



Federal Ministry
for Economic Cooperation
and Development

Evaluating German Development Cooperation

BMZ Evaluation Policy

BMZ PAPER

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

Contents

Abbreviations and acronyms	5
1. Introduction	6
2. Fundamentals – Definitions, guiding principles and criteria	9
2.1 What we mean by evaluation	9
2.2 Other instruments of review	13
2.3 Evaluation and review of results and efficiency in the german legal context	13
3. Standards – Implementing the fundamentals	15
3.1 Appropriate institutional structures – a precondition for credible evaluations	15
3.1.1 Central, independent evaluations	15
3.1.2 Further types of evaluation as part of an organisation-specific evaluation system	16
3.2 Evaluation process and products – milestones of high-quality evaluations	18
4. The BMZ evaluation system	24
4.1 Key actors in the evaluation system	24
4.2 Different channels for implementing development cooperation – adapted approaches for evaluation	26
4.2.1 General aspects for all channels	26
4.2.2 Official bilateral cooperation with partner countries	28
4.2.3 Cooperation with civil society and municipalities	30
4.2.4 Cooperation with international organisations and the EU	31
4.2.5 Thematic cooperation – special initiatives and transitional development assistance	33
4.2.6 Cooperation with the private sector	33
5. Implementation and the way forward	35
5.1 Dialogue and cooperation with partners	35
5.2 How evaluation is institutionalised	37
5.3 Further development of evaluation and quality assurance	38
6. Literature	41

Abbreviations and acronyms

BGR	Federal Institute for Geosciences and Natural Resources
BHO	Federal Budget Code
BMZ	Federal Ministry for Economic Cooperation and Development
BMZ-E	BMZ evaluation unit
BRH	Bundesrechnungshof [Supreme federal audit office]
CSOs	Civil society organisations
DeGEval	DeGEval – Evaluation Society
DEval	German Institute for Development Evaluation
ECD	Evaluation Capacity Development
ECG	Evaluation Cooperation Group
EU	European Union
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH [German agency for international cooperation]
IO	International Organisation
KfW	KfW Development Bank
M&E	Monitoring and Evaluation
ODA	Official Development Assistance
OECD	Organisation for Economic Co-operation and Development
OECD DAC	Development Assistance Committee of the OECD
Policy	BMZ Evaluation Policy for German Development Cooperation
PTB	Physikalisch-Technische Bundesanstalt [National Metrology Institute]
ToC	Theory of Change
ToR	Terms of Reference
RIE	Rigorous Impact Evaluations
UNEG	United Nations Evaluation Group
VV-BHO	Allgemeine Verwaltungsvorschriften zur Bundeshaushaltsordnung [General Administrative Regulations on the Federal Budget Code]

1. Introduction

By adopting the 2030 Agenda for Sustainable Development and the Paris Climate Agreement, the international community – with Germany as an active player – has created a collective guiding framework for action. All countries have pledged to harmonise economic progress with social justice, and keep it within the planet’s ecological boundaries. Integral components of the framework include the safeguarding of human rights, the imperative of peace and security, the rule of law, democracy and good governance.

Germany’s Federal Ministry for Economic Cooperation and Development (BMZ) is set to contribute tangibly to the Sustainable Development Goals at all levels – globally, in Germany and Europe, and in its partner countries.

Spending resources wisely to achieve policy goals is a cross-cutting budgetary principle, hence also for development policy, i.e. not only spending money according to purpose and economically, but also effectively and efficiently. Changing opportunities and challenges at all levels require the course of action to be put to the test periodically. Sound and evidence-based information on effects, causal relationships and the performance of development cooperation is essential for policy and operational decision-making.

THE ROLE OF EVALUATION IN GERMAN DEVELOPMENT COOPERATION

Since its early years, the BMZ has recognised evaluations as an effective means of learning and accountability. Evaluations aim to improve development effectiveness and legitimacy by supplying credible analyses and assessments of the results achieved by development interventions.

They address the following key questions:

- To what extent does German development cooperation bring about positive changes with regard to relevant development objectives, in a way that is sufficiently cost-effective and sustainable? What are the causes of the strengths and weaknesses identified?
- What, if anything, should be changed in order to improve development effectiveness?

Beyond compliance with legal requirements and international agreements, BMZ is also committed to shaping and adapting development policy steadily based on knowledge gained from evidence of what has been achieved.

Given an open learning culture, independent and high-quality evaluations that are useful for decision-making processes can and should add value to the effective design and implementation of interventions both at the BMZ and in development organisations (i.e. official, non-governmental and international organisations working with BMZ funds). They can and should also promote informed discussions and decision-making within organisations, and by supervisory bodies or donors to whom organisations are accountable. Similarly, in appropriate cases, they enrich the dialogue with partner governments and partner organisations. They can also promote an informed, evidence-based dialogue with development professionals and the interested public.

PURPOSE AND SCOPE OF THE POLICY

This Policy defines and explains the objectives, principles, and standards, as well as the distribution of tasks and the procedures in Germany's development evaluation system. It aims to ensure the quality and use of evaluations, as well as the coherence and efficiency of the system as a whole. Based on relevant principles and standards of the OECD's Development Assistance Committee (DAC), and the cross-policy area standards of the DeGEval - Evaluation Society, it also incorporates international best evaluation practices, as well as the principles of the Aid Effectiveness Agenda and the 2030 Agenda for Sustainable Development (UN 2015).

This Policy reflects the state of the art in German development evaluation work, put into practice already to a considerable extent. It is designed to help align the shared understanding already achieved even more systematically with the remaining challenges in the coming years.

Applying principles and standards consistently

- permits **division of labour in the evaluation system**, and an approach that is quality assured. Essentially, the system comprises two levels: The first level is broad-based. Development organisations that implement the BMZ's policy priorities and budget funds, manage evaluations especially at project level. The second level is more selective. The German Institute for Development Evaluation (DEval) conducts evaluations at the strategic policy level; in exceptional cases these are commissioned by the BMZ. Concerning supplementary, actor-specific mid-level strategic evaluations, see [Section 4.2](#).
- facilitates **syntheses and use of evaluation findings** as well as **cooperation** between the actors, and thus the achievement of **synergy effects**.
- provides the basis for the further **institutionalisation** of evaluation as a function as well as the **further development** of evaluations and their use within the BMZ and the development organisations.

The **scope of application** of this Policy encompasses all Official Development Assistance (ODA) from the BMZ budget, which the BMZ implements through development organisations together with partners in developing and emerging countries, internationally, and in Germany and Europe. It also includes market and other third-party funds for which the BMZ bears political (co-)responsibility.

This Policy is **binding** for the BMZ and Germany's official implementing organisations (BGR, GIZ, KfW and PTB) as well as Engagement Global. It formulates **minimum requirements** for the evaluation work of DEval and provides **guidance** for German civil society organisations – in each case in conjunction with contractual agreements, administrative regulations or funding guidelines of the BMZ with respect to these organisations. At the same time, it provides **information** for international organisations and the EU. Due to their independence and international membership, these entities usually have evaluation policies or similar regulations of their own (which are compatible with the BMZ policy). Adapted accordingly, this Policy also provides guidance for the BMZ's financial contributions to these organisations (see [Section 4.2](#)).

The following chapters explain what the BMZ means by evaluation (also as distinct from other instruments of review), as well as the principles and the criteria of evaluation ([Chapter 2](#)); the key organisational, process and product standards ([Chapter 3](#)); the evaluation system, i.e. the division of labour between the main actors and their respective roles and responsibilities, and the corresponding procedures in detail ([Chapter 4](#)); finally, aspects of implementation and further development to address remaining challenges,

like the firmer integration of evaluation into a BMZ results-based management model and a more thorough involvement of partner country organisations ([Chapter 5](#)).

This Evaluation Policy is supplemented by a **Glossary of key terms** used therein ([BMZ 2021a](#)), the **Guidelines for using evaluation criteria** ([BMZ 2021b](#)) and the **Guidelines for joint ministerial evaluations** ([BMZ 2021c](#)). To leave space for updates, these are kept as separate documents.

In the following chapters, the present tense is used either to describe well established facts, or to describe something which must be observed. Descriptions worded “need(s) to”, “has/have to”, “will” or indirect prescriptions such as “is a prerequisite” are used interchangeably to indicate a requirement. The word “should”, or any other recognisably similar wording, points to a recommendation.

2. Fundamentals – Definitions, guiding principles and criteria

An evaluation system that involves a division of labour requires a consistent and shared conceptual understanding of evaluation. The BMZ Evaluation Policy is guided by the standard-setting agreements reached by the OECD DAC. A further frame of reference is provided by the cross-policy area standards of the DeGEval - Evaluation Society (see [Section 2.1](#)). To enhance understanding, related instruments for review ([Section 2.2](#)) and links to the German Federal Budget Code (see [Section 2.3](#)) are also outlined below. Concerning further operationalisation of the fundamentals in standards, see [Chapter 3](#).

2.1 WHAT WE MEAN BY EVALUATION

Since its early years of existence as a ministry in its own right, the BMZ has recognised evaluation as an effective means of learning from experience and providing a credible account of the results achieved by development interventions. Evolving over time, the following key issues emerge:

- (1) The **overarching objective of development evaluation** is to help improve the development effectiveness of interventions. Evaluations aim to generate evidence of relevance to the intended main users (the knowledge function). This evidence serves two purposes: (a) to provide robust underpinnings for managerial decisions – with respect to either specific development interventions, or with respect to strategic policy issues (the learning and management function); (b) to provide accountability for the extent of results achieved with the tax funds used (the control and accountability function). If evaluation results are communicated to a wider audience, they can also contribute to a better understanding of development cooperation, its potential and its risks. Depending on their timing and intended use, evaluations can serve both learning and accountability, or primarily one of the two objectives.
- (2) **In development cooperation, evaluations are defined** as systematic and objective analyses and assessments of ongoing or completed development interventions. These assessments usually encompass the design, implementation and results of development interventions. They need to provide information that is credible and useful ([OECD DAC 1991](#) and [2002/2009](#)) and, in appropriate cases, recommendations that are actionable. Or, defined in a nutshell: the systematic analysis of the “merit or worth of an object” ([DeGEval 2016](#), [OECD DAC 2019](#)).

The term ‘evaluation’ presupposes the application of **principles** based on **standards**, including a methodologically robust, transparent and verifiable analysis and assessment of empirical data in relation to predefined **criteria** and benchmarks (concerning principles and criteria see below, and concerning standards see also [Chapter 3](#)).

In German development cooperation, ex ante analyses are not included under the term “evaluation”. However, they should be guided by the evaluation criteria and promote evaluability (see [Section 5.2](#)).

(3) Core principles are

- **usefulness** of the results (e.g. through **participation** by affected parties and stakeholders in the process, **transparency**, i.e. dissemination and publication of the report, and **follow-up** of actionable recommendations);
- **credibility** of the findings (through **professionalism**, i.e. quality of methods and processes, such as **accuracy** and **verifiability** of the analysis as well as **fairness** and transparency of the process);
- **independence** of assessments (through **impartiality** of the approach and **unbiased** analysis in reporting and presentation of evaluation results, supported by the absence of undue influence, e.g. by the organisation commissioning the evaluation, achieved through appropriate structural measures and by avoiding conflicts of interest).

