called upon to do, and have accordingly acted to staff their central departments with adequate human resources while at the same time assigning them a high de-

gree of independence. These divisions are now attempting to defend and preserve their privileges against internal and external encroachment (including by the BMZ). As a result, the coordination process has begun to falter, as the BMZ is clinging to the status quo while the implementing agencies have been increasing their budgets, beefing up their staff resources and enhancing the status of their evaluation departments. In spite of more demanding international requirements, evaluations which have become on the whole more complex and demanding, and the effort to create a common foundation for the German development cooperation evaluation system, the BMZ division in charge has not seen an increase in its staff resources or its budget on any sizeable scale in over a decade (in spite of its being transformed into a central unit), nor has its status been enhanced within the BMZ. Especially because the governmental implementing organisations, and even many NGOs, have been expanding their evaluation units, the BMZ’s Evaluation and Audit Division has come under increasing pressure, as it is no longer able to adequately steer this dynamic development with its stagnating resources. The skewed situation which has come about in the evaluation area of development cooperation as well as the steering deficit at the BMZ’s Evaluation and Audit Division are actually the result of successful change – which the BMZ’s Evaluation and Audit Division has not been able to keep pace with due to its stagnating financial and human resources. This steering deficit is causing a number of problems which are significantly hampering a consistent (conscious and deliberate) development of the system.

### **3 Recommendations**

#### **System development**

An evaluation system for German development cooperation should be created which features a consistent, balanced structure between the vari-

ous actors themselves, between the BMZ and its implementing organisations and in cooperation with the organisations of civil society – a structure which functions according to uniform rules and clearly spells out the tasks of the individual organisations.

The BMZ's goal of developing a master plan for the overall structure of the evaluation in development cooperation which will coordinate the tasks of the individual actors among each other and strengthen the role of the BMZ thus warrants support. The instruments used to date (setting out rules and regulations, annual meetings, the "Joined Up Evaluation" working group) are not sufficient for this. The BMZ's Evaluation and Audit Division needs to be augmented considerably in terms of its human resources and budget. Its institutional status also needs to be enhanced in order to make its leadership role appear convincing and for it to become more assertive.

Three possible ways of increasing the independence, credibility, usefulness and quality of German development cooperation evaluation systems on the whole are proposed here for further analysis:

- (A) Strengthen the authority and powers of the BMZ's Evaluation and Audit Division
- (B) Establish an independent evaluation agency which advises and supports the BMZ in system development and the development cooperation agencies in the development of instruments and in improving quality. Such an agency should be independent and carry out evaluations, in particular on impact and sustainability of development cooperation, in a manner which allows a comparison of countries, sectors and institutions at a high level of competence and is thus taken seriously.
- (C) Establishment of an independent evalua-

tion advisory council which performs strategic, quality assurance and control functions for the entire development cooperation evaluation system.

All three versions could be combined with one another. An independent commission should review the feasibility of the proposals as quickly as possible in order to take advantage of the pressure for action and the overall positive attitude towards reform prevailing at present.

### **Foundations and objectives**

The development cooperation agencies should close existing gaps in the basic foundations underlying their programmes, organisation and methodologies and adapt to national and international standards, goals and targets. The 1992 DAC Principles should continue to guide all development cooperation agencies in the design of their evaluation systems along with the new 2006 DAC Quality Standards for evaluation systems. The BMZ should actively promote this process by means of a regular dialogue with the various groups of development cooperation agencies and by agreeing on and periodically reviewing sufficiently detailed specifications.

### **Institutionalisation**

To boost the credibility of their evaluation systems, development cooperation agencies which have not made the evaluation function institutionally autonomous should assign independent evaluation units or officers to perform this task. Smaller organisations should also review this possibility; as an alternative, the evaluation function could be outsourced. In the case of larger organisations – including the BMZ – independence could be further strengthened by assigning external evaluators with distinct term-limitations. In addition, rules could be instituted preventing such external evaluators from being employed by the organisation later.

