

“SYSTEM REVIEW” – A CATALYST FOR BETTER EVALUATION PRACTICE IN GERMAN DEVELOPMENT COOPERATION?

Summary

Robust findings on the effectiveness of German development cooperation are of central importance – whether for reasons of accountability, for improving the management of project and programmes, or for securing conceptual lessons learnt. Along with results-oriented planning and monitoring, independent evaluation is considered by the OECD/DAC donor community to be the instrument of choice for acquiring vital knowledge.

In seeking to assess the current status of evaluation practised in the field of German development cooperation, the Federal Ministry for Economic Cooperation and Development (BMZ) conducted a system review (Systemprüfung) at the beginning of 2007 (cf. Borrmann and Stockmann, 2009). The system review resulted in the formulation of specific recommendations for action to help development organisations in the field of German development cooperation improve their internal evaluation practices. Additionally, the review provided system-level recommendations such as the founding of an independent evaluation institute. This recommendation in particular was put into practice in 2012 with the launch of the German Institute for Development Evaluation (DEval).

The consequences of the first system review of evaluation practice in German development cooperation, as well as wider developments in the evaluation field, were then studied by DEval in 2014. The following conclusions were drawn:

- The system review has proved to be a valuable instrument in itself. Both in the organisations examined and at the system level, this instrument served as an important catalyst for advancing evaluation practices. Thus, it is recommended that such reviews of the evaluation system be continued in the future.
- Considerable progress has been made in institutionalising evaluation practice and there is greater receptiveness to the importance of evaluation. In the relevant organisations, not only are these changes reflected in increased (human) resourcing of evaluation but structures and processes for evaluation have also been strengthened (including increased application of evaluation guidelines and plans, standardised processes and an enhanced institutional status of evaluation units).
- For a majority of the development organisations surveyed, the most important objectives of evaluations are project improvement and (institutional) learning. Unfortunately, the knowledge gained from an evaluation frequently remains within the immediate circle of project stakeholders, and it is still rare for evaluation reports to be published in full. In order to achieve higher potentials for conceptual and institutional learning, more must be done to create a culture of “transparent learning”.
- A central challenge for many evaluations continues to be the failure to apply sophisticated methods designed in line with clear impact hypotheses. Development cooperation organisations are, however, showing a strong interest in training opportunities in order to address this problem.
- There is frequently room for improvement when it comes to involving partners in the evaluation process in a systematic manner. On this point, the study found considerable differences between the development cooperation organisations surveyed.
- The development cooperation activities within the portfolio of the BMZ are in fact subject to comprehensive evaluation. Such a high range of evaluation is not found in other German ministries delivering Official Development Assistance (ODA), nor in other policy fields.

Evaluation practices in German development cooperation

Whether considering the Millennium Development Goals, the Paris Declaration, or the 2030 Agenda for Sustainable Development, ambitious global targets draw attention to the need for effective, sustainable, and coherent development policy. Evaluations supply evidence-based knowledge not only critically important for targeting future development policy measures but also for providing accountability towards the usage of taxpayers' money.

It is in this context that evaluation has become increasingly important – all the more so given that an effective common framework for monitoring and evaluation (M&E) must still be negotiated for the 2030 Agenda. But how far has evaluation developed in the field of German development cooperation and what is the quality of the knowledge now being generated? This question was examined as part of a system review commissioned by the BMZ in 2007 and based on case studies of both governmental and non-governmental development cooperation organisations (cf. information box). In this review it became clear that the heterogeneity of German development cooperation organisations is also reflected in their varying systems of evaluation. Thus, a key system-level recommendation was to set up an independent evaluation institute. The purpose of this agency was to support the BMZ in strategic decision-making by providing independent evaluations of the effectiveness, efficiency, and sustainability of development cooperation (Borrmann and Stockmann, 2009: 17, 65, 165). At the same time, conceptual and interinstitutional learning at the system level should also be promoted through this agency. The BMZ acted on this recommendation by founding DEval, the German Institute for Development Evaluation.

The system review also shed light on how the various development cooperation organisations evaluate their interventions funded by BMZ. It found that there are many aspects determining how useful and meaningful evaluation findings can be. Key differences lie in the objectives set for the evaluation (e.g. learning lessons or providing accountability), in the form of the evaluations conducted (e.g. mid-term, or ex-post evaluation, project or cross-sectional evaluation), as well as in the designs and methods used. These are also the characteristics which mainly determine how evidence-based and useful the results will be. Additionally, evaluation results and their uses are influenced by the available resourcing, as well as by structures and processes (such as the independence of the evaluation team), the formats for institutional learning from evaluation, and the integration of partners.

