



# SUSTAINABILITY IN GERMAN DEVELOPMENT COOPERATION

*Evaluation synthesis*

*Executive Summary*

2018



---

**DEval**  
GERMAN  
INSTITUTE FOR  
DEVELOPMENT  
EVALUATION

---

## Imprint

### Published by

German Institute for Development Evaluation (DEval)  
Fritz-Schäfer-Straße 26  
53113 Bonn, Germany

Tel: +49 (0)228 33 69 07-0

Email: [info@DEval.org](mailto:info@DEval.org)  
[www.DEval.org](http://www.DEval.org)

### Authors

Dr. Martin Noltze  
Dr. Michael Euler  
Ida Verspohl

### Responsible

Prof. Dr. Jörg Faust (until June 2016)  
Dr. Sven Harten (since June 2016)

### Design

MedienMélange:Kommunikation!, Hamburg  
[www.medienmelange.de](http://www.medienmelange.de)

### Translation

Dr. John Cochrane

### Photo credits

Gui Yongnian/123rf.com (Cover), FO Travel/Alamy Stock Foto (Chap. 1), themacx/iStock.com (Chap. 2), Nikon'as/Fotolia.com (Chap. 3), epicurean/iStock.com (Chap. 4 + 6), andresr/iStock.com (Chap. 5), Steve Bloom Images/Alamy Stock Foto (Chap. 7)

### Bibliographical reference

Noltze, M., M. Euler and I. Verspohl (2018), *Evaluation synthesis of sustainability in German development cooperation*, German Institute for Development Evaluation (DEval), Bonn

### Printing

Bonifatius,  
Paderborn



© Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), January 2018

ISBN 978-3-96126-065-2 (print)  
ISBN 978-3-96126-066-9 (PDF)

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.

The Institute's evaluation reports contribute to the transparency of development results and provide policymakers with evidence and lessons learned, based on which they can shape and improve their development policies.

This report can be downloaded as a PDF file from the DEval website:  
[www.deval.org/en/evaluation-reports.html](http://www.deval.org/en/evaluation-reports.html)

Requests for print copies of this report should be sent to:  
[info@DEval.org](mailto:info@DEval.org)

# EXECUTIVE SUMMARY

## **Background, purpose and object of the evaluation**

The 2030 Agenda for Sustainable Development makes sustainability the guiding principle for global action by humankind. The Sustainable Development Goals (SDGs) defined in the 2030 Agenda combine economic progress with social justice and the sound management of environmental resources. Responsibility for implementing the 2030 Agenda rests with all countries. At the same time, implementation requires new arrangements for cooperation between governments, the private sector, the scientific and academic community, and civil society.

The international development cooperation community has also pledged to reorient its approach accordingly. In the future, the design and implementation of development cooperation must comply with the goals and principles of the 2030 Agenda. This is a key challenge for international development cooperation. At the level of individual projects, it requires planners to reflect in particular on social, economic and environmental interactions, and effects on disadvantaged groups. To support this process, evidence-based recommendations are required. Currently there are only a limited number of projects that were designed explicitly in line with the 2030 Agenda and its principles. Nonetheless it is possible to study the sustainability of development cooperation projects empirically.

In evaluations of German development cooperation projects, sustainability has been systematically assessed since 2006. In that year the Federal Ministry for Economic Cooperation and Development (BMZ) published its 'Evaluation criteria for German bilateral development cooperation. A guideline for evaluations performed by the BMZ and the implementing organisations'. Based on the Principles for Evaluation of Development Assistance adopted by the Development Assistance Committee (DAC) of the Organisation for Economic Cooperation and Development (OECD) in 1991, this guideline contains instructions on assessing the evaluation criteria relevance, effectiveness, efficiency, impact and sustainability. Pursuant to the guideline, the sustainability of specific projects is assessed using mandatory key questions. The outcome of this assessment is the award of a sustainability score. Conceptually, the sustainability of projects is assessed in close conjunction with impact. It is therefore to be expected that

evaluation practice to date – through the criterion of impact – already covers several of the principles of the 2030 Agenda.

The evaluation synthesis conducted here aims to better understand the interactions between various determinants when assessing the sustainability of projects. The purpose of the study is to help better align the strategic and operational orientation of German development cooperation with the new requirements of the modern understanding of sustainability contained in the 2030 Agenda. This is in response to the increased importance of sustainability when evaluating German development cooperation projects in conformity with the SDGs.