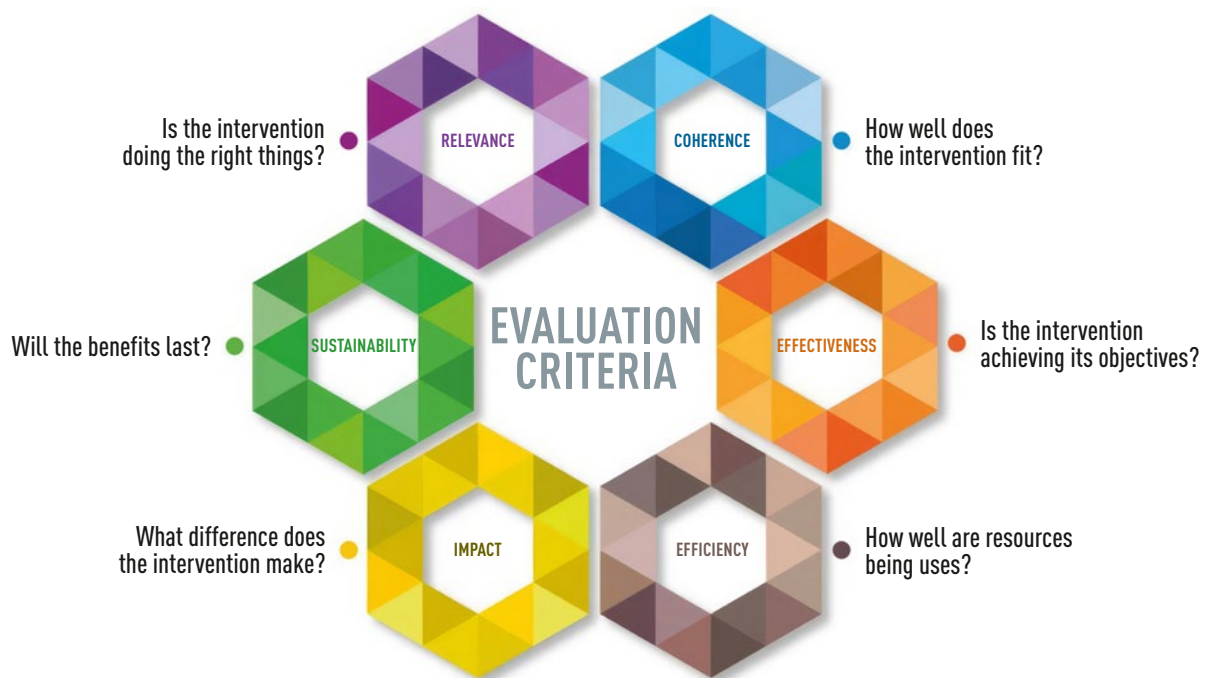
In German development evaluation, depending on the object of the evaluation, two further principles are especially important:

- **partnership**, i.e. first of all – insofar as appropriate and feasible given the object, the question and the context of an evaluation – participation by the responsible partner organisations, particularly with the aim of strengthening their ownership; secondly, participation – especially in joint programmes – by development partners, and in the case of BMZ and DEval evaluations, other donors and ministries to ensure coherence and efficiency;
- **ethical standards that incorporate human right principles**, especially with activities targeting or affecting people directly. Evaluations of such interventions must respect human dignity and diversity and must not harm people during the evaluation process. Furthermore, they must assess (i) the intervention's provisions for non-discrimination and for equality of opportunity to access the activities, and (ii) the outcomes of the intervention, disaggregated by group (such as gender, age, ethnicity or religious affiliation). Concerning application of the human rights principles of transparency, accountability and participation by beneficiaries and affected parties, see below as well as [Section 3.2](#) on process and product standards.

These evaluation principles, and the associated standards and application of the criteria (see below), are closely aligned with the principles of the **Aid Effectiveness Agenda** of [Paris \(2005\)](#), [Accra \(2008\)](#) and [Busan \(2011\)](#), and the **2030 Agenda for Sustainable Development** ([UN 2015](#)).

They are not mutually exclusive and can be interdependent. For example, without independence there is no credibility, and without credibility there is no usefulness. Usefulness implies partnership, and ethical principles are key to credibility. Goal-conflicts arising from applying these principles need to be considered carefully and transparently.

- (4) **Standardised criteria** support comparability, and thus the potential for cooperation between actors as well as for synthesis products. For assessments of development effectiveness (as a composite and comprehensive consideration of the worth or significance of an intervention overall) the six (reformed) DAC criteria (OECD DAC 2019) are regularly applied – relevance, coherence, effectiveness, efficiency, impact and sustainability – and the guiding questions linked to them:



Source: OECD DAC 2019. For the application of the criteria in German bilateral cooperation see BMZ 2021b.

Using the DAC criteria enables a comprehensive analysis of the **development effectiveness** of an intervention, i.e. the extent to which the intervention generates relevant outcomes and impact in relation to (direct and strategic) objectives – including unintended (positive or negative) effects – efficiently and in a sustainable manner. The criteria need to be operationalised appropriately using concrete and contextualised evaluation questions, in a way that reflects the purpose and the object of the evaluation. If individual criteria are not applied or other criteria (see below) are selected, the reasons why need to be explained. Reasons may include e.g. the purpose, the object or the timing of an evaluation.

The criteria take into account the **effects or risks with respect to the three inseparable core dimensions of sustainable development** (i.e. the social, environmental and economic dimensions, and where relevant their interactions and conflicting objectives), as well as coherence with international and national norms and standards (e.g. human rights).

Cross-cutting themes therefore have to be evaluated in principle and irrespective of the objectives of the intervention and preferably within the scope of the criteria (see also the BMZ quality criteria for German development cooperation, BMZ 2020). The analysis should be conducted in a manner that is appropriate to the object of the analysis (evaluand), is risk-oriented, and takes into account any unintended effects, including negative ones (see also [Chapter 3](#)) when assessing:

- a. compliance with human-rights-based due diligence obligations, and social and environmental safeguards, in planning and implementation;
- b. equality of opportunity to access outputs of development interventions financed or co-financed by the BMZ, effects or risks with regard to gender equality (*gender-sensitivity*) and entitlements of different population groups, particularly with respect to disadvantaged or vulnerable individuals and groups (*leave no one behind*); depending on the reach of the development intervention, also reduction of inequality;
- c. effects and risks with respect to the climate and the environment, bearing in mind intergenerational equality (*climate-sensitivity*);
- d. effects and risks with respect to violent conflicts (*conflict-sensitivity*).

Context matters. In particular the political framework conditions can have considerable influence on the results of development interventions – and especially on their sustainability. Specific circumstances may lead to adaptation of the criteria: For interventions in the context of acute natural disasters or in conflict settings (like most interventions for transitional development assistance), the criteria can be interpreted appropriately or supplemented with the criteria for humanitarian assistance (see [ALNAP 2006](#)). In this case, “sustainability” can be replaced by “connectedness”, and “coverage” can serve as a criterion in its own right. Observing the do no harm principle is especially important in this context. This is to be achieved through conflict sensitivity (of the development intervention and the evaluation procedure). Evaluations therefore need to also consider effects or risks of the intervention with regard to lines of conflict, conflict dynamics and local capacities for conflict management. The evaluability can be affected by limitations on field studies, especially for safety and security reasons. This can be partially offset by using digital methods (see [Section 5.2](#)).

On the further need to adapt the criteria, see also the section on cooperation with the private sector, which is gaining importance ([Section 4.2.6](#)).

- (5) **Basically, two types of evaluations are distinguished**, depending on the scope of the evaluation, namely ***project evaluations*** and ***strategic evaluations***. ***Project evaluations*** focus on individual development interventions (a single project, which may also have a programme-type structure, a single policy-based approach). ***Strategic evaluations*** focus on comprehensive interventions, e.g. the engagement in a whole sector or thematic area, in a partner country or with a particular instrument (e.g. financing mechanism, mode of delivery, approach, strategy). They are usually based on case studies of several individual interventions. Strategic evaluations also include ***institutional evaluations***, which focus on organisations, networks or their sub-systems, as well as ***meta-evaluations*** and ***evaluation syntheses***. Concerning these and further types of evaluation, see the Glossary of key terms ([BMZ 2021a](#)); from an institutional perspective, see also [Section 3.1](#).

This basic distinction between ***project evaluations*** and ***strategic evaluations***, which is also common internationally, is particularly relevant to the division of labour and corresponding tasks within the German development evaluation system. For further details, see [Chapter 4](#).

2.2 OTHER INSTRUMENTS OF REVIEW

The following approaches for reviewing development cooperation are related to, but distinct from, evaluation:

- (1) **Monitoring** by the entity or entities responsible for implementing (and/ or supervising implementation of) the intervention, for management purposes. This involves the ongoing, systematic collection of data on changes critical to results and progress achieved in the course of implementing development interventions. Good monitoring data capture relevant changes in the context and within the project executing agency, particularly with respect to risks. They also reflect changes in the actual course of implementation as compared to the original time schedule, disbursement plan, and planned activities, outputs and their use (implementation or performance monitoring). Furthermore, they capture changes at the level of outcomes and impact (impact-oriented monitoring). Monitoring data are key for self-assessments, and for evaluations. They save time and costs when performing evaluations. In turn, evaluations can generate valuable information for improving monitoring.
- (2) **Reporting** to the BMZ on the status achieved by the organisation accountable to the BMZ, based on self-assessments (progress and final reports); see also [Section 2.3](#).
- (3) **Data collection and aggregation of standard indicators**, especially for the purpose of communicating with the public.
- (4) **Audit of the performance and regularity** by independent units within a development organisation, at the BMZ (auditing, audit group) or by external bodies (particularly chartered accountants) to assess compliance of administrative and business practices with the law and the principle of economy. Together with evaluations, these reviews of the use of funds provide systematic and independent analyses and assessments of development cooperation work from different perspectives, which are in principle complementary. In addition, Germany's supreme audit institution (BRH) provides external financial oversight in line with its constitutional mandate.

Closely related to evaluation and potentially overlapping is empirical research. Accompanying research of the effects of interventions, particularly in the form of rigorous impact analyses, can be counted as a specific approach to evaluation, provided that core quality standards of evaluation are observed (see [Chapter 3](#)). Conversely, science-based evaluation such as that of DEval is also a specific approach of applied research.

2.3 EVALUATION AND REVIEW OF RESULTS AND EFFICIENCY IN THE GERMAN LEGAL CONTEXT

In Germany, reviewing results and efficiency is prescribed “for all measures having financial (fiscal) impact” of the federal government by law (VV-BHO on Section 7 of the German Federal Budget Code). Ultimate responsibility for this rests with the ministries. These reviews consist of two steps: the study proper (so-called “efficiency review”) and the follow-up. The *study proper* must encompass a summative, systematic and verifiable analysis and assessment of (a) the achievement of (project and higher-level objectives), (b) the causality and appropriateness of the intervention for the extent of achieved objectives, and (c) efficiency. In addition, there is also a step two, *the follow-up*. Pursuant to the BHO regulations, the purpose of these reviews is not just the analysis of the collected data and rating (the study proper), but also particularly the use of the findings for managing ongoing measures, or for designing follow-on or similar projects or programmes. Concerning complementary requirements for the planning of measures, see [Section 5.2](#).

The BMZ implements this legal requirement mostly by using regular **reports** (self-assessments) required from the development organisations working with BMZ funds (see [Section 2.2](#)). These are supplemented by (independent) **evaluations** selected on the basis of transparent criteria (see [Section 4.2](#)). Following the BMZ business model of financing, the **unit of account** for reviewing results and efficiency is usually an individual development intervention (“project”).

Evaluation reports can be studies on which the legally required review of results and efficiency is based since:

- the functions are similar (accountability, learning and follow-up on the evaluation results);
- the OECD DAC criteria also encompass those of the BHO; the DAC criteria “effectiveness” and “impact” include the achievement of objectives as well as outcomes and impact achieved in relation to both (direct) project objectives and (higher-level) strategic objectives of a development intervention; both OECD DAC and BHO use the criterion “efficiency”;
- the methodological requirements of verifiability and robustness of data collection and analysis are basically similar.
- In BMZ practice, evaluations are to a large extent already being used systematically for the mandatory reviews of results and efficiency (see also [Section 4.2](#)). To the extent that evaluations are also intended to meet results and efficiency review needs, this presupposes that they comply with the BHO requirements.