DEval has surveyed the current state of evaluation practices across the German development cooperation field and identified changes that have occurred since the last system review. In addition to an online questionnaire survey of 34 development cooperation organisations, interviews were conducted with those organisations which took part in the earlier system review. DEval also undertook a meta-evaluation of 45 evaluation reports.

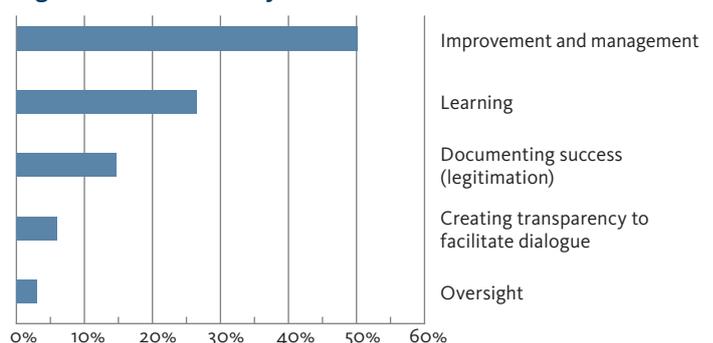
“System reviews” in German development cooperation

In 1997/1998, the BMZ commissioned, for the first time, a systematic study of evaluation practices in German development cooperation (Borrmann et al., 1999). This was followed by a second study in 2007/2008 (Borrmann and Stockmann, 2009). Based on twenty individual case studies, the study provided a number of recommendations for action to advance evaluation practice, some concerning the system as a whole and others specific to certain organisations. This led to agreements between the BMZ and the relevant organisations on steps to be taken to implement the recommendations accepted by those organisations.

Current state of evaluation practice and the impetus gained from the system review

The majority of the surveyed development cooperation organisations report that the most important objectives of evaluations are to improve projects and project management as well as increase learning. Fewer respondents mentioned other objectives, such as legitimising measures taken, creating transparency, or exercising oversight (cf. diagram on page 3; also Lücking, Freund and Bettighofer, 2015, and Borrmann and Stockmann, 2009: 130).

Figure 1: Evaluation objectives



Source: online survey DEval 2014; single choice selection from list, n=34

Additionally, the organisations are subject to evaluation obligations towards the BMZ. In practice, the results of these obligatory project evaluations are mainly used by those directly involved in the project and are fed into programme or project management. In order to foster conceptual learning that goes beyond a specific project there has been a stronger focus on cross-sectional evaluations since the last system review.

The oversight function of evaluation can particularly lead stakeholders to be resistant to taking part in an evaluation. However, compared to the situation recorded by the first system review, there have been improvements in this area, with those working on a project showing greater openness to the benefits of evaluation. They now tend to regard evaluations as a part of their own professional responsibility, which can be seen specifically in growing demand for evaluations. Acceptance is also reflected in structural changes such as increased financial and human resources allocated to evaluations since the system review.

The growing institutionalisation of evaluation is expressed by an intensified introduction of evaluation guidelines which are increasingly aligned to established national and international standards, such as those set by the Gesellschaft für Evaluation (DeGEval) or the Development Assistance Committee of the Organisation for Economic Cooperation and Development (OECD/DAC). The study also reveals an intensification of dialogue among development cooperation organisations on issues of evaluation and records that this exchange is regarded by all sides as being very helpful. Nevertheless, much more can be done to intensify exchange with academic and research institutions on evaluation methods and designs.

What is the potential for advancing evaluation practices in German development cooperation?

In addition to the positive developments described above, there are some potential areas for further progress:

The original system review criticised the fact that too few cross-project, impact and ex-post evaluations were being conducted. Although singular project evaluations continue to make up the majority of evaluations being performed, the current study finds that there has mostly been a rise in the number of cross-sectional evaluations. A factor impeding similar increases in ex-post and impact evaluations is the lack of available financial and human resources. Moreover, doubts are raised regarding the value added by these types of evaluation and the capacity for applying robust

methods, especially in light of the resources required to apply these kinds of methods.

