



# META-EVALUATION ON THE QUALITY OF (PROJECT) EVALUATIONS IN GERMAN DEVELOPMENT COOPERATION

*Executive Summary*  
2022

# IMPRINT

## Authors

Dr Kerstin Guffler  
Laura Kunert  
Marian Wittenberg  
Dr Nico Herforth

## Responsible team leader

Dr Kerstin Guffler

## Responsible head of department

Amélie Gräfin zu Eulenburg

## Cover and graphic design

Katharina Mayer

## Editing

Marcus Klein, PhD

## Translation

Dr John Cochrane

## Photo credits

Cover: VektorMine, Shutterstock

## Bibliographical reference

Guffler, K., L. Kunert, M. Wittenberg and N. Herforth (2022), *Meta-evaluation on the Quality of (Project)Evaluations in German Development Cooperation*, German Institute for Development Evaluation (DEval), Bonn.

## Published by

German Institute for  
Development Evaluation (DEval)  
Fritz-Schäffer-Straße 26  
53113 Bonn

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

E-mail: [info@DEval.org](mailto:info@DEval.org)

[www.DEval.org](http://www.DEval.org)

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 interventions.

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:

<https://www.deval.org/en/publications>

A BMZ response to this evaluation is available at <https://www.bmz.de/de/ministerium/evaluierung/bmz-responses-19422>

© German Institute for  
Development Evaluation (DEval), 2022

ISBN 978-3-96126-191-8 (PDF)

This is an excerpt from the publication "Meta-evaluation on the Quality of (Project)Evaluations in German Development Cooperation". Download the full report here: <https://www.deval.org/en/evaluations/our-evaluations/meta-evaluation-on-the-quality-of-project-evaluations-in-the-german-development-cooperation>

# EXECUTIVE SUMMARY

## Introduction

**Meta-evaluations can be described as evaluations of evaluations. They are becoming increasingly important in development cooperation.** According to Caracelli and Cooksy (2009: 2), meta-evaluations are conducted “to improve an evaluation in process, reflect systematically on the strengths and weaknesses of an evaluation to enhance one's future evaluation practice, or provide information about the credibility of the findings to users.” This understanding is also adopted in the present meta-evaluation. There are now a growing number of meta-evaluations in international development cooperation. In German development cooperation too, intra-organisational and cross-organisational meta-evaluations are now being carried out.

**This meta-evaluation ran concurrently with the preparation of the Evaluation Policy for German Development Cooperation, published by the Federal Ministry for Economic Cooperation and Development (BMZ). It follows on among others from previous assessments of the German development cooperation system, the study on monitoring of a system assessment, and the sustainability meta-evaluation (Noltze et al., 2018).** In 1999, the BMZ commissioned a first systematic assessment of evaluation practice in German development cooperation (Borrmann et al., 1999). Ten years later, a second system assessment was completed (Borrmann and Stockmann, 2009). Building on the latter, in 2015 DEval conducted a study to monitor implementation of its findings and recommendations (Lücking et al., 2015). In 2018, DEval published a cross-organisational meta-evaluation<sup>1</sup> on the quality on the quality of project evaluations of the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH and KfW Development Bank (Noltze et al., 2018).

**The BMZ published its Evaluation Policy for German Development Cooperation (BMZ, 2021a) prior to the completion of this meta-evaluation, thus setting the most recent milestone. This policy further consolidates the understanding of evaluation quality in German development cooperation.** In particular, the policy states that both the standards of the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD), and the standards of the Evaluation Society in Germany (DeGEval), are binding for the official implementing organisations. It also states that they provide guidance for the non-governmental organisations. Since the period of the meta-evaluation preceded the adoption of the policy, it was not possible to include it in the analysis. Nevertheless, important findings were generated for the future unfolding of the policy.

**One particular feature of this meta-evaluation is that it examines the application of quality standards in evaluations for a variety of both official implementing (governmental) and non-governmental organisations.** The following organisations were involved in the meta-evaluation: the Federal Institute for Geosciences and Natural Resources (BGR), CARE Germany, the Institute for International Cooperation of the German Adult Education Association (DVV), the German Red Cross (DRK), the Protestant Agency for Diakonie and Development (EWDE), GIZ, the Heinrich Böll Foundation (hbs), the Konrad Adenauer Foundation (KAS), KfW, MISEREOR and the National Metrology Institute (PTB). This meta-evaluation examined project evaluations conducted between October 2016 and December 2022 for which German organisations were (co-)responsible. Germany's official implementing (governmental) organisations were all included. The non-governmental organisations were selected using criteria that reflected their structural heterogeneity. This enabled the findings to reflect the broadest possible range of experience with the application of quality standards.