The present evaluation synthesis contains a first comprehensive and systematic aggregate assessment of sustainability in evaluations of German Financial and Technical Cooperation (FC and TC). The study is confined to evaluations of the two major official implementing organisations – the KfW Development Bank (KfW) and the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH. As the object of the evaluation is to be addressed as comprehensively as possible, the study is not restricted either to particular sectors, or to particular regions or types of project. In addition to purely bilateral projects in specific countries, the study also covers regional, sectoral and global projects.

## **Methodology**

The factors affecting the sustainability score were analysed using multivariate regression models. These models allow investigators to ascertain the effect of various factors on the variable to be explained – in this case the score awarded for the sustainability of projects. Due to the limited availability of data it was only possible to include certain factors. Consequently the study is restricted to specific features of projects, factors associated with their implementation and available contextual information. The latter include both specific features of the immediate context of the development projects, and macro quantitative indicators at the level of partner countries. Furthermore, the analysis also draws on findings of the accompanying meta-evaluation on sustainability in German development cooperation. The findings of the meta-evaluation allowed the evaluation team to include in

their analyses the criteria used to assess sustainability. Secondly, the investigators used the assessment of evaluation quality performed in the meta-evaluation as a weighting factor for individual observations in the regression models. No observations were ruled out of the analysis. However, the weighting of individual observations does ensure that the most credible findings received the greatest weighting in the synthesis.

### **Key findings, conclusions and recommendations**

#### *Factors affecting the rating of project sustainability*

In the evaluations conducted by the KfW and GIZ the sustainability score varies only slightly. Over 84 per cent of the evaluations included awarded a score of 2 or 3 for sustainability. Furthermore, the higher the score awarded for the DAC criteria relevance, effectiveness, efficiency and impact, the higher the sustainability score (1 = highest score, 6 = lowest score). Consequently, in all regression models the average score for all DAC criteria (excluding sustainability) – according to statistical level of significance and effect size – is the key determinant of the sustainability score.

Hence sustainability is an overarching evaluation criterion. It contains barely any genuine determinants that can be strictly separated from the remaining DAC criteria. Nonetheless, the regression models do demonstrate that certain factors are particularly important with regard to sustainability rating. In particular, the information obtained from the accompanying meta-evaluation on sustainability permits conclusions regarding the sustainability of specific projects. The findings of the accompanying meta-evaluation also demonstrate that although sustainability is assessed on the basis of comprehensive criteria in practice, this assessment is at the same time performed unsystematically and inconsistently. Through the assessment of impact, the assessment of sustainability is also always linked to the assessment of the other DAC criteria.

Differences in the assessment of sustainability also arise according to the type of evaluation used. While ex-post evaluations base their assessments on observations, in project progress reviews (PPRs), project evaluations (PEs) and final evaluations sustainability is assessed on the basis of a

prognosis. Compared to the other types of evaluation, ex-post evaluations tend to award the lowest scores for project sustainability. But it is not only the scores that differ depending on the type of evaluation. So too do the criteria on which they are based. Comparing the sustainability scores between different projects is thus only possible to a limited extent. Generally speaking, however, it can be concluded that in ex-post evaluations the role and the contributions of development partners and target groups are particularly important for the sustainability of projects. By contrast, when sustainability is assessed in PPRs, PEs and final evaluations it is primarily the direct outputs, the implementation of the project and the context of implementation that are taken into account.

Alongside these differences, however, the determinants identified in the different types of evaluation also display commonalities. For instance, in both ex-post evaluations and in PPRs, PEs and final evaluations, the predictability of the continuation of results has a significant positive effect on project sustainability. This shows that in all types of evaluation, not only the outputs and results of projects, but also the durability of results – a key conceptual element in the assessment of sustainability – has a significant effect on the sustainability score.

#### *Recommendations on boosting the sustainability of projects*

The recommendations below result from the findings and conclusions of the evaluation synthesis. Due to their complexity, the recommendations are supplemented – in the various sub-points – by suggestions and ideas that relate primarily to their application.

The evaluation team recommends that when planning and implementing projects, the BMZ and the implementing organisations should take greater account of the capacities of the partners and executing agencies on the ground, and systematically support their development.