Although it is vital to draw reliable conclusions about the effectiveness of development interventions, the evaluation reports analysed by the DEval study reveal that evaluation designs do not include operationalised impact hypotheses which can be further verified. Another problem, already criticised in the system review, is a failure to use the full spectrum of potential evaluation methods. For the most part, evaluations rely on qualitative document analyses and interviews. About half of the reports examined were based solely on one of these methods or a combination of the two. Far less used were quantitative surveys and analyses, which for example are based on standardised questionnaires. In this regard, more use should be made of mixed-methods approaches and an explicit triangulation of data drawn from a number of sources. (Quasi-) experimental designs continue to be a rarity.

Financial limitations are partly responsible for preventing more sophisticated designs and methods. Another factor is that consultants brought in for an evaluation are often chosen mainly for their know-how in a specialist field or country context. Therefore, evaluations are frequently weakened by a lack of expertise in robust qualitative and, above all, quantitative methods of evaluation. Respondents in the study often expressed an interest in training in more sophisticated methods.

Cross-project conceptual learning is regarded by respondents as an important evaluation objective (cf. diagram on page 3), and the study records an increasing use of cross-sectional evaluations. However, evaluation processes are often not systematically aligned to the objective of institutional learning. Lessons learned are all too often shared only among the immediate stakeholders of the evaluated project or programme in question – and even at this level the lessons are often not widely shared, as seen by insufficient sharing of evaluation reports with partners and a lack of partner involvement in evaluation processes. Interinstitutional or collaborative learning is also hindered by the fact that evaluation reports are rarely published in full. Due to interests in protecting partners or out of concern that donations may be threatened it is common practice to publish only summary versions of the evaluations conducted. From the perspective of an individual organisation, the sharing of evaluation results therefore remains both internally and externally limited. This can restrict existing opportunities to learn lessons from each other and inhibit the widest possible use of evaluation findings.

Conclusion

The system review (Systemprüfung) has strengthened the evaluation system of German development cooperation. Its recommendations have been driving forward discussion and sharing of the benefits and uses of evaluation.

In addition to the founding of an independent evaluation institute designed to strengthen evaluation capacities for complex and strategic evaluations and improve methods and standards, there has been expansion and improvement of evaluation practice within many development cooperation organisations. Nevertheless, room for improvement remains in a number of areas:

- In order to promote evidence-based conceptual learning, the use of impact, cross-project, and ex-post evaluations must continue to be strengthened (financially as well as methodologically). To this end, a stronger focus is needed on evaluation design and applied methods.
- With regard to conflicting objectives that can arise in evaluations, it seems advisable to create a politically accepted culture of “transparent learning” in which critical evaluation findings are published and discussed more widely.
- Comparatively speaking, the BMZ portfolio is extensively evaluated. This is an example that could be followed by other governmental institutions providing ODA, as well as those in other policy fields.
- It seems reasonable to conduct further reviews of the evaluation system at appropriate intervals of time in order to track ongoing developments or to instigate new ones.

References

Borrmann, A., K. Fasbender, M. Holthus, A. von Gleich, B. Reichl and R. Shams (1999), *Erfolgskontrolle in der deutschen Entwicklungszusammenarbeit. Studie im Auftrag des Bundesministeriums für wirtschaftliche Zusammenarbeit und Entwicklung*, NOMOS-Verlagsgesellschaft, Baden-Baden.

Borrmann, A. and R. Stockmann (2009), *Evaluation in der deutschen Entwicklungszusammenarbeit*, Volume 1 – Systemanalyse, Waxmann, Münster.

Lücking, K., S. Freund and S. Bettighofer (2015), *Evaluierungspraxis in der deutschen Entwicklungszusammenarbeit. Umsetzungsmonitoring der letzten Systemprüfung und Charakterisierung wesentlicher Elemente*, Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), Bonn.

Authors



Dr Kim Lücking
Senior Evaluator, DEval



Simon Bettighofer
Former evaluator, DEval



Simon Freund
Evaluator, DEval

The German Institute for Development Evaluation (DEval) is mandated by the German Federal Ministry for Economic Cooperation and Development (BMZ) to independently analyse and assess German development cooperation interventions. Evaluation reports contribute to the transparency of development results and provide policy-makers with evidence and lessons learned, based on which they can shape and improve their development policies.