<sup>1</sup> The BMZ's Evaluation Policy for German Development Cooperation states that DEval meta-evaluations are part of the quality assurance framework for the evaluation system (BMZ, 2021a).

The present meta-evaluation looked at the application of the quality criteria in relation to the requirements for the involved organisations to apply the OECD-DAC and/or the DeGEval standards. It also examined the evaluations of GIZ and KfW in relation to the quality criteria of the sustainability meta-evaluation.

**This cross-organisational meta-evaluation aims to generate findings on the involved organisations' understanding of evaluation quality, and on strengths and weaknesses in the application of the quality standards. It also identifies and analyses factors linked to the application of the quality standards.** To support future learning, the meta-evaluation also aims to find explanations for the non-application of quality standards. To achieve the above, the meta-evaluation addressed the following evaluation questions.

### Evaluation questions

1. What understanding of evaluation quality do the involved German development cooperation organisations have?
2. To what extent are quality standards applied in evaluations of the involved German development cooperation organisations?
  - a) To what extent are strengths and weaknesses evident in the application of the OECD-DAC and the DeGEval standards in the evaluations of the involved German development cooperation organisations?
  - b) To what extent are strengths and weaknesses evident in the application of the organisation-specific quality standards in the evaluations of the involved German development cooperation organisations?
  - c) To what extent are strengths and weaknesses evident in the application of the sustainability meta-evaluation quality criteria in the evaluations of the GIZ and KfW?
3. To what extent are country-specific, evaluation-specific and organisation-specific factors linked to the application of quality standards?

## Theoretical and empirical background

### *The understanding of quality*

**This meta-evaluation equates evaluation quality with the application of the relevant quality standards, i.e. the quality standards that are required for the involved organisations. It then examines evaluation quality accordingly.** This definition of quality includes – but is not limited to – the content of the OECD-DAC and DeGEval standards documents. The basis for identifying evidence of good evaluations is provided by the OECD-DAC and DeGEval standards. This is due to their international recognition, their link to development cooperation and their relevance to German development cooperation organisations. The term “application of quality standards” was chosen in consensus with the reference group<sup>2</sup>. It describes the extent to which evidence can be obtained as to whether and how the quality standards were included in the evaluations examined – i.e. whether this was documented in writing, or was reported back in writing upon request by the evaluation team. The OECD-DAC and DeGEval standards are maximum standards. This means that the involved organisations do not have to apply all quality standards in all evaluations. These internationally recognised quality standards were complemented with organisation-specific quality standards and the quality criteria of the sustainability meta-evaluation.

<sup>2</sup> The reference group comprised representatives of the involved organisations and VENRO, and officers from BMZ Division GS 22 “Evaluation and development research, DEval, IDOS”. Its members accompanied the evaluation process during all phases of the evaluation (for example through virtual meetings or comments on evaluation documents DEval, 2021a).