The evaluation units and officers should moreover receive much greater budgets which they are able to use at their sole discretion. Existing restrictions on the planning, execution and quality assurance authority and powers should be eliminated without delay. The highest leadership level should waive its right to issue orders to strengthen the independence of evaluation units and officers.

The evaluation systems should be linked as effectively as possible to quality and knowledge management.

### **Planning and process**

To eliminate existing evaluation deficits, evaluations must explicitly be made mandatory in the planning phase of new development cooperation measures. Responsibilities and frequencies need to be specified and the required financial resources separately included in the cost calculation. A minimum size could be specified for small measures to be evaluated.

Development cooperation agencies which especially emphasise the responsibility (“ownership”) of their partners should not lower the priority of their own interests and their responsibility for evaluations, but rather actively initiate independent evaluations, plan and execute them jointly with their partners and use their results.

Evaluation contracts should be publicly tendered whenever possible. When contracts are freely awarded, competitive public bids and transparent criteria should be used. This would allow a larger market to be reached, avoid dependence on a few evaluators and strengthen competition for quality.

The independence of evaluators must be systematically reviewed and secured through formal declarations issued by the evaluators and the “four-eye principle”.

Independent evaluators are to be assigned unrestricted responsibility for their reports both in formal and de facto terms.

An inception report is to be provided for and should contain a detailed description of the methodological conception.

### **Forms and scope**

Greater advantage should be taken of the potential offered by independent external evaluations. More multi-project and ex post evaluations must be planned and executed in a systematic manner to promote conceptual learning and ensure reliable accountability regarding impact and sustainability of organisations’ own projects vis-à-vis the outside world.

The evaluation systems of most development cooperation agencies are considerably underfunded. The BMZ has the greatest need to catch up here. It is all the more necessary to increase financial resources when evaluations need to face new challenges posed by recent changes in development cooperation and the focus on impact. At least one per cent of the portfolio (annual expenditures on projects and programmes) should be made available for independent evaluations. Because smaller development cooperation agencies have difficulties evaluating on a multi-project and ex post basis, the BMZ should set up an external evaluation facility.

### **Quality**

The conceptual focus of development cooperation on impact is not yet sufficiently reflected in the evaluations which are currently being carried out. The number of ex post evaluations should be boosted significantly in order to be able to review the long-term impact of development cooperation.

Rigorous methods which allow the measurement of impact and the identification of causal factors



based on probabilistic hypotheses in a manner which is as reliable as possible must be developed for regular operations and taken advantage of in a standard manner.

There appears to be a clear increase in more comprehensive evaluations taking place at ever higher levels of aggregation. In the course of this development there will be more and more joint evaluations by different development cooperation agencies and donors at the sectoral, regional and national levels. Moreover, new development cooperation instruments will have to be evaluated as well. This implies new challenges not only as a result of the unavoidable harmonisation, but also due to the need to use suitable evaluation methods. The BMZ and the other large development cooperation agencies should therefore promote the targeted development, the testing and application of appropriate evaluation designs, data collection and analysis methods.

Process quality has already attained a remarkable level among German development cooperation agencies. To further enhance this quality, above all quality assurance needs to be improved. This means that every organisation - depending upon its size - should have an evaluation unit or an evaluation officer who is in charge of quality assurance for all evaluations and who effectively ensures that quality.

Generally speaking, the German evaluation system does not provide for comparative analyses of programmes of development cooperation organisations within a sector or country as well as inter-institutional comparisons between organisations. Such comparative evaluations, which are facilitated by efforts towards joined up evaluations, should be included in the evaluation portfolio, as they hold out the promise of considerable benefits, especially for conceptual learning.

### **Staff**

Many German development cooperation agencies have competent staff working in their evaluation units, but they have often not been adequately prepared for their very specific evaluation tasks. That is why development cooperation agencies should recruit trained and experienced external evaluators for their evaluation units and/or prepare internal personnel carefully and provide advanced training on an ongoing basis. The requirements applying to evaluations on the whole have increased, and the qualification profile of the staff that are to be used in this area must reflect this.