However, roles and standards of evaluations in the BMZ’s remit go far beyond serving the need for results and efficiency reviews in various respects:

- In line with internationally shared objectives of development policy, in German development cooperation evaluating **sustainability** is equally essential to determine the extent of results achieved by interventions. This includes capturing two distinct concepts: For one, the three dimensions of sustainable development of the 2030 Agenda. These are covered by several OECD DAC criteria. In addition, durability of the benefits after completion of the support, when interventions have been transferred to sole partner responsibility. This latter aspect is covered by the OECD DAC criterion of “sustainability”. The OECD DAC evaluation criteria also include further standard criteria that go beyond the BHO, such as relevance and coherence (see [Section 2.1](#) and Guidelines for using evaluation criteria, [BMZ 2021b](#)).
- **Evaluations are also more demanding**, or at least more explicit, in terms of the required **quality standards**, particularly with respect to independence, and also in terms of many process and product standards (see [Section 2.1](#) and [Chapter 3](#) for further details). Moreover, they are subject to institutionalised **quality assurance** (see [Sections 3.2](#) and [4.2](#)) as well as **follow-up** on the implementation of recommendations and of **learning** from the knowledge gained (see [Sections 3.2, 4.2](#) and [5.3](#)).
- The frame of reference for evaluations is also different. DEval’s strategic evaluations in particular aim to pursue development policy-relevant issues, which lie outside any immediate results and efficiency review requirements (mostly already fulfilled by project-type evaluations), such as evaluations of instruments and country-specific or thematic strategies. While serving primarily information and knowledge needs at the policy level, these evaluations can provide relevant and empirically sound additional information for the purpose of reviewing results and efficiency as well.

3. Standards – Implementing the fundamentals

The fundamentals provided in [Chapter 2](#) entail concrete organisational, process and product standards. These are explained below. The organisational standards ([Section 3.1](#)) are based in particular on the OECD DAC principles ([OECD DAC 1991](#)), the OECD DAC working tool on evaluation systems and use ([OECD DAC 2006](#)), and best practice. The process and product standards ([Section 3.2](#), covered to some extent in more detail in [Section 4.2](#)) are based on the OECD DAC quality standards ([OECD DAC 2010](#)) and the DeGEval standards ([DeGEval 2016](#)). These standards can only be applied successfully, however, when other units within the BMZ and the development organisations support them in a spirit of shared responsibility with different roles (see also [Section 5.2](#)).

3.1 APPROPRIATE INSTITUTIONAL STRUCTURES – A PRECONDITION FOR CREDIBLE EVALUATIONS

In line with good international development evaluation practice, independent evaluations are managed or carried out by central, structurally independent evaluation units (see [Section 3.1.1](#)), sometimes supplemented by other types of evaluation (see [Section 3.1.2](#)). As a policy innovation, the BMZ has largely externalised the central evaluation function, and transferred the task of conducting policy-relevant strategic evaluations to DEval. For further details on the tasks and the differentiated responsibilities in the German development evaluation system, see [Chapter 4](#).

3.1.1 *Central, independent evaluations*

Central, independent evaluation units, in smaller organisations, possibly in conjunction with (non-competing) policy tasks and separate from operational tasks, should preferably be attached directly to the highest level of the organisation, or enjoy a status that is equivalent to being **structurally independent** by virtue of not being bound by instructions of middle-level management concerning evaluative tasks. Their tasks include in particular

- **evaluation programming**: preparation of a preferably organisation-wide, annual or multi-year **evaluation plan**, covering all development interventions;
- **managing evaluations: design, process management and quality assurance** of evaluations as well as unaffected (e.g. without co-signature requirements) reporting to the managing director or executive board of the organisation and, where such exists, to a supervisory body; also **publication, dissemination** (for further details see [Section 3.2](#)) and **documentation** of the reports;
- **quality assurance** of decentralised evaluations, if applicable (see below);

- **provisions for implementing evaluation recommendations** (see [Section 3.2](#)) and well established processes for internal, systematic and targeted **communication** of evaluation results and lessons learned. Possibly commission of **evaluation syntheses**;
- drafting of an evaluation policy or appropriate regulation, preferably supplemented by guidance documents that take into account this Policy and the aforementioned OECD DAC documents.

These features can also have an important signalling function vis-à-vis other public and private funders.

Central evaluation units need to be provided with sufficient resources by the development organisations. This includes an appropriate budget and a professional staff with good knowledge of development cooperation and the work of the organisation, and especially of evaluation standards and methods, as well as management skills. Relevant training can help and should also be provided.

The **evaluation plan** should preferably take into account information and learning needs of the organisation as a whole (irrespective of funding source), as well as duties of accountability towards funders and supervisory bodies. In any case it needs to include an appropriate selection of BMZ-funded interventions, based on an organisation-specific or budget item-specific selection procedure (for further details see [Section 4.2](#)).

Before an evaluation is included in this plan, or at the latest before it is commenced, it is advisable to check both the **evaluability** (i.e. feasibility of the evaluation) and its **efficiency**. Enabling conditions for conducting an evaluation include e.g. a sufficiently long implementation period, availability of data and accessibility, and especially in conflict regions security (however, see digitalisation in [Section 5.3](#)). Furthermore, to ensure efficiency, an appropriate balance should be struck between the costs of an evaluation and the intended knowledge gain. Field studies are not always the most suitable approach here. This is the case for instance when sufficient robust evidence is already available, e.g. through data surveys conducted by development partners. In such cases a desk study may suffice (see e.g. [Section 4.2](#), co-financing).

Development organisations need to take this Policy into account in their **evaluation policies** or through other appropriate regulations, explaining inter alia the types of evaluation (cf. Sections [2.1\(5\)](#) and [3.2\(1\)](#)) and their role as well as procedures, division of tasks and the form of quality assurance. Evaluation policies and any supplementary guidance documents, and the evaluation plan, should be documented, accessible and communicated widely within the organisation.

For **BMZ**, this system-wide Policy replaces a BMZ-only policy, supplemented by published and internal guidelines and further guidance documents.

3.1.2 Further types of evaluation as part of an organisation-specific evaluation system

Development organisations differ in terms of their size and decision-making structures, as well as their procedures for designing and implementing development interventions. Consequently, in some development organisations evaluations managed by a dedicated central unit are supplemented by evaluations **managed or co-managed by other bodies**, including such outside the organisation. The organisation-specific arrangements should be concretised in appropriate regulations such as an evaluation policy (see [Section 3.1.1](#)).

Where the term “evaluation” is used, the principles as well as the process and product standards contained in this Policy (see [Chapter 2](#) and [Section 3.2](#)) need to be complied with. This also includes a sufficient degree of independence, and quality assurance mechanisms. Regarding the **managing unit**, evaluations managed by a central evaluation unit are sometimes supplemented by:

- (1) **Decentralised evaluations managed by the development organisations:** These sometimes even represent a major part of the evaluation system, especially with smaller organisations or organisations with a high degree of decision-making delegation to local offices. To avoid obvious conflicts of interest, decentralised evaluations are not to be managed or commissioned by those operationally responsible for the design or implementation of the development intervention. They are to be carried out exclusively by external evaluators with proven expertise and independence. Arrangements have to be in place to ensure the quality and independence of the evaluation, including quality assurance, e.g. through a central evaluation unit.

Further evaluative approaches provided by or managed by those units or bodies which are operationally responsible for the intervention evaluated (sometimes referred to as “self-evaluations”), however, do not fall under the term “evaluation” in the sense of this Policy, notwithstanding their importance and use of data and empirical analysis. Some are geared towards *joint learning* between functional units in development organisations and partner organisations, or empowerment of target groups. These valuable, participatory approaches can help those involved to reflect on and improve the quality of an intervention. *Reporting* is another case in point, despite its importance for reviewing results and efficiency.

At the **BMZ**, decentralised evaluations are not a regular part of evaluation work (however, see the special case of funding programmes, [Section 4.2.3](#)). When these are occasionally initiated at the BMZ, they take into account the fundamentals and standards of this Policy. In the future, in order to avoid the duplication of activities, decentralised evaluations and other reviews similar to evaluations that assess the effectiveness of development cooperation in aggregate terms in relation to policy objectives (including (2) and (3) below), as well as synthesis studies, meta-reviews, evidence gap maps and the like (see also [Section 5.3](#) on syntheses of rigorous impact evaluations) will be registered with the BMZ evaluation unit by the commissioning functional unit at the planning stage and communicated to DEval.

- (2) **Partner-led evaluations:** Some developing and emerging countries already have binding regulations that make evaluations of government programmes mandatory. Civil society partner organisations are increasingly seeking self-managed evaluations. These can include interventions funded by German development cooperation. In line with the 2030 Agenda and the Aid Effectiveness Agenda (ownership), development organisations should support these as far as possible and participate appropriately, also the BMZ in the case of strategic evaluations.
- (3) **Joint evaluations with development partners and/or partner organisations:** Particularly in the case of joint programmes, joint evaluations with the participation of partner organisations are an appropriate choice in line with the Paris Declaration (2005) and the European Consensus on Development (2017), see [Section 5.1](#). This also includes cross-ministerial evaluations (see [Sections 4.2](#) and [5.1](#)).

These evaluations are recognised as being equivalent to central evaluations, if evaluation units are involved, or the evaluations described at (2) and (3) above apply appropriate quality standards including quality assurance.

3.2 EVALUATION PROCESS AND PRODUCTS – MILESTONES OF HIGH-QUALITY EVALUATIONS

The following overview of actors, process steps and products is based on the model of an evaluation managed by the central evaluation unit of a development organisation. It is aligned with the OECD DAC Quality Standards for Development Evaluation (2010) and the DeGEval Standards (2016). For non-centrally managed evaluations, the process and product standards have to be applied as well, using an adapted but functionally equivalent form of implementation of the tasks intended for the central evaluation unit. This is to ensure independence and quality to a sufficient degree.

These procedures apply as well to **BMZ** and **DEval** evaluations, sometimes with appropriate variations. In the case of DEval evaluations, as an independent evaluation institute, DEval largely assumes the tasks of an evaluation unit of the BMZ as well as the evaluation team (see also [Section 4.1](#)). The quality of its evaluations is assured through internal and external peer reviews, as well as in a defined reference group process.

Actors in the process

The **main actors** are the evaluation unit, the evaluation team, and relevant stakeholders.

Evaluation units are responsible for process management and process quality throughout the entire cycle, from planning to disseminating and using the evaluation results for management purposes. They ensure independence, as well as the quality of implementation and of the products of the evaluation team. They also ensure transparency and participation by the relevant stakeholders in the evaluation during the process.

Evaluation teams are responsible for the professional conduct of evaluations. They are committed to a fair, transparent, and ethically sound approach, in compliance with the principles and standards of this Policy and the sources mentioned above. They maintain confidentiality and follow data protection guidelines.

Relevant stakeholders are those actors which are affected primarily by the evaluation. They include the functional unit responsible for the development intervention being evaluated, and other affected work units within the development organisation. Usually there are further relevant stakeholders, like partner organisations. They contribute their information needs early on and provide relevant information about the intervention evaluated including context, on a timely basis throughout the process. In evaluations conducted by development organisations, depending on the scope of the evaluation, the BMZ is also one of the relevant stakeholders.

Process and products

Appropriate **stakeholder participation** throughout the process is crucially important for the quality and subsequent use of the evaluation (in-process learning). In strategic evaluations of the **BMZ** and **DEval**, a **reference group** serves as a “sounding board” for the evaluation at predefined milestones. Besides the BMZ units concerned, the reference group includes development organisations affected, partner organisations in appropriate cases and possibly other knowledge bearers.

The evaluation process typically **comprises four phases**, each with one or more specific products, depending on the scope of the evaluation, and with different actors responsible for and contributing to each phase: (1) preparation (design); (2) implementation (data collection and analysis); (3) reporting (preparation and submission of the report); (4) follow-up (outreach and implementation of recommendations).

(1) Preparation (preliminary design). Responsible: evaluation unit; milestone: ToR or equivalent (concept note, approach paper)

The **conceptual design**, to be prepared in all cases, usually takes the form of a **Terms of Reference (ToR)** for the selection of the evaluation team.