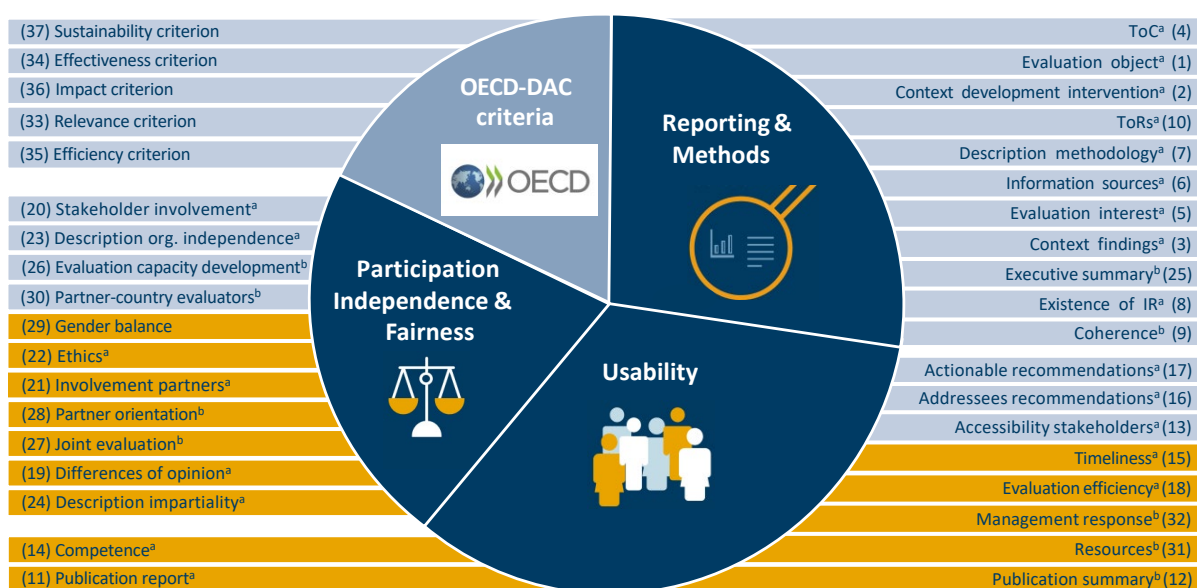
## Analysis grid

The analysis grid includes quality criteria derived from the OECD-DAC and DeGEval standards documents. It also includes organisation-specific quality criteria and quality criteria from the sustainability meta-evaluation (Noltze et al., 2018). The OECD-DAC and DeGEval quality criteria can be classified into three areas. These are: 1) the overlap between the standards documents of OECD-DAC and DeGEval, 2) the OECD-DAC standards document minus the overlap with the DeGEval standards document (OECD-DAC only) and 3) the OECD-DAC criteria. Since all involved organisations were committed to the “OECD-DAC criteria” (BMZ, 2006), these are listed as a separate area, although their application is part of the OECD-DAC standards (quality standard 2.8 OECD-DAC, 2010). For DRK, EWDE, GIZ and hbs, the analysis grid also included organisation-specific quality criteria.<sup>3</sup> For GIZ and KfW it included quality criteria already addressed in the previous sustainability meta-evaluation. criteria already addressed in the previous sustainability meta-evaluation.

## Standard clusters

With the exception of the five OECD-DAC criteria, the quality criteria were assigned to three standard clusters – “reporting and methods”, “participation, independence and fairness” and “usability”.<sup>4</sup> The standard cluster “reporting and methods” mainly comprises quality criteria that relate to either the presentation of information on the evaluation methodology, or the existence or content of selected evaluation documents (Figure 1). The standard cluster “participation, independence and fairness” is primarily assigned quality criteria that concern the inclusion of different groups of people in the evaluation. The standard cluster “usability” focuses primarily on the usefulness of evaluations, while active use plays only a minor role, and benefit is not examined.<sup>5</sup>

**Figure 1** Assignment of the 37 quality criteria to the standard clusters and OECD-DAC criteria



Source: DEval, authors' own graphic

Note: blue bar = quality criterion examined by evaluation; yellow bar = quality criterion examined at the level of the organisation across all evaluations; org. = organisational, IR = Inception Report; ToC = Theory of Change; ToRs = Terms of References. <sup>a</sup> Quality criterion comes from the area of overlap between OECD-DAC and DeGEval standards; <sup>b</sup> Quality criterion comes from the "OECD-DAC only" area. The organisation-specific quality criteria and the quality criteria of the sustainability meta-evaluation are not shown, as they were not examined for all organisations.

<sup>3</sup> Organisation-specific quality criteria were defined as requirements that are of great importance to an organisation for the quality of its evaluations, independently of the OECD-DAC or DeGEval standards.

<sup>4</sup> The terms used to designate these three standard clusters display similarities to the designations of the DeGEval standards groupings (1. usefulness, 2. feasibility, 3. fairness and 4. accuracy). However, as the identified quality criteria partly represent the overlap between the OECD-DAC and the DeGEval standards, naming the clusters identically would not properly reflect their content. The OECD-DAC standards are structured largely in relation to evaluation phases, hence these terms were not used.

<sup>5</sup> Details on the definitions can be found in section 2.1 of the main body of the evaluation report.

### *Factors affecting application of the quality criteria*

The factors described below were analysed if they possessed three attributes, namely: **1) They had a clear cross-organisational definition. 2) It was possible to describe clear links to selected quality criteria. 3) Data were available either from the organisations or in secondary databases.** To identify the factors, three focus group discussions were conducted with the responsible officers of the involved organisations, and scientific and empirical literature was reviewed. The factors were then systematically categorised as either 1) country-specific, 2) evaluation-specific or 3) organisation-specific.