- With this in mind, an explicit assessment of the capacities of all relevant partners and agencies might also be taken into consideration when deciding on the eligibility for support of a module during project

planning. Here it should be ensured that the partners and agencies possess the technical, financial and institutional capacities to continue the activities and outputs previously generated by the project.

- Furthermore, the capacities of the partners and agencies could be analysed repeatedly at regular intervals in the course of an ongoing project. Successfully transferring the outputs to the partners at the end of the project could also be underpinned by developing long-term exit strategies.
- Strengthening the partner system might ensure partner-country ownership of implementation of the 2030 Agenda.

The evaluation team recommends that the GIZ and KfW in future understand the factors relevant to the management of the project not only in relation to effectiveness, but also in direct relation to sustainability, and take this into account accordingly.

- These include particularly the use of institutional structures on the ground, the systematic analysis of lessons learned and the development of scaling-up and exit strategies.

#### *Systematic learning from evaluations*

The comparability of evaluation findings is a key prerequisite for conducting evaluation syntheses. Aggregating findings from individual evaluation reports promotes systematic, strategic and cross-institutional learning. Unfortunately, the findings on the sustainability of development cooperation projects found in the evaluation reports are only comparable to a certain extent. There are various reasons for this.

First of all, although the key questions do provide guidance for assessing sustainability, they are not sufficiently operationalised. This is reflected by the fact that the specific criteria underlying each individual score are manifold, and cannot always be specified unequivocally. Given the diversity of the portfolio of implemented measures a certain flexibility in assessment is necessary; even so, the assessment of sustainability must also be comprehensible and comparable for outsiders. This idea is

also reflected in the principle of joint accountability in the 2030 Agenda.

Secondly, the implementing organisations studied here display systematic differences in the practice and management of evaluation. The findings demonstrate that GIZ evaluations award significantly higher sustainability scores than KfW evaluations – even though the same number of criteria are rated positively. Furthermore, the use of different types of evaluation both within and between the implementing organisations leads to structural differences in the assessment of sustainability. There are also fundamental differences in the way the two implementing organisations manage evaluations. At the KfW all ex-post evaluations are audited by the evaluation department. Here the assessment of individual measures is placed in the context of the assessment of comparable measures. By contrast, the conduct of PPRs and PEs is decentralised. Responsibility rests with the officer responsible for the commission in question. Whereas at the KfW a core team of staff members checks all reports, thus establishing a minimum degree of comparability, the decentralised evaluation system at the GIZ precludes the organisation-wide comparison of individual reports. It is therefore to be assumed that overall, evaluations of GIZ projects are more heterogeneous and depend more heavily on attributes of the authors than is the case at the KfW.

Thirdly, the meta-data from evaluations and projects that are recorded by the implementing organisations tally only to a certain extent. Information relevant to the present analysis was in some cases incomplete, or was systematically recorded by only one implementing organisation.

The sketchy comparability of sustainability ratings makes it more difficult to identify enabling factors for sustainability. For instance, on the basis of the information available it is not possible to establish definitively whether the macroeconomic and political indicators integrated into the models actually have no effect on the sustainability of projects, or whether it is not possible to establish any link at all due to the lack of comparability and transparency of the criteria on which the assessment was reached. The potential for obtaining from evaluation syntheses strategic findings and findings that

would be relevant to the management response is thus very limited.

*Recommendations on boosting systematic, strategic and cross-institutional learning*

The recommendations below are also supplemented with suggestions and ideas that relate chiefly to their application.

To guarantee the systematic assessment of sustainability, the evaluation team recommends that the BMZ and the implementing organisations develop standardised and binding criteria. These should serve as a basis for the award of scores, and should be weighted transparently for this purpose.

- To take due account of the heterogeneous portfolio of German Technical and Financial Cooperation, the criteria should possess an appropriate degree of sector- and region-specific flexibility. Binding instructions on applying the criteria might also be defined separately for each sector or for TC/FC modules.

The evaluation team recommends that the BMZ and the implementing organisations – where possible – harmonise meta-data on projects and their evaluations and record this information at a central point.

- The systematic and central recording of meta-data from projects and evaluations would make cross-institutional, aggregated analyses considerably easier to perform, and therefore quicker.
- With this in mind, the BMZ and the implementing organisations might explore how they could meet the requirements of joint accountability articulated in the 2030 Agenda by recording and systematically preparing meta-data.