The evaluation stakeholders should be involved in the design, and their information needs should be taken into account, e.g. through appropriate evaluation questions. Timeline and budget have to be in line with the task and the intended use. If an inception report is planned (see below), it is recommended that the evaluation team be given room for further conceptual refinement, balancing inter alia the knowledge and information needs, and feasibility.

The **evaluation team** is selected transparently by the evaluation unit in accordance with the procurement regulations of the respective development organisation – unless the analysis and assessment is carried out by appropriately qualified personnel of the evaluation unit itself. Hybrid forms are also possible. Evaluation teams have to (as a team) possess in particular good methodological skills, plus thematic knowledge and experience with the subject matter and the context of the intervention, and relevant cross-cutting competences (e.g. intercultural competence, gender competence). Safeguarding independence, they must not have been previously involved in the planning or implementation of the development interventions under review. During the selection process, attention also needs to be paid to further possible conflicts of interest, such as the financial dependence of team members on other contracts from the same organisation. Possible conflicts of interest have to be scrutinised, avoided if possible and disclosed where necessary. In order to ensure the dual control (two pairs of eyes) principle as well as the competence of the team, i.e. evaluation and thematic competence, at least two members are generally planned. The teams are gender-diverse (at least across the evaluation programme). In the case of country-specific interventions, they preferably include qualified evaluators from the region in which the development intervention is or was implemented – also with a view to evaluation capacity development.

(2) Implementation (data collection and analysis). Responsible: evaluation team; milestone for strategic evaluations: inception report, further products as appropriate

At least in the case of strategic evaluations, an **inception report** needs to be provided for further development of the ToR/ the preliminary design. This contains an initial analysis of the object of the evaluation, the context and the theory of change (ToC), reconstructed if necessary. It also describes the further procedure, including the planned methods of data collection and analysis. The evaluation questions have to reflect the knowledge and information needs of the intended users and be backed up by an (at least preliminary) evaluation matrix.

The **context**, i.e. in particular the political and socio-economic (non-influenceable) framework conditions in a partner country, is of considerable importance for assessing development effectiveness. The relevant baseline conditions for the intervention, as well as any changes over time, need to be considered. Responsiveness and risk management in the implementation of the intervention are also assessed as a critical factor for effectiveness.

The **implementation of data collection** varies according to the purpose or object of the analysis. The selection and scope of the information collected needs to enable the evaluation questions to be answered appropriately. Undue burden on respondents is to be avoided through appropriate selection and procedures for

obtaining necessary information (principle of data economy). In the course of interviews, care must be taken to obtain prior, informed consent while assuring confidentiality (non-traceability of information sources).

If the **selection of (more in-depth) case studies** is involved, the selection has to be based on transparent criteria, taking efficiency issues into account.

In the case of **field studies in partner countries**, the evaluation team presents the findings and any initial conclusions to the partner organisation and, where appropriate, other stakeholders, in order to obtain partner and stakeholder perspectives.

If, in the course of data collection, evaluation teams receive indications of sexual abuse or ethically questionable behaviour that violates dignity, they must inform the evaluation unit confidentially in consultation with the persons concerned. The evaluation unit in turn informs the appropriate bodies within their organisation. Any additional risks for affected persons must be avoided. Evaluation teams also need to inform the evaluation unit confidentially in the event of indications of corruption (and any other criminal offences) or serious conflicts of interest in connection with the implementation of the development intervention under analysis.

Methods and approaches: Evaluations are based on objective (replicable), i.e. at least robust and comprehensible findings. The intersubjective verifiability and accuracy of the findings are of key importance. The findings are based on data collection and data analysis in accordance with recognised (and in justified cases also innovative) designs and methods. The designs and methods applied need to be appropriate to the objectives or purpose of an evaluation, the evaluation questions, and the object. They also need to be feasible within the budget (“form follows function”). Using both quantitative and qualitative methods has proven to be particularly effective. In any case, an appropriate validation of the findings needs to be ensured, e.g. through triangulation of methods, cross-checking of data and information sources, and application of the dual control (two pairs of eyes) principle.

Evaluations not only address the question of whether and to what extent objectives have been achieved, but also and in particular how and in what way (including key success or failure factors). They also consider whether the objectives were appropriate. The basis for answering these questions is a theory-based approach that usually starts with a description (or reconstruction) of the **ToC** of the development intervention. The evaluation assesses the realisation of the ToC, and explicitly assesses unintended, positive and negative effects. If the development intervention is especially complex, it also considers non-linear causal pathways and larger indirect effects.

When assessing **effects** (in particular **outcomes and impact**), it is necessary to consider the counterfactual (key question: What would have happened /What changes would there have been if the development intervention had not been implemented?). This includes an enquiry about the baselines before the intervention began (key question: What has changed over time as a result of the intervention?). In order to be able to speak of effects, a causal link between observed changes and the development intervention – i.e. as opposed to external influencing factors – has to be comprehensibly established, regardless of the methods used. Potentially relevant external factors are therefore also explicitly considered.

In the course of data collection and analysis, the Guidelines for using evaluation criteria ([BMZ 2021b](#)) and [Section 2.1](#) need to be taken into account.

The experiences and perspectives of the **beneficiaries** of the development intervention, **and of groups and individuals affected by the intervention (also indirectly)**, must be ascertained. Particularly in the case of interventions that directly affect individuals, the do no harm principle has to be observed. This means remaining aware of e.g. risks of stigmatisation for certain groups, such as indigenous minorities, or risks of re-traumatisation in the context of conflict. Data on access to development cooperation services (in this case: activities and outputs), and on effects or risks, are disaggregated by gender and by vulnerable group.

(3) Reporting (preparation and submission of report). Responsible: evaluation team and the evaluation unit; milestone: final report

The **evaluation team** prepares the **final report** based on the data analysis. In the case of very complex evaluations, intermediate steps such as the thematic validation of the findings and conclusions prior to preparation of the recommendations have proven useful. Evaluation units provide quality assurance on the formal quality of the report. Relevant stakeholders need to have the opportunity to comment on the draft report, particularly with regard to its factual accuracy. Their comments need to be given due consideration in the final version. Any differences of opinion can be reflected e.g. in management responses or in implementation plans (see [4]).

The **quality of evaluation reports** (as the final product of the evaluation team and basis for further use) is particularly important. These reports must adhere to core quality standards as well. Reports that are comprehensible and use language that the target audience can understand, as well as the timely completion of reports, add significantly to their usefulness.

Strengths and weaknesses have to be presented in a balanced way on the basis of the findings, taking into account different perspectives. When rating scales are used (either verbal scales e.g. ‘good’, or numerical scales), appropriate scales and benchmarks are applied. Rights and legitimate interests, including those of people in partner countries, as well as company and business secrets, including those of cooperation partners, must be protected. Confidential information is anonymised accordingly. The evaluation questions are to be answered in full. The methodological approach and the assessments must be presented transparently (using an evaluation matrix). Any limitations – also with regard to the evaluation questions – need to be specified (e.g. data availability, time constraints, access difficulty due to conflict). If an evaluation is carried out by equal partners (e.g. a consortium), a transparent procedure to voice dissent among individual partners has to be provided for.

Findings, conclusions and recommendations are clearly separated. They are logically inferred from the evidence, and comprehensibly explained and documented. Recommendations are always provided in the case of an ongoing development intervention, or if the intervention will subsequently be continued by a similar intervention. Recommendations are clearly formulated, realistic in terms of number, form and addressee (i.e. understandable, workable and feasible) and allow verification of their implementation later on. The addressees of recommendations are always clearly named. As a matter of principle, the **BMZ** is not the addressee of recommendations in evaluations of development organisations.

If **‘lessons learned’** (generalisation of evaluation insights) are included in reports, these should be validated across several comparable cases.

Evaluation reports contain a meaningful **executive summary** that appropriately reflects the main body of the report and is written in language appropriate to the intended main users. The reports identify those responsible for the content of the report, and the commissioning party.

The evaluation unit will ensure that the **full reports are made available to the evaluation stakeholders** within and outside of the organisation. In the case of country-specific evaluations, executive summaries need to be translated into the partner country's national/ most widely used UN official language (currently at least English, French and Spanish). Depending on the intended main users of the evaluation, a converse procedure is also possible (full report in one of the foreign languages, executive summary: translated into German). This same principle also applies to joint evaluations with other partners.

The BMZ also always receives the full report with a German summary in addition, if applicable.

In the spirit of transparency, reports should preferably be published in full. Where legitimate interests militate against full publication, at least a meaningful summary report in German, prepared or authorised by the evaluation team, will be published. The **BMZ** publishes corresponding summaries of its evaluations. The full reports are provided upon request. **DEval** always publishes the full versions of its evaluations.

(4) Follow-up (outreach and implementation of recommendations). Responsible: functional and evaluation unit; milestone: implementation plan

Communication of evaluation results as well as follow-up on the implementation of recommendations are an integral component of development evaluation (see also [Section 4.2](#)).

Various formats of **communication** of evaluation results to a larger audience have proven effective. They should be targeted and tailored to the intended users, to support learning and enhance knowledge.

Concerning **recommendations**, evaluation units have to make provisions for **follow-up on their implementation**. At least in the case of *strategic evaluations*, evaluation units will ensure a formal **management response (implementation plan)** in their organisation. They will coordinate the management response process, and quality-assure the completeness and logicity of its content. Formats for this make clear which recommendations are accepted (or not, explaining why not) by those responsible, and which concrete implementation steps are planned and by when. The evaluation units will monitor the actual implementation of the steps foreseen in management responses at least on the basis of random checks and document the results (**implementation monitoring**). Responsibility for planning (and ultimately implementing) the response rests with the functional units. Evaluation teams can play an important advisory role in drafting the management response.

Development organisations need to notify the BMZ of their take-away and the steps taken (or intended) to follow-up on recommendations.

At the **BMZ**, the follow-up to strategic evaluations by DEval or the BMZ is divided into three parts; it comprises: a) a **public response** by the BMZ (part of the summary report in the case of BMZ evaluations or a separate response published at the same time as the report in the case of DEval evaluations); b) a formalised, internal **implementation plan** with concrete, time-bound and verifiable steps, and c) subsequent **implementation monitoring**.

Partner countries/partner organisations are equally important as a place of learning, particularly in many project evaluations and, depending on the issue, also strategic (e.g. country programme) evaluations. Development organisations need to use the evaluation results for joint reflection with their partner organisation, in order to draw joint lessons for the future from the strengths and weaknesses of the intervention. Ideally, options for joint learning and capacity development will already be incorporated into the design and implementation of the evaluation.

Concerning **evaluation syntheses** and further aggregate analyses, and their use, see Sections [4.2](#), [5.2](#) and [5.3](#).

4. The BMZ evaluation system

The BMZ evaluation system is structured so as to ensure coherence and efficiency of the system as a whole. This structure is based on a division of labour between the main actors and provides clearly defined roles and responsibilities for the actors involved (Section 4.1). It is aligned with the BMZ's way of implementing its development policy and hence leaves room for appropriate variations (Section 4.2).

4.1 KEY ACTORS IN THE EVALUATION SYSTEM

The key actors in the BMZ evaluation system are the BMZ's evaluation unit, DEval, and the independent evaluation units within the development organisations, with the roles and tasks outlined below; concerning complementary roles and key tasks in organisations, see also Section 5.2.

- (1) **BMZ's evaluation unit (BMZ-E)** gives direction to the evaluation system as a whole, and has a key role as an interface between the BMZ and DEval. BMZ-E:
 - sets standards for the evaluation of German development cooperation, and promotes their application and refinement within the evaluation system;
 - ensures structural complementarity of the evaluations carried out by DEval and by the development organisations in order to avoid duplication or gaps (cost-effectiveness of the overall system);
 - forms the interface and, if necessary, acts as a mediator between the BMZ's functional units and DEval on fundamental evaluation issues, particularly with a view to safeguarding the independence of the Institute's evaluation work, taking into account issues of usefulness;
 - coordinates and manages the collective evaluation needs of the BMZ by ensuring the participation of all Directorates-General. It also provides quality assurance of BMZ public responses and internal implementation plans for DEval evaluations.