## **Methodology**

### *Data and data analysis*

**Once the organisations had been included in the meta-evaluation, a stratified random sample of a total of 296 evaluations was drawn for analysis of the application of quality standards.** The organisations were selected using four criteria. This was in order to cover the greatest possible structural heterogeneity (criteria 1 to 3), and enable the analysis of a sufficient number of evaluations per organisation (criterion 4). In total, the evaluation units/desks of the organisations were (co-)responsible for 839 evaluations in Germany during the period under review from October 2016 to December 2020. The population then encompassed 576 evaluations that the BMZ had either (co-)funded, or in which a development intervention (co-)funded by the BMZ was examined. The sample drawn from that comprised 296 evaluations.

**The understanding of quality (evaluation question 1):** To determine the understanding of quality among the organisations involved, and which quality standards they were required to apply, the meta-evaluation examined documents of the organisations as well as relevant agreements with the BMZ. It also conducted a qualitative content analysis.

**OECD-DAC and/or DeGEval standards (evaluation question 2a):** To examine the application of the quality criteria in the 296 evaluations, evaluation documents (evaluation reports and annexes, ToRs, inception reports) and organisational documents (such as evaluation plans, guides and manuals for conducting evaluations) were used. In the inter-coding phase, for 14 of the 37 OECD-DAC and DeGEval quality criteria either very little information or none at all could be coded in the evaluation documents provided by the organisations. To avoid drawing erroneous conclusions about non-application, in a further step the responsible officers of the evaluation units/desks were therefore surveyed concerning the application of these quality criteria in their organisation.<sup>6</sup> The quality criteria described in the analysis grid were coded on the basis of ordinal or binary scores. Values were also assigned to the average frequencies of application of the quality criteria reported by the responsible officers of the evaluation units/desks in the online survey. For each quality criterion, different statistics (e.g. mean values) were calculated for each organisation. These statistics were subsequently converted into percentages and assigned to the predefined thresholds. A mean value across all organisations was then calculated, and analysed using descriptive statistical methods.

**Organisation-specific quality standards (evaluation question 2b):** To examine the application of the eleven organisation-specific quality criteria, these were coded in the evaluation documents of the four organisations concerned. Values were then calculated and analysed as with the OECD-DAC and the DeGEval quality criteria.

<sup>6</sup> Since the quality criteria of the online survey were applied on average about 6 per cent less than the quality criteria of the document analysis, there was no reason to assume that the organisations systematically rated themselves more highly than they would have been rated by the objective coding.



**Quality criteria of the sustainability meta-evaluation (evaluation question 2c):** To analyse the application/repeated application of the quality criteria from the sustainability meta-evaluation, 15 quality criteria from the previous sustainability meta-evaluation of Noltze et al. (2018) were used. Eight quality criteria were already included as quality criteria in the OECD-DAC and in the DeGEval analysis grid and were transformed. The remaining seven were coded anew in the current evaluations. Descriptive statistics were calculated for the analysis. Structural equation models were calculated for the difference in findings between the sustainability meta-evaluation and the present one (Weiber and Mühlhaus, 2010).

**To investigate links between selected factors and the application of the quality criteria, multivariate regression analyses were estimated (evaluation question 3).** Regression analyses allow the identification of statistical correlations between the factors (independent variables) and the quality criteria as well as the standard cluster “reporting and methods” (dependent variables; Backhaus et al., 2011). The factors examined were either 1) country-specific, 2) evaluation-specific or 3) organisation-specific. The information for each factor was obtained from data submitted by the organisations or obtained from secondary databases.