This also applies to the evaluators used in external evaluations, for whom evaluation and methodological competence is becoming ever more important in view of the orientation of development cooperation towards results and the growing complexity of programmes. Greater attention thus needs to be devoted to appropriate further education and training in recruiting evaluators.

### **Utility**

In order to enhance the usefulness of evaluation results, the implementation of recommendations needs to be monitored to a greater degree than in the past. To this end it is first of all necessary for commissioning agencies to provide a management response which states who is responsible for what recommendations during which period of time and what recommendations have not been implemented and why. This implementation plan is to be accompanied by the monitoring of implementation which is as independent as possible. Any deviations from this plan must be reported to the persons in charge. To intensify institutional learning, each new plan should indicate whether and what evaluation results from other evaluations have been taken into account.

Greater transparency is needed to facilitate inter-institutional learning and to improve the quality

and the credibility of evaluations. It is not enough to merely make brief summaries publicly available – the entire evaluation report (with the exception of narrowly delineated data warranting protection) needs to be published.

This will not only allow donor organisations to tap the learning potential offered by evaluations – it will include the partner organisations to a much greater degree than in the past in the evaluation planning, execution and analysis (see the section on the involvement of partners).

### **Partners**

To arouse the interest of partners in evaluations, it is also necessary to actively involve them in the planning and execution of evaluations and enlist the support of local evaluators while making an effort to inform and make them aware of the benefits of evaluation.

Because there is a lack of internal (that is, working within the partner organisations) as well as exter-

nal trained evaluation experts, further education and training should be promoted in the field of evaluation, e.g. by increasing the number of programmes offered by development cooperation agencies such as InWEnt.

### **Networking**

The level of networking among German development cooperation agencies in the area of evaluation with each other as well as with international organisations and research has picked up considerably over the last few years, but is still in need of improvement. The potential offered by research in particular needs to be more effectively used by the development cooperation agencies.

To live up to the Paris Declaration and other international agenda and respective commitments, joint evaluations should increasingly be carried out by the German development cooperation agencies together as well as with other international donors. Clear institutional arrangements should be devised to this end.

**Comments of BMZ's Management**

BMZ welcomes this candid report explicitly. We take our responsibility to account for the money spent for development cooperation seriously: Taxpayers in Germany and recipients in partner countries need to know whether development aid is used effectively and efficiently. Only transparent and high-capacity evaluations using comparable metrics can assure this.

The main thrust of the report confirms our line of thinking about a coherent system-wide evaluation approach including in particular our implementing agencies. Strengthening the role and the capacities at the development policy level is a prerequisite. The report points to solutions which merit further in-depth consideration. We will explore the feasibility of implementing one or more of the options presented as main recommendations.

The study has triggered important changes among the agencies reviewed already. With a view to harmonisation within the German development cooperation system, the evaluation and audit division of BMZ has been requested to draft an evaluation policy and guidelines pertaining to all areas of BMZ's development cooperation. We are aware that this is only a beginning.

**Published by the**

Federal Ministry for Economic  
Cooperation and Development

**Bonn Office**

Dahlmannstraße 4  
53113 Bonn  
Germany

Phone +49 (0) 228 99 535 - 0

Fax +49 (0) 228 99 535 - 3500

**Berlin Office**

Stresemannstraße 94  
10963 Berlin  
Germany

Phone +49 (0) 30 18 535 - 0

Fax +49 (0) 30 18 535 - 2501

[poststelle@bmz.bund.de](mailto:poststelle@bmz.bund.de)

[www.bmz.de](http://www.bmz.de)

**Editors-in-chief**

Michaela Zintl

*Division 120: Evaluation of development cooperation; auditing*

Jutta Wagner

*Division 114: Development Education and Information*

**Official responsible**

Michaela Zintl

*Division 120: Evaluation of development cooperation; auditing*

Klaus Krämer

*Division 114: Development education and information*

**Design and layout**

Typo-Druck GmbH, Bonn

As at

July 2009