Since the establishment of an independent evaluation institute (DEval), BMZ-E itself now plans and manages evaluations conducted by external evaluation teams only in exceptional cases. These include, in particular, international joint evaluations with other donors (usually not in a leadings role), and joint evaluations with other German federal ministries. DEval supports BMZ-E in the methodological quality assurance of these evaluations, except when it plays an active role in the evaluation itself. Decentralised BMZ evaluations can supplement these evaluations. If BMZ-E commissions evaluations itself, it assumes the characteristics of a central, independent evaluation unit with the corresponding tasks (see Section 3.1). A BMZ evaluation plan is substituted by the DEval evaluation plan, based on a prior consultative process within BMZ and between BMZ-E and DEval.

BMZ-E also promotes evaluation capacity development in partner countries of development cooperation, taking into account the special role of DEval in this context. Further tasks arise from its role as policy unit for the legally required reviews of results and efficiency and for methodologically robust

impact analyses. Concerning the representation of German development evaluation in international bodies together with DEval, taking account of different roles, see [Section 5.1](#).

- (2) As an independent evaluation institute, **DEval** is mandated to analyse and assess all ODA-related interventions (including special cases beyond ODA) in an independent and transparent manner. This also applies to the ODA of other federal ministries, provided that all parties involved agree.

The focus of its evaluation work (as its key area of work) includes in particular:

- scientifically sound, strategic and at the same time policy-relevant evaluations and reviews that focus on overarching thematic or country/regional development policy goals and their implementation, as well as on instruments and structures of development cooperation;
- cross-organisational evaluations as well as synthesis studies that provide broader insights (lessons learned), based on solid empirical evidence generated both for German development cooperation and/ or by other donors and academia;
- meta-evaluations (quality review) of evaluations carried out by development organisations.

DEval prepares its multi-year evaluation plan independently. This responds primarily to the evaluation needs of the BMZ and recommendations of the DEval Advisory Board, and especially the representatives of the political parties in the German Bundestag (the country's federal parliament). It also includes evaluation proposals put forward by DEval itself. The main intended users of the DEval evaluations are the BMZ and, indirectly, the German Bundestag. The plan requires the approval in full of the BMZ state secretary. DEval's independence in selecting its evaluation projects is safeguarded by transparent processes. DEval bases its evaluations on its own quality standards, which need to correspond to the standards of this Policy or go beyond them.

DEval reports to the BMZ state secretary and, through the latter, to relevant committees of the German Bundestag. The Institute publishes and disseminates all reports resulting from the evaluation plan. An advisory board comprising members of the German Bundestag and representatives of governmental and non-governmental German development organisations, international organisations and academia advises the Institute on all relevant issues, and participates in the selection of the Institute's managing director.

As an institute, DEval enjoys the greatest possible independence in selecting, conducting and disseminating its evaluations. Nonetheless, DEval does not work in isolation. Its autonomy is embedded in its federal funding purpose and its statutes. The institutionalised accumulation of knowledge acquired through its familiarity with the German development cooperation system and its largely self-reliant implementation of evaluations combined with its leading methodological competence for development evaluations, and its targeted knowledge transfer, are distinct features of its usefulness.

Further work focuses on adapting scientific methods and approaches to the context of development evaluation and strengthening evaluation capacities in partner countries. By disseminating evaluation results as well as sharing knowledge on methods and approaches, DEval can raise the profile of German development evaluation in international debates and play an active role in shaping these.

- (3) **Central, structurally independent evaluation units of development organisations** are a constituent part of the division of labour in Germany's development evaluation system (see [Section 4.2](#)). Concerning their characteristics and tasks, see [Section 3.1](#).

The evaluation units conduct evaluations of BMZ-funded projects and programmes on their own responsibility, in accordance with this Policy and, where appropriate, with organisation-specific procedures agreed with the BMZ. They also support DEval or BMZ evaluations as the entry point for access to data and documents and ensure access within a reasonable period of time. In addition, they help to ensure that evaluations are used by their respective organisations in several ways: for improving design and implementation of development interventions, for dialogue with partner organisations as well as for accountability vis-à-vis the BMZ.

4.2 DIFFERENT CHANNELS FOR IMPLEMENTING DEVELOPMENT COOPERATION – ADAPTED APPROACHES FOR EVALUATION

The BMZ implements its development policy objectives through three essentially **actor-specific channels** of development cooperation. These are reflected appropriately in the German evaluation system:

- Official bilateral cooperation with partner countries ([Section 4.2.2](#))
- Cooperation with civil society and municipalities ([Section 4.2.3](#))
- Cooperation with international organisations and the EU ([Section 4.2.4](#))

The **thematic focussing** on the BMZ's core areas of cooperation takes place firstly within the three actor-specific channels or pathways, using the relevant instruments. Secondly it is achieved through so-called **special initiatives and transitional development assistance** ([Section 4.2.5](#)), which can be realised through all three actor specific implementation channels. Another case of thematic focussing is e.g. cooperation with the **private sector**, through which private funds and in some cases expertise can be mobilised ([Section 4.2.6](#)).

This section will outline roles and responsibilities of key evaluation actors and of functional units of the BMZ as well as procedures (including selection and financing of evaluations) for evaluating development interventions implemented through various channels (or pathways). The guiding principle is to achieve consistency between channels while allowing for appropriate variations.

4.2.1 General aspects for all channels

- (1) **Evaluations of individual development interventions (project evaluations)** are generally the task of the development organisations operating with BMZ budget funds (see [Chapter 2](#)). These organisations manage evaluations on their own responsibility in line with this Policy (see previous chapters and further details in the following sections). Evaluations are (co-)financed from the same budget items as the interventions under review – whether as part of the financing of a development intervention or as a stand-alone measure.

The BMZ promotes and calls for an **appropriate and systematic selection** of evaluations of individual development interventions within both the actor-specific and the cross-actor implementation pathways. The selection criteria and procedural details are based on the two basic functions of evaluations (learning and accountability) and are adapted to the specific characteristics of the budget item or

organisation (see below for further details). The selection can rely either on a representative (mostly random) or a purposive sampling, taking into account a sufficient coverage of BMZ funding as well. In any case, conducting an evaluation has to be considered where the intervention is extensively funded by the BMZ and has a larger scope, including prospective scope, e.g. in the case of pilot measures. Other key reasons include an expected major gain in knowledge (e.g. also in the case of identifiable risks) in relation to the costs of an evaluation.

In the future, to support the **quality assurance** of project evaluations and, where appropriate, strategic evaluations conducted by implementing organisations and civil society organisations, cross-organisational **DEval** meta-evaluations will be carried out regularly, in addition to the development organisations' own procedures. These will be based on an analytical grid derived from recognised standards – especially those of OECD DAC and DeGEval – and will reflect the quality standards of this Policy appropriately.

- (2) Strategic evaluations are generally the task of DEval. In selected cases at mid-level, they are also conducted by development organisations (see below for further details). If strategic evaluations are highly policy-relevant (such as evaluations of instruments and those designed across implementation channels or across organisations) they always fall within the remit of DEval or, in exceptional cases, of the BMZ. The former are financed institutionally (core contributions), the latter from the special budget items earmarked for them.

DEval's selection criteria are primarily geared to the political and strategic relevance for its intended main users, while taking into account both risks, and potential for learning and innovation. The selection criteria also aim to achieve an appropriate degree of coverage of the BMZ portfolio in the longer term. These selection criteria, leaning more towards the learning than the accountability function, are also applied by BMZ for aggregating and prioritising its proposals for DEval's evaluation plan.

- (3) Cross-ministerial evaluations offer the opportunity for joint learning by the ministries involved, and thus promote policy coherence in cases where two or more ministries meet or overlap in their respective engagement (see also [Section 5.1](#)). Key interfaces between development policy and other policy areas include in particular the Humanitarian-Development-Peace Nexus, and climate and environmental issues. Especially when other ministries are engaged significantly and additional knowledge is expected to be gained, the BMZ aims for cross-ministerial evaluations. These may be managed by DEval, or by the ministries concerned with DEval participating in implementation, see also [Section 5.1](#) and the Guidelines for joint ministerial evaluations ([BMZ 2021c](#)). Joint ministerial evaluations are usually co-managed by BMZ-E.
- (4) For details **on the evaluation process**, including the involvement of partner organisations, the reporting procedure, publication, and follow-up by the responsible organisation, see [Section 3.2](#).

4.2.2 Official bilateral cooperation with partner countries

Cooperation with **partner countries** is the core of BMZ's development policy and accounts for the largest share of its budget. BMZ country strategies, with thematic development objectives agreed between governments, provide the strategic framework. These strategies are usually operationalised through (internal and cross-organisational) thematic "development cooperation programmes". These programmes comprise mutually complementary individual interventions ("projects") implemented by Germany's official implementing organisations (BGR, GIZ, KfW and PTB) with project executing agencies in partner countries.

Evaluations by the implementing organisations need to comply with the principles and standards of this Policy (see especially Chapters 2 and 3) and the guidance below.

- (1) **Evaluations of individual development interventions** ("projects") form the broad and institutionalised foundation of independent analysis and assessment. The **central evaluation units of the implementing organisations** conduct evaluations of a representative **sample** of completed and, where appropriate, ongoing projects in accordance with this Policy. The selection criteria and procedures for the sampling need to be agreed with the BMZ. The focus is on final and ex-post evaluations, since outcomes, development impact, and sustainability can usually only be soundly assessed after completion of the support. Evaluations of completed phases or components of ongoing programme-type projects (with sufficiently similar individual interventions) can provide valuable insights for further phases and will be further pursued. The sample population comprises budget-financed interventions as well as so-called development loans (mixed financing). The inclusion of other interventions financed with market or third-party funds, for which BMZ bears political (co-)responsibility, is envisaged in the medium term. In addition to the sample, the implementing organisations will also include requests from the BMZ in their evaluation plan.

The implementing organisations will inform BMZ-E and DEval in good time and annually about their **evaluation plan**, comprising the annual sampling, and any other evaluation projects (see also [2]). Amendments during the current year will be communicated without delay.

Insofar as the (project-type) intervention to be analysed includes a policy or strategic orientation or component, the evaluation units of the implementing organisations will enable timely involvement of the BMZ during evaluations, based on milestones (see Section 3.2).

Duplicating evaluations of **joint programmes** funded by different donors must be avoided. If no joint evaluation by the development partners is planned, evaluations of **co-financing arrangements**, which are **German financial contributions** to development interventions carried out under the lead responsibility of an international organisation or the EU or other donors, will be conducted as far as possible by the implementing organisations as desk reviews, based on the analyses prepared by the lead organisation, or on existing data. Evaluations of **co-financing contributions from other donors** or an international organisation or the EU to interventions that are led or implemented by German development cooperation (sometimes referred to as "combined financing" arrangements) will take into account, as far as possible, the evaluation needs of the co-financing donor.

To ensure comparability, the implementing organisations use standardised (but adaptable) **evaluation questions** (see Guidelines for using evaluation criteria, BMZ 2021b) as well as standard benchmarks ('rating scales'). Given the varying importance of the six DAC evaluation criteria with regard to an overall assessment of the development effectiveness, they also use standard 'knock-out criteria' that mark

out certain evaluation criteria (effectiveness, impact and sustainability) as key for the overall rating. To make these evaluations more useful at the strategic level, they will analyse and assess, inter alia, the **contribution** of the project to the (higher-order) development policy objectives specified by the BMZ. They also analyse or reconstruct the underlying **intervention logic** – preferably the ToC – with a view to the project’s suitability for achieving the development objective. To improve verifiability, the assessments are based on an **evaluation matrix** and the data and documents used are at least documented internally. The implementing organisations use harmonised **reporting formats** to facilitate synthesis products, and to ensure uniform external presentation. Regardless of the form of publication, they send the complete reports to the BMZ in accordance with the defined distribution list.