### *Evaluating the application of quality standards*

**For organisations required to apply the quality criteria (group 1), application of the quality criteria was both analysed and rated. For organisations with no such requirement (group 2), application of the quality criteria was only analysed.** The thresholds for rating the application of the quality criteria were defined in dialogue with the reference group. The threshold represented the ex ante assessment of when a quality criterion was considered to be barely applied, partly applied, largely applied or applied in an evaluation. When setting the thresholds in increments of 25 per cent ( $0 \leq 25$  per cent = “barely applied”,  $> 25 \leq 50$  per cent = “partly applied”,  $> 50 \leq 75$  per cent = “largely applied”,  $> 75 \leq 100$  per cent = “applied”), the evaluation took into account the fact that the quality standards are to be understood as maximum standards. The rating was based on the defined thresholds for application, and by adding the extreme values 0 (not achieved) and 100 (exceeded) for group 1. Since application of the quality criteria of the sustainability meta-evaluation was now being analysed once again, at this point it was postulated that the application of the quality criteria had improved since that meta-evaluation. Both groups were analysed with regard to their degree of application of the quality criteria for the OECD-DAC and the DeGEval standards. Group 1 was also rated. For the OECD-DAC criteria (BMZ, 2006), the organisation-specific criteria and the quality criteria of the sustainability meta-evaluation, only organisations in group 1 were analysed and rated.

### *Strengths and challenges in the methodology*

**The selection of the organisations based on their structural heterogeneity enabled the meta-evaluation to analyse and describe the application of individual quality criteria across a correspondingly wide range.** The findings of this meta-evaluation are valid for the involved organisations. The official implementing organisations are fully described. The selection of the involved non-governmental organisations is not representative for all non-governmental development organisations. The selection of organisations enables non-governmental organisations that were not involved to locate themselves within this range and use the findings of the meta-evaluation for themselves. The transferability of the findings to the population of evaluations by organisation was ensured by the selected statistical parameters for sampling. Transferability to other evaluation types of an organisation is not.

**Since the analysis grid was derived from the OECD-DAC and DeGEval standards, it can also be used by other organisations.** The analysis grid of this meta-evaluation can thus be used in the future to produce an analysis grid based on the BMZ Evaluation Policy.

**When analysing the quality criteria from the online survey, there were limitations with regard to the triangulation of methods. Given the cost-benefit ratio, however, adding a further data collection method would not have been appropriate.** Besides surveying the responsible officers of the evaluation units/desks, the triangulation of data might also have included the assessment of the evaluation teams for the respective evaluations. However, this meta-evaluation covered a large number of evaluations.

It was therefore beyond its scope to interview former evaluators from those evaluations instead of or in addition to the officers responsible for the evaluations. Furthermore, in some cases staff turnover would have made it no longer possible to locate and interview some of the evaluators. The sample size would thus have been reduced.

**Generally speaking, there were limits to measuring some quality criteria. For certain quality criteria, it would take a lot of effort to investigate a “good” application. There are quality criteria that can only be examined in depth with a great deal of effort.** For example, for the quality criteria “stakeholder involvement” and “accessibility for stakeholders”, both the appropriate number of stakeholders who can be involved and the intensity of their involvement in the various evaluation phases are difficult to determine.

**The operationalisation of the quality criteria was developed across all organisations. Some of these operationalisation choices did not match the evaluation practice of all organisations. This meant that for these organisations, application received a lower score than it would have if the criteria had been operationalised otherwise.** There are quality criteria for which there was greater leeway (such as “quality assurance processes”). Accordingly, there is a conflict of objectives here between the meta-evaluation's interest in the application of selected quality criteria across organisations, and the heterogeneity of the application of the quality criteria.

**Due to an inconsistent understanding of the measurement of “evaluation costs” across organisations, it was only possible to analyse explanations of the application/non-application of quality standards to a limited extent.**

**Examining the quality criteria of the sustainability meta-evaluation once again enabled the present meta-evaluation to analyse the difference in the application of the quality criteria by the GIZ and KfW over time. This provides evidence of the extent to which application of the quality criteria was improved through internal organisational reforms and the support of external actors (BMZ and DEval).** It also brought to light challenges associated with a longitudinal study (such as raising the thresholds, and where necessary making appropriate adjustments to quality criteria over time).

## Findings

### *Evaluation question 1: What understanding of evaluation quality do the involved German development cooperation organisations have?*

The involved organisations' understanding of quality was predominantly based on the OECD-DAC and/or DeGEval standards and, where applicable, organisation-specific quality standards. When the meta-evaluation began, the involved organisations had in some cases not addressed these quality standards systematically. Furthermore, the BMZ's requirements concerning the application of the quality standards varied in the selected budget items during the period under review – in some cases the OECD-DAC standards were marked as mandatory, in others no requirements were specified.

### *Evaluation question 2a: To what extent are strengths and weaknesses evident in the application of the OECD-DAC and the DeGEval standards in the evaluations of the involved German development cooperation organisations?*

Overall, a positive picture emerged regarding the application of the OECD-DAC and DeGEval standards. The involved German development cooperation organisations applied the quality standards in about two thirds of their evaluations. This was also the case – to a somewhat lesser degree – for organisations not required to apply the quality standards. Application of the quality standards sometimes varied widely between the organisations. This was to be expected, given the selection criteria for the inclusion of the involved organisations in the sample. It was thus possible to obtain a heterogeneous picture across the varying degrees of application.

However, it emerged that the organisations had largely not yet fully identified the quality standards in their organisational documents. Nor had they systematically prescribed the application/non-application of these quality standards. This was also the case for the traceability of the application or



**non-application of some selected quality standards at the level of the individual evaluation.** It should be noted that, for various reasons, the application of some quality standards was recorded not at the evaluation level, but at the organisational level. This might be due to either the way in which the meta-evaluation chose to operationalise some quality criteria, or a lack of documentation on application/non-application, or the fact that application was documented exclusively at the organisational level rather than at the level of the individual evaluation. There is a clear need for improvement here, as without information, an external investigation of an (explained) application/non-application of the quality standards at the level of the evaluation is only possible to a limited extent. It was thus not possible to trace whether a quality standard was either not applied (with or without explanation), or was applied but was not documented.

**The OECD-DAC criteria were very largely achieved by the involved German development cooperation organisations.** However, it should be explicitly pointed out that operationalisation took place in relation to the OECD-DAC standards rather than the OECD-DAC criteria. This made it easy to achieve the threshold. In the application of the OECD-DAC criteria, first documentation of non-application was also available at the organisational and evaluation levels. In this respect, the application of the OECD-DAC criteria already differed – albeit to a minor extent – from the application of most of the other quality criteria. It can be assumed that in future evaluations, documentation of the non-application of the OECD-DAC criteria will continue to increase. This is because since 2020 (BMZ, 2020), the updated BMZ guidelines on evaluation criteria require priority setting that is explained and transparent.

In the annex to the report in section 7.1, the findings for the four official implementing organisations BGR, GIZ, KfW and PTB are also presented and classified at the level of the individual organisation.

*Evaluation question 2b: To what extent are strengths and weaknesses evident in the application of the organisation-specific quality standards in the evaluations of the involved German development cooperation organisations?*

**A positive picture also emerged for the application of the organisation-specific quality standards by DRK, EWDE, GIZ and hbs.** On average, these quality criteria were “largely achieved”. Once again, there was potential for improvement in the explanation of non-application at the evaluation level.

*Evaluation question 2c: To what extent are strengths and weaknesses evident in the application of the sustainability meta-evaluation quality criteria in the evaluations of the GIZ and KfW?*

**Regarding the application of these quality criteria, the picture was a positive one – with a few exceptions. Specifically, the quality criteria were achieved on average to a degree of about 75 per cent.** This represents a somewhat higher degree of application than was found for the OECD-DAC and DeGEval standards. However, challenges remained in the application of the quality criteria “selection procedure for interviewees described” and “control/comparison groups included”. Furthermore, it should be noted that the application of all the quality criteria has improved – in some cases clearly – since the sustainability meta-evaluation. Overall, an average difference of 36 per cent was observed. These changes indicate that the measures implemented after the sustainability meta-evaluation to improve the evaluation practices of GIZ and KfW might have affected application. This is a very positive result in light of the extensive efforts made by a large number of actors in connection with the measures. It should be noted, however, that alternative explanations cannot be ruled out (for example, operationalisation of the quality criteria in ways that make it relatively easy to achieve the thresholds, or changed documentation methods).

***Evaluation question 3: To what extent are country-specific, evaluation-specific and organisation-specific factors linked to the application of quality standards?***

Overall, the findings provide little evidence of significant links between the factors identified in the literature and the focus group discussions, and the application of the quality criteria. This involves mainly evaluation-specific factors: “number of internal and external evaluators” and the implementation of various “quality assurance processes”.

**Conclusions and recommendations**

**The meta-evaluation's recommendations are derived primarily from the evaluation questions on the application of selected quality standards and quality criteria (OECD-DAC, DeGEval and/or organisation-specific quality standards, as well as quality criteria of the sustainability meta-evaluation, evaluation questions 2a to 2c).** The recommendations are worded in general terms. This means that each involved organisation must consider its own organisation-specific findings in order to determine the particular relevance of the recommendations to it. This is because the cross-organisational mean value presented in the findings is not sufficiently meaningful for evaluating the application of quality standards by specific organisations. The criteria-based selection of non-governmental organisations focuses on their structural heterogeneity, and thus reflects the range of possible degrees and forms of application for different organisations. This means that non-governmental organisations which were not involved can also determine the relevance of the findings to them, and thus draw guidance from the conclusions and recommendations. The recommendations addressed to the BMZ are intended for the BMZ Evaluation Division.