In the long term, the **BMZ and the implementing organisations** intend to systematically strengthen **rigorous impact evaluations** (see [Section 5.2](#)).

- (2) **Analyses at the strategic level**, such as the aggregate achievement of development policy objectives at country level, are carried out at intervals by **BMZ** internally, sometimes supplemented e.g. by DEval country portfolio reviews. In the future, DEval evaluations will complement the internal reviews more systematically. Commensurate with a country-level view, they will in particular also assess impact, defined as the extent (scale, reach or significance) of higher-level effects generated or likely to be generated by the country portfolio. This assessment of contributions to higher-level changes will also include effects on both structural changes (systems or norms) and/or effects on changes in people’s well-being. It will also consider the extent of the respective effects. These evaluations will either focus on country portfolios as a whole, or have a focus on selected thematic core areas of the BMZ, aiming for thematic cross-country syntheses. The selection will take into account different country categories. Depending on the purpose, the evaluations should preferably also include other development interventions with a relevant country component, which are not defined (in Germany) as “official bilateral cooperation”, particularly the projects and programmes of the special initiatives (see [Section 4.2.5](#)) and earmarked contributions to international organisations (see [Section 4.2.4](#)).

Some **implementing organisations**, while respecting the division of labour with BMZ and DEval, conduct “**mid-level strategic**” **evaluative studies** that concern their organisation. These include *joint evaluations* focussing on the performance of several implementing organisations (see below). They also include, *organisation-specific, cross-cutting syntheses*, i.e. thematic or instrument-specific analyses based on several project evaluations of the respective organisation, sometimes combined with complementary field studies. These allow valid conclusions to be drawn for the future operational design or implementation of development cooperation. They thus contribute substantially to institutional learning in their respective organisations. At the same time, they can make an evidence-based contribution to BMZ thematic strategies (see [Section 5.2](#)). The implementing organisations proactively notify the BMZ (BMZ-E and the functional units) and DEval about their planning and conceptual design, and enable the BMZ to feed its knowledge and information needs into the process in good time.

Harmonised evaluation practice facilitates the **piloting of joint evaluations of development cooperation programmes or parts thereof by the implementing organisations**, involving shared but differentiated responsibilities. The conceptual approach and selection of case studies for these evaluations are discussed and agreed on between BMZ and the implementing organisations, while DEval has an important advisory role. BMZ’s decision about the future course of action will be based on the experience gained from piloting this new evaluation instrument, taking into account its usefulness for functional units for managing development cooperation programmes as well as its fit in the evaluation system’s architecture and its efficiency.

4.2.3 Cooperation with civil society and municipalities

A strong civil society plays a crucial role as a driver of sustainable development. Municipalities possess know-how and networks that are relevant for solving global challenges. The BMZ therefore provides substantial funding for civil society and municipal engagement in partner countries, and for development education and engagement in Germany. Funding is usually provided through German host agencies.

Evaluations managed by civil society (host) organisations (CSOs) are to be guided by this Policy (see [Chapter 2](#) and [Section 3.2](#) in conjunction with [2.3](#)).

- (1) In the **selection of CSO evaluations of individual development interventions**, the focus is usually on learning objectives of the development organisations and the BMZ. Over an extended period of time, they should cover relevant parts of the portfolio of budget items.

Financing of evaluations by CSOs is usually provided as part of project expenditure. In addition, earmarking funds of budget items are feasible and tend to work better as an incentive to conduct evaluations. They also allow financing of e.g. ex-post evaluations, rigorous impact evaluations, and evaluation syntheses (where appropriate also across agencies) as well. The BMZ will explore this option further.

To support **evaluation capacity development**, the BMZ also promotes (host) CSO interventions aiming at strengthening partner organisation capacities for monitoring and evaluation (M&E), as well as the empowerment of (end-) beneficiaries for monitoring governmental and non-governmental actions, either as a project component or as a stand-alone intervention.

BMZ grants for individual development interventions sometimes comprise a substantial *volume* and a substantial *share of the total funding* of development interventions with a host agency. In such cases, for *larger* development organisations, organisational and procedural arrangements for systematic evaluations (including evaluation system, see [Section 3.1](#)), based on this Policy, will be used by the BMZ as a key criterion in deciding whether to continue funding development interventions.

- (2) At the **strategic level** the responsibility for evaluations is determined by the designated purpose of the funding guidelines or the budget item provisions. Furthermore, development interventions of CSOs may be included in strategic evaluations of DEval, including meta-evaluations, or BMZ evaluations (see [Section 4.2.1](#)).

In connection with the **promotion of the development engagement of selected CSOs** (churches and political foundations), their central evaluation units will ensure a systematic selection of strategic evaluations in consultation with the BMZ. The main focus of these is on the learning objectives of the organisation, with an appropriate degree of coverage of the development interventions supported by the BMZ. Evaluation policies of CSOs take BMZ-specific procedures into account. The development organisations involve the BMZ at the predefined milestones along the evaluation timeline (see [Section 3.2](#)). Policy-relevant conclusions and their follow-up will be part of the agenda of the regular institutional dialogue between BMZ and the respective CSO.

Larger funding programmes (which provide access to funding for projects for a wide range of CSOs) – or relevant parts of them – will be evaluated even more systematically and will cover an extended period of time in the future, in addition to internal analyses. If a funding programme is not evaluated by DEval, the **BMZ** (functional unit) will arrange for an evaluation and manage it in accordance with the

process and product standards (Section 3.2). The participation of the relevant CSOs in the reference group and appropriate quality assurance are of particular importance here. Here, too, the focus may be on learning objectives, e.g. by setting thematic priorities (see also Section 2.3, results and efficiency reviews).

For **Engagement Global (EG)** this Policy is binding. Its central evaluation unit will support BMZ evaluations of *EG-administered funding* programmes through all phases and assist with quality assurance. In addition, within the framework of the EG-evaluation system, the evaluation unit will ensure systematic evaluations of larger *EG-managed programmes*.

For **institutional funding** (core contributions) separate procedures apply that take into account federal regulations on results and efficiency reviews (see Section 2.3) as well as the purpose of the organisation, such as the evaluation procedures that are customary for research institutions.

4.2.4 Cooperation with international organisations and the EU

Multilateral organisations are key to a rules-based international order, which the BMZ supports and helps shape as part of its ministerial responsibility. These and other international organisations (see BMZ glossary for key terms; hereinafter, for the purposes of this Policy, no further distinction is made) offer comparative advantages for the implementation of development interventions. This is due to the high legitimacy they possess as a result of their broad membership and political neutrality, in conjunction with expertise and in some cases capital.

BMZ funding includes (1) **core funding** of selected international organisations (assessed contributions and shares, as well as non-earmarked, voluntary contributions, combined here for the purpose of evaluation, and (2) **earmarked contributions**, whose purpose ranges from strengthening the work of an international organisation (IO) with respect to an important development-policy theme (institutional objectives), to using its comparative advantage as an implementing organisation (thematic or country-specific objectives).

Since IOs have their own decision-making, supervisory and administrative structures, **the BMZ acknowledges and uses their mechanisms and procedures, also for evaluation**, provided they follow OECD DAC, UNEG or ECG standards.

- (1) With **core funding**, the BMZ exercises its membership rights by participating in the supervisory or steering body of the organisation concerned.

In its work in these bodies and as part of its bilateral institutional policy dialogue, the BMZ actively supports the work of central, independent evaluation units. Alternatively, especially in the case of larger organisations, it advocates establishing such units. It aims to strengthen the evaluation function, ensure the independence and quality of the work, and make sure that the evaluation results and recommendations are used. The BMZ supports, and where appropriate calls for, the adoption (and at long intervals the updating) of an appropriate evaluation policy.

For **continued core contributions**, the BMZ also bases its funding decisions on the presence of an appropriate system for reporting (sometimes referred to as “self-evaluation”) as well as for independent evaluation (central, independent evaluation, sometimes supplemented by decentralised evaluations). Systems’ reviews of evaluation (or parts thereof) are provided by IO-Peer Reviews, assessments by the United Nations Joint Inspection Unit and by the Multilateral Organisation Performance Assessment Network (MOPAN). Some IO evaluation units conduct assessments of the quality of decentralised evaluations. BMZ uses this information for its decision making.

(2) **Earmarked contributions** range from the funding of individual projects to the funding of joint institutional or joint donor programmes, or trust funds of international organisations.

(2.1) An appropriate selection of **evaluations of individual development interventions** may already be ensured by an existing evaluation system of an IO or the EU. Either by (usually decentralised) project evaluations, or – depending on the designated purpose of the financial contribution – by (central) strategic evaluations (e.g. country programme evaluations). Where these are not available or planned (possibly also by other donors) in sufficient numbers or are not of sufficient quality, in the future the **BMZ** will ensure a systematic and appropriate selection of evaluations preferably to be conducted by IOs. Depending on BMZ’s information needs, this selection can be thematically based, organisation-specific or cross-organisational. It will cover relevant parts of earmarked contributions, including co-financing arrangements through implementing organisations, over an extended period of time. The BMZ’s framework or project agreements with IO will stipulate the option for BMZ getting involved at milestones of evaluations. Where interventions are co-funded by several donors (e.g. in the case of multi-donor trust funds), evaluations of IOs should preferably permit appropriate participation by all donors. Where **co-financing** through implementing organisations is selected as a form of funding, see the corresponding paragraph in [Section 4.2.2](#).

In cases where **several federal ministries** are engaged with the same IO to a significant extent, the BMZ – provided that it has ministerial responsibility for an organisation within the German government – will strive to ensure a coordinated approach with a view to policy coherence.

In addition to co-financing project or programme evaluations by IOs, innovative approaches that support further strategic development of the evaluation function of an IO may also be financed, particularly where the IO is also being or will be funded through core contributions.

(2.2) **At the strategic level**, analyses of the aggregate appropriateness and effectiveness of earmarked contributions with respect to BMZ policy objectives, whether thematic, country-specific or institutional (organisation-specific), are usually conducted by the BMZ internally. In addition, independent evaluations can add value, e.g. by including earmarked contributions and co-financing arrangements in DEval country programme evaluations (see [Section 4.2.2](#)) or joint ministerial evaluations.

In IOs **with no separate independent evaluation unit of their own** as defined in this Policy, the **BMZ** works towards systematic, independent evaluations of the portfolio, with an appropriate governance structure to ensure independence, quality and usefulness. It also involves itself (or DEval) in these evaluations, depending on their strategic and financial significance for the BMZ.

(3) With the **EU as a supranational organisation**, the BMZ works at the institutional level to support an appropriate evaluation system along the lines of core financing contributions to international organisations. Concerning (“earmarked”) co-financing arrangements for individual interventions, ranging from projects to comprehensive programmes, see [Section 4.2.4](#) on official bilateral assistance. In the case of joint programming and joint implementation, the BMZ participates in EU evaluations appropriately. Conversely, in (German) country portfolio evaluations DEval takes into account the respective procedure agreed between the EU and the member states.

4.2.5 Thematic cooperation – special initiatives and transitional development assistance

Special initiatives are a BMZ instrument for strengthening selected priority development-policy themes over a long, though usually limited, period of time. By contrast, transitional development assistance is one of the BMZ's permanent tasks. It is a crisis management instrument in its own right, designed to help boost the resilience of particularly affected people and local structures flexibly, but in the medium and long term. Both instruments use all three actor-specific channels for implementing their respective strategies (see Sections 4.2.1 to Section 4.2.4).

- (1) The procedure for **selecting evaluations of individual development interventions** follows the rules for actor-specific channels contained in this Policy (see relevant paragraphs above). However, the population for sampling can be aligned with the respective budget item. Where “project” funding is provided for global (multi-country) programmes, BMZ has to be involved by development organisations early on.