***Identification of relevant/non-relevant quality standards and prescription thereof in organisational documents***

**Recommendation 1**

- a) As part of a revision of their evaluation practice, the evaluation units/desks of the BGR, CARE, DRK, DVV, EWDE, GIZ, hbs, KAS, KfW, MISEREOR and PTB should (if they have not already done so) identify the quality standards that are required for their organisation. They should explicitly prescribe these in organisational documents, and define their application in evaluation processes. The organisations should review at regular intervals the identification and systematic prescription of quality standards. When doing so, they should specify the degree of application they require for each of the quality standards.
- b) In the context of upcoming updates of funding guidelines or ancillary provisions for individual budget items, the BMZ should make a contribution towards strengthening its Evaluation Policy as a reference document for evaluations. As part of these updates, together with the non-governmental organisations concerned the BMZ should establish and prescribe special conditions for particular organisations (e.g. as in the case of the funding guidelines for political foundations). The principle of maximum standards should be retained here.
- c) Based on its Evaluation Policy, and in dialogue with the official implementing and non-governmental organisations, the BMZ should develop an analysis grid for the application of the quality standards, taking into account the analysis grid of the present meta-evaluation. It should also make this available to the official implementing and non-governmental organisations.

### **Recommendation 2**

- a) The evaluation units/desks of the BGR, CARE, DRK, DVV, EWDE, GIZ, hbs, KAS, KfW, MISEREOR and PTB should explain and prescribe in organisational documents any general non-application of particular quality standards that are required for them.
- b) The BMZ should reach an agreement with the official implementing organisations on the application and (explained) non-application of the quality standards described in the BMZ Evaluation Policy. It should do so either in order to jointly determine non-application at the organisational level or to document discrepancies.

### ***Ensuring the application and traceability of the application/non-application of relevant quality standards at evaluation level***

### **Recommendation 3**

- a) The evaluation units/desks of the BGR, CARE, DRK, DVV, EWDE, GIZ, hbs, KAS, KfW, MISEREOR and PTB should, if they have not already done so, further improve the application of the quality standards required at the organisational level (recommendation 1) in individual evaluations, and especially those quality standards that are barely or partly applied. Furthermore, the application or (explained) non-application of all quality standards should be traceable at the level of each evaluation and regularly reviewed by the organisations.
- b) The BMZ should urge the official implementing organisations to ensure the application of the relevant quality standards, and the traceability of their application/non-application, at evaluation level.

### ***Joint learning***

### **Recommendation 4**

- a) The evaluation units/desks of the BGR, CARE, DRK, DVV, EWDE, GIZ, hbs, KAS, KfW, MISEREOR and PTB, and representatives of VENRO, should regularly share their various lessons learned in identifying, prescribing, assuring and tracing the application/non-application of all quality standards. This dialogue should also integrate non-involved organisations and include further types of evaluation – such as decentralised evaluations – in order to continue improving the application of quality standards.
- b) The BMZ should financially support the dialogue with and between the organisations on identifying, prescribing, assuring and tracing the application/non-application of the quality standards.

### ***Ensuring the application and traceability of the application/non-application of the sustainability meta-evaluation quality criteria***

### **Recommendation 5**

- a) When developing the analysis grid for the quality standards described in the BMZ Evaluation Policy (recommendation 1), the BMZ should consider adopting the quality criteria from the sustainability meta-evaluation. If appropriate, it should also include them in the analysis grid.
- b) Based on recommendation 5a, the GIZ and KfW should ensure/improve the application/non-application of the quality criteria from the sustainability meta-evaluation that have been incorporated into a BMZ analysis grid. They should also guarantee the traceability of the (explained) application/non-application for each evaluation.

This is an excerpt from the publication "Meta-evaluation on the Quality of (Project)Evaluations in German Development Cooperation". Download the full report here: <https://www.deval.org/en/evaluations/our-evaluations/meta-evaluation-on-the-quality-of-project-evaluations-in-the-german-development-cooperation>