Concerning **evaluation criteria and procedures**, assessing risks or *effects of exit* and *connectedness* (as special features of the criterion “sustainability”) are especially important in all special initiatives and in transitional development assistance. *Conflict sensitivity* needs to be observed particularly in transitional development assistance (see Chapter 2).

- (2) At the **strategic level**, empirically robust and transparent analyses of the aggregated results of individual development interventions in relation to development policy objectives of the special initiatives will be carried out at intervals in the future, if appropriate separately for specific fields of action. For questions that go beyond analysing the extent of results achieved, strategic evaluations of special initiatives can add value.

In crisis countries and conflict settings, German development cooperation is increasingly becoming involved in implementing the Humanitarian-Development-Peace (HDP) Nexus. This is designed to manage crises sustainably, while mitigating humanitarian needs by improving coordination, cooperation and communication of collective objectives and results. The HDP Nexus offers various entry points for strategic evaluations, including across countries, partners, and ministries. These should possibly be designed and implemented in the next few years.

4.2.6 Cooperation with the private sector

The international community's ambitious Sustainable Development Goals can only be achieved together with the private sector, and the expertise, innovative capability, and potential funding it has to offer.

The BMZ therefore supports development interventions, using its actor-specific implementation channels, to promote the private sector *in partner countries* – directly or indirectly, addressing financing bottlenecks or policy frameworks, or by using private-sector solutions in connection with public tasks. For the evaluation of these approaches, the principles, standards, and evaluation criteria apply without restriction – as always, as appropriate to the object of the evaluation.

In addition, BMZ promotes the engagement of the private sector *in Germany, Europe or internationally* (through IOs) for reaching international development goals (“private sector for development”). Working with other ministries, it uses a broad range of approaches and financing instruments to mobilise private sector investment, loans and equity, as well as expertise.

Evaluations have to take into account special characteristics of private sector financing, without losing sight of development goals and public responsibility. Compliance with human-rights-based due diligence obligations, and social and environmental safeguards, is particularly important in this context. Quality standards have to be applied thoughtfully, for instance by balancing the usual transparency standards with legitimate confidentiality concerns of companies. Some **evaluation criteria** require further work to appropriately reflect the subject matter (see [OECD DAC 2019](#) and [BMZ 2021b](#)). For instance, when assessing *efficiency*, an additional dimension should be assessed, namely whether and to what extent the use of official funds has led to an additional, sustainable mobilisation of private investment (so-called financial additionality). There is also a need for clarification when assessing **impact** with respect to relevant fields of observation/dimensions of results (so-called development additionality). The criterion of **coherence** can also be interpreted differently than in the case of purely official interventions.

Internationally, the debate on a joint OECD DAC-wide understanding of the evaluation criteria for private sector engagement supported by public financing is still unfolding. Given the major importance of this policy area for achieving the Sustainable Development Goals, in its capacity as a method developer DEval will continue to participate in international discussions with the aim of developing international standards. As the debate unfolds, DEval will also help to adapt the evaluation criteria of German development cooperation together with knowledge bearers at the BMZ and the implementing organisations.

The selection of and responsibilities for evaluations of individual development interventions, and evaluations at the strategic level, need to comply with the general and actor-specific rules, respectively.

5. Implementation and the way forward

Evaluation thrives on interaction: on dialogue and cooperation with evaluation partners (Section 5.1) and on exchange and cooperation within organisations, with policy units and the evaluation unit performing complementary roles for the common objective of enhancing development effectiveness. (Section 5.2). In addition, to keep up with a changing environment, evaluation needs to adapt over time. Finally, the evaluation system – and this Policy – should be independently reviewed at extended intervals (Section 5.3).

5.1 DIALOGUE AND COOPERATION WITH PARTNERS

National and international networks and partnerships are proven fora for sharing lessons learned, with the aim of joint learning and further development of good practice. They are complemented by thematic collaborations and working groups to address specific issues.

... with development organisations and other federal ministries in Germany

The **annual meetings of BMZ-E with the evaluation units of the German implementing and civil society organisations and DEval** are well established and particularly important for the evaluation system. These meetings are organised by one of the development organisations or by DEval, in consultation with the BMZ. This valuable tradition will be maintained in the future.

Joint ministerial approaches are gaining importance. These aim to support and improve the overall effectiveness of the federal engagement across ministerial boundaries, with a view to policy coherence for development (EU 2017). In partnership with the Federal Foreign Office and the Federal Ministry for the Environment, Nature Conservation, Nuclear Safety, and Consumer Protection the BMZ will therefore continue to pursue **annual meetings of the interministerial consultative group on ODA evaluation**. In addition to exchange, it will be important to achieve sufficient harmonisation of standards and procedures. Furthermore, in appropriate cases the BMZ will continue to initiate **joint ministerial evaluations** or participate in **cross-ministerial evaluations** led by DEval (see Section 4.2 and BMZ 2021c).

... with other donors

The BMZ considers the **OECD DAC Network on Development Evaluation** as the most important international forum for harmonising norms and standards and knowledge sharing. Germany is represented by the BMZ together with DEval. BMZ and DEval, respectively, will continue contributing to DAC working groups on standards and guidelines involving methodological issues. BMZ also values informal donor networks, such as Nordic Plus and DACH (DE, AT and CH, as well as further evaluation units with at least passive German language skills), which are particularly conducive for a very open exchange of experiences and practices and joint learning.

Furthermore, BMZ-E together with DEval represents Germany in the expert group of **EU Heads of Evaluation Services (EUHES)**, which deliberates on the evaluation work of the European Commission. Where concerns of the implementing organisations are affected, BMZ-E involves them in advance.

International joint evaluations with other partners, especially of joint programmes, help to lower transaction costs, and are at the same time a good opportunity for knowledge transfer between the partners. BMZ-E participates in strategic joint evaluations by central evaluation units of other donors, in cases where the evaluation is of high political/strategic importance for the BMZ, involving DEval as well. In particular, these include selected EU evaluations (see [Section 4.2](#)). German development cooperation organisations also already make use of corresponding options for joint evaluations with their partners.

... with professional networks

Cross-policy professional networks support and are crucial for a broader perspective beyond development policy. The DeGEval Evaluation Society provides Germany-wide standards. Through its cross-policy conferences, it also provides a platform for dialogue among organisations commissioning evaluations, evaluators and researchers on the current state of the professional debate. The BMZ supports DeGEval as an institutional member, and in accordance with its role participates in the working group on development policy and humanitarian assistance. The European Evaluation Society (EES) and world-wide professional networks and conferences enable dialogue beyond national boundaries. Especially the latter also support exchange with actors from developing and emerging countries. BMZ-E uses these opportunities in accordance with its mandate.

... with partner countries

Evaluation Capacity Development (ECD) interventions support partner country ownership. ECD aims to enable partner countries to evaluate their own government programmes – including those co-financed by donors such as Germany – using their own capacities, and to use this for good governance. Building on the Aid Effectiveness Agenda, the 2030 Agenda for Sustainable Development places strong emphasis on the role of country-led evaluations. It also underlines the need to strengthen national capacities for monitoring and evaluation. National evaluation capacities are considered a prerequisite for implementing internationally agreed principles on ownership, managing for development results and mutual accountability.

The **BMZ** has been supporting ECD for many years in various ways. ECD has been enshrined as a core area of work in DEval's founding statutes and has been promoted either as a component of development interventions or as a free-standing intervention by the BMZ. To meet global ECD needs efficiently, while maintaining their respective roles and responsibilities, the BMZ and DEval will explore options to implement tried and tested methods and formats more closely together with other donors and international partners. Furthermore, ECD as a cross-cutting task also during evaluations should be used wherever possible as a guiding principle for all development evaluations, in the spirit of partner orientation (see also [Section 5.3](#)).

5.2 HOW EVALUATION IS INSTITUTIONALISED

To generate impact, evaluations should be embedded in a results-based management model and be accompanied by complementary measures.

Evaluation as part of the BMZ management model

Evaluation is an essential source of evidence. Embedding it in a systematically results-based, institutionalised way of working in development organisations enables it to perform this role effectively.

Internationally, the following approach has been tried and tested: a) consistent impact orientation in key decisions; b) use of appropriate instruments to utilize existing evidence and generate new evidence, with a particular role for independent evaluation; c) regular feedback of evidence into processes of decision-making as well as proper knowledge management and data management. An enabling framework for this is created by a culture of learning that (i) also allows for mistakes; (ii) sets incentives for adaptation; (iii) makes a point of using pilot interventions (as “experiments” that may involve a high rate of failure) with stringent and empirically verified analysis of lessons learned; and (iv) provides sufficient resources (financial resources and time, possibly also training measures) for those responsible for managing development resources, and for the collection and analysis of empirical data.

This also applies to the **BMZ**, which has further refined results-based management and the use of evidence as part of its BMZ 2030 reform process. In its portfolio management, the BMZ will continue to preserve, and as required strengthen, the traditional role of evaluation as an integral part of learning-oriented, evidence-based analysis of development effectiveness. It will also continue to promote the use of evaluation results at the BMZ. In this context, particular importance is attached to (i) preserving the independence of DEval as enshrined in its statutes; (ii) providing it with the resources it needs to fulfil its mandate, and (iii) ensuring the development-policy relevance of its evaluation plan and the relevance of recommendations contained in its reports for the strategic management of the BMZ portfolio.

In its regular **institutional dialogue** with the development organisations with which the BMZ primarily works, the BMZ will place (or continue to place) evaluation on the agenda, with the aim of maintaining or strengthening the evaluation function along the lines of this Policy, including the use of quality evaluations for learning in the organisations themselves as well as for accountability towards the BMZ.

Complementary tasks – Shared responsibility

As part of systematic results-based management the **BMZ’s** functional and policy units play a key role throughout the entire cycle of planning, implementation, M&E, learning and adaptation. They thus play an important complementary role with respect to independent evaluations.

(1) Ensuring evaluability during planning

The BMZ’s regulations (as well as their application) for planning, monitoring and reporting will serve to ensure the evaluability of **individual development interventions**. BMZ funding decisions have to be based on planning documents which contain a) operationalised (project or programme) *objectives* with SMART indicators at the level of results (usually the lower outcome level) that are consistent with the ministry’s strategic objectives; b) an analysis that clearly indicates the relevance of the intervention for reaching strategic objectives. This analysis also includes a logical *results framework* for the planned intervention that links different levels of results through appropriate results chains; c) an analysis of

sustainability as well as of possible *risks*; d) *costs and financing*. *Lessons learned* from similar interventions, and especially through evaluations, should be taken into consideration transparently.

The implementation of development interventions needs to be accompanied by appropriate **monitoring** (see [Section 2.2](#)) and corresponding **reporting** by the development organisation working with BMZ funds (see [Section 2.3](#)).

BMZ strategies and programmes include *strategic objectives*. These need to describe concretely the targeted (higher-level) development changes (upper outcome or impact level), taking into account the SDGs. Every effort should be made to operationalise **strategies** with indicators. However, given the wide variety and context dependency of the individual development interventions underlying the strategies, this will not always be possible. By contrast, time-bound **programmes** encompassing several development interventions and backed up with funds always need to have operationalised objectives and a logical results framework. Setting *milestones* of implementation (see [Section 2.2](#), monitoring) and conducting internal analyses of performance and results achieved to date support the management of interventions and improve evaluability, and therefore are to be used throughout.

In the future, functional and/or policy units will ensure even more systematically

- (2) **appropriate and early selection of evaluations** of individual interventions unless an established system is already in place (see [Section 4.2](#)). They will also ensure that they contribute to the selection of strategic evaluations during evaluation programming as well as to the selection of evaluation questions during the consultation in the design phase of an evaluation (see [Sections 4.1](#) and [3.2](#)). In particular, rigorous impact evaluations (see [Section 5.3](#)) or accompanying research of effects of interventions require early selection.
- (3) **appropriate support**, in particular for strategic evaluations, and **follow-up** on recommendations, (see [Section 3.2](#)), as well as support or participation in decentralised (project or programme) evaluations led by other federal ministries, IOs and the EU.
- (4) **Systematic use of evaluation results** including evaluation syntheses for shaping development policy, in addition to other sources of evidence. In the future, units responsible for strategies, such as country or thematic strategies, or policies, such as for BMZ (cross-cutting) quality criteria, will make greater and more active use of evaluations at an early stage before drawing up a new strategy.

5.3 FURTHER DEVELOPMENT OF EVALUATION AND QUALITY ASSURANCE

Good quality evaluation requires regular adaptation with regard to a) quality, including methodological quality; b) changing environment (context) for development cooperation and changes within development cooperation itself. New challenges include for instance stronger promotion of private sector engagement, the growing importance of policy coherence and more extensive normative requirements, e.g. concerning human rights (see previous chapters). The evaluation of German development cooperation must be responsive to these changes, while preserving its brand essence.

The BMZ is committed to pursue the following points in particular

- (1) **Greater use of digitalisation.** The availability of existing, standardised data sets in partner countries, collected by development partners and the scientific (and business) communities, is increasing steadily. More and more, this is also enabling the evaluating agencies to use quantitative methods. Furthermore, increasing digitalisation, also in partner countries, is offering opportunities for newer approaches that are particularly important in settings of acute crisis, and can help to reduce the carbon footprint. The use of technological innovations such as artificial intelligence – also in evaluations – is only just beginning. Development evaluation should make more intensified use of these possibilities. It is important to keep in mind, however, the key and continued importance of contextualisation (knowledge of local conditions and causal mechanisms) and qualitative methods especially in development cooperation and its evaluation. In line with one of its core areas of work, DEval in particular is called upon to help further develop appropriate methods and designs. Furthermore, the BMZ will work towards providing applied development research, which includes DEval evaluations, with access to standardised data sets collected in development cooperation (as “open data”).
- (2) **Systematic expansion of rigorous impact evaluations (RIEs).** In suitable cases RIEs, applying experimental or quasi-experimental designs and methods, not only provide particularly robust evidence of the effects of interventions, when combined with qualitative methods and embedded in a theory-driven approach. Selected thoughtfully, they are also an efficient (cost-effective) tool, despite their usually high costs. Pilot interventions, innovative programmes, or settings where little robust evidence is available internationally (evidence gaps), are cases in point. In the future the BMZ will therefore make greater use of RIEs in appropriate cases, and promote a more systematic and more efficient application of the approach, starting with better information sharing (e.g. public RIE data base). For further details see the forthcoming (internal) guidance document. In the medium term the targeted selection, promotion and quality assurance of RIE as well as improvement of their use will be further developed, clarifying the role of DEval.
- (3) **Piloting of more rapid instruments and mid-term assessments.** Rapid reviews (a familiar tool in humanitarian assistance), adaptive reviews and process reviews, which sometimes systematically examine only partial aspects, provide quick results for decision making. They can and should enrich the evaluation toolbox, provided that the brand essence (i.e. fundamental quality standards) of evaluations is retained.
- (4) **Improving the use of evaluations by communicating evaluation results more user-friendly and providing more syntheses.** BMZ and the development organisations have already taken several steps to bring evaluation results and lessons learned into practice. In addition to *implementation planning for recommendations*, these steps include the tried and tested involvement of functional units at the predefined milestones along the evaluation timeline (in-process learning) and various formats for communicating evaluation results after completion. However, the use of evaluation results can and should be enhanced considerably, if reports (containing a concise summary) are stored in a user-friendly IT-based system. In addition, evaluation units (and DEval) should also support the use of evaluations by *better processing evaluation results*, using more customised, user-friendly formats, including visualisation, to “translate” the findings.

Furthermore, evaluation units can help to improve conceptual, intra-organisational learning by providing cross-cutting syntheses based on their organisation specific project evaluations. Cross-organisational evaluation syntheses (or synthesis studies) by DEval that preferably also include robust insights from outside German development cooperation offer major system-wide potential for learning. In the future, they need to be more systematically incorporated into BMZ’s thematic or country strategies or cross-cutting policies and considered during their implementation as well.

- (5) **Recalibrating partnership.** Joint engagement by development partners (“donors”) has to be matched by evaluations, with the aim of supporting policy coherence for development. In addition, today’s understanding of partnership requires a shift in attitudes towards partner country organisations. This means strengthening partner-country ownership and capabilities, including evaluation capacity. The BMZ will increasingly require that a) BMZ-supported development interventions *integrate M&E*; b) in German evaluations, *partner-country organisations* are consistently involved at the predefined milestones along the timeline of an evaluation and *professional evaluators from the region* are regularly deployed; and c) specific evaluation capacity development interventions continue to address the *supply and the demand side* for evaluations. Capitalising on the extensive knowledge gained on the German side, BMZ-E and DEval will pursue ECD interventions preferably in concert with other donors and partners.
- (6) **Completing the components of the evaluation system, and supplement them with complementary measures.** Many development organisations and the BMZ already have good practices for institutionalising quality evaluations in place. However, there are unmet needs in both the actor-specific and cross-actor implementation pathways (see [Section 4.2](#) in conjunction with [Chapters 2 and 3](#)). Going forward, these needs should be studied more closely and in the future be met appropriately. The complementary measures will include better linking of monitoring and evaluation (and further sources of evidence), so as to better use their potential as impact- and risk-oriented learning instruments, use them more coherently and systematically for managing development cooperation and reflect on their respective use early on, i.e. already in the planning process.

To implement these elements, the BMZ will continue or strengthen the well-established institutional dialogue on several levels with DEval and the development organisations (see [Sections 4.2.2, 4.2.3 and 4.2.4](#)). Given its importance, BMZ-E is institutionalising in particular the joint work with the official implementing organisations in the *working group on evaluation*, in which DEval is involved as an advisor. In addition to regular exchange on current issues of mutual concern, the working group aims to systematically address issues to further improve coherence, the division of labour and the functionality of the implementing organisations’ evaluations. The BMZ will also systematically continue the *dialogue with civil society organisations* (e.g. regular dialogue with VENRO, the umbrella organisation of development non-governmental organisations in Germany) accordingly.

Ensuring the quality of evaluations and reviewing the evaluation system

Several elements of systematic quality assurance and review are already in place or are increasingly being applied, some by development organisations themselves. Cross-organisational meta-evaluations and the evaluation system’s monitoring by DEval provide an external perspective, complemented by occasional institutional evaluations by or of DEval. In the future, DEval will include quality standards set out in this Policy as a reference for its evaluations. Furthermore, regular DAC Peer Reviews include an assessment of several of the system’s standards. In addition, as in the past, external, comprehensive independent system reviews encompassing all of the major German actors in the evaluation system should be conducted at extended intervals, possibly by peer review. This will require a governance structure that is carefully selected in order to avoid conflicts of interest and ensure professional quality.

This Policy itself will also be put to the test in this context.

6. Literature

Accra (2008): Accra Agenda for Action.

<https://www.oecd.org/dac/effectiveness/45827311.pdf>

ALNAP (2006): Evaluating Humanitarian Action Using the OECD DAC Criteria.

<https://evaluation.msf.org/sites/default/files/2021-12/Evaluating%20humanitarian%20action%20using%20the%20OECD%20DAC%20criteria.pdf>

BMZ (2021a): Glossar der Schlüsselbegriffe im Bereich Evaluierung der Entwicklungszusammenarbeit. [(BMZ-)Glossary of Key Terms for German Development Evaluation; German only]

<https://www.bmz.de/resource/blob/92890/glossar-evaluierung.pdf>

BMZ (2021b): Evaluierungskriterien der bilateralen Entwicklungszusammenarbeit; Orientierungslinien des BMZ. [Evaluation criteria for German bilateral development cooperation; BMZ guidelines. German only]

<https://www.bmz.de/resource/blob/92894/evaluierungskriterien.pdf>

BMZ (2021c): Ressortgemeinsame Evaluierungen; Orientierungslinien des BMZ.

[Joint ministerial evaluations. BMZ guidelines. German only]

<https://www.bmz.de/resource/blob/92898/ressortgemeinsame-evaluierungen.pdf>

BMZ (2020): Reform Strategy “BMZ 2030”

Bundesregierung (2020a): Bundeshaushaltsordnung (as at: April 2020). [Federal Budget Code; German only]

Bundesregierung (2020b): Allgemeine Verwaltungsvorschriften zur Bundeshaushaltsordnung. Herausgeber: Bundesregierung (in Kraft gesetzt durch BMF-Rundschreiben) (as at: April 2020). [General Administrative Regulations on the Federal Budget Code; German only]

Busan (2011): Busan Partnership for Effective Development Co-operation.

<https://www.oecd.org/dac/effectiveness/49650173.pdf>

BWV (2013): Bd. 18: Anforderungen an Wirtschaftlichkeitsuntersuchungen finanzwirksamer Maßnahmen nach § 7 Bundeshaushaltsordnung. [German only]

BWV (2016): Bd. 10, 2., überarbeitete Auflage: Prüfung der Vergabe und Bewirtschaftung von Zuwendungen. [German only]

DeGEval (2016): Standards für Evaluation; 1. Revision. [German only]

https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards_fuer_Evaluation.pdf

EU (2017): Joint statement by the Council and the representatives of the governments of the Member States meeting within the Council, the European Parliament and the Commission: “THE NEW EUROPEAN CONSENSUS ON DEVELOPMENT” (2017/C 210/01).

<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:C:2017:210:FULL>

IEG and DAC Evalnet (2008): Sourcebook for Evaluating Global and Regional Partnership Programs.

<https://openknowledge.worldbank.org/handle/10986/6601>

OECD DAC (1991): Principles for Evaluation of Development Assistance.

<https://www.oecd.org/dac/evaluation/dcdndep/41029845.pdf>

OECD DAC (2002): Glossary of Key Terms in Evaluation and Results Based Management. (English/French/Spanish)

<https://www.oecd.org/dac/evaluation/2754804.pdf>

OECD DAC (2006): Evaluation Systems and Use: a Working Tool for Peer Reviews and Assessments.

<https://www.oecd.org/dac/evaluation/dcdndep/35857765.pdf>

OECD DAC (2009): Glossar entwicklungspolitischer Schlüsselbegriffe aus den Bereichen Evaluierung und ergebnisorientiertes Management. [German translation of the Glossary 2002]

<https://www.oecd.org/dac/evaluation/dcdndep/43184177.pdf>

OECD DAC (2010): Quality Standards for Development Evaluation.

<https://www.oecd.org/development/evaluation/qualitystandards.pdf>

OECD DAC (2019): Better Criteria for Better Evaluation; Revised Evaluation Criteria Definitions and Principles for Use.

<https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>

OECD DAC (2021): Applying Evaluation Criteria Thoughtfully.

<https://doi.org/10.1787/543e84ed-en>

Paris (2005): Paris Declaration on Aid Effectiveness.

https://www.oecd-ilibrary.org/development/paris-declaration-on-aid-effectiveness_9789264098084-en

UN (2015): Transforming our world: the 2030 Agenda for Sustainable Development; Resolution adopted by the General Assembly of the United Nations, A/RES/70/1.

https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf

UNEG (2016): Norms and Standards for Evaluation.

Impressum

PUBLISHED BY THE

Federal Ministry for Economic Cooperation
and Development (BMZ)
Division Public Relations, Digital Communication

CONTACT

poststelle@bmz.bund.de
www.bmz.de

EDITED BY THE

BMZ, Division Evaluation, Development Research, DEval, IDOS

AS AT

October 2023

ADDRESSES OF THE BMZ OFFICES

→ BMZ Bonn

Dahlmannstraße 4

53113 Bonn

Tel. +49 228 99535-0

Fax +49 228 9910535-3500

→ BMZ Berlin

Stresemannstraße 94 (Europahaus)

10963 Berlin

Tel. +49 30 18535-0

Fax +49 228 9910535-3500