RESULTS ORIENTATION AND EVALUABILITY OF DEVELOPMENT COOPERATION PROGRAMMES

Experiences from the sustainable economic development programmes in Myanmar, Mozambique and Egypt

Executive Summary

2021

IMPRINT

This is an excerpt from the publication "Wirkungsorientierung und Evaluierbarkeit von EZ-Programmen. Erfahrungen aus den Programmen zur nachhaltigen Wirtschaftsentwicklung in Ägypten, Mosambik und Myanmar". Download the full report here:
https://www.deval.org/en/evaluations/our-evaluations/results-orientation-and-evaluability-of-development-cooperation-programmes-experiences-from-the-sustainable-economic-development-programmes-in-myanmar-mozambique-and-egypt

EXECUTIVE SUMMARY

Background and objectives

The **Joint Procedural Reform (Gemeinsame Verfahrensreform, GVR)** enacted by the Federal Ministry for Economic Cooperation and Development (BMZ) in 2017 redefined the procedures to be followed by the Ministry and its governmental implementing organisations[1] in their German development cooperation (GDC) activities. The GVR provides specific procedural steps for use by GDC actors at the **three triad levels (country strategy, programmes and modules)** for planning, implementation and steering purposes. The aim is to enhance the impact, efficiency and effectiveness of bilateral development cooperation while strengthening the BMZ's steering capacity (Hartmann et al., 2019; OECD DAC, 2021)[2].

GDC programmes are supposed to pool the specific measures (modules) managed by the implementing organisations. By forming an interface between the programme and module levels, they operationalise the BMZ strategic objectives set out in the country strategy (BRH, 2021; Hartmann et al., 2019). The objective of GDC programmes is to contribute to high-level development strategies to initiate change in partner countries (OECD DAC, 2021). The GVR increases the importance attached to the programme level, especially for steering purposes, and is intended to improve GDC programme design and implementation. Although the GVR stipulated programme-level evaluations, they have not yet been conducted.

This report examines whether and to what degree **GDC programmes are oriented towards results. Special attention is given to GDC programme design, reporting and evaluability**. Further matters of interest include the specific demand for and the purpose of future GDC programme evaluations and the most suitable timing for such evaluations. With the insights generated, this report aims to contribute to improved GDC programme planning and steering while providing content-based and methodological guidance for the GDC programme evaluations to come.

Results orientation

To ensure that GDC programmes are adequately oriented towards results, the **intended results and necessary resources** need to be assessed from the planning and design phase onwards. To address those aspects, proper procedures for verifying implementation progress, identifying where improvement is necessary and initiating corresponding optimisation measures throughout the implementation process are necessary. Such an approach can also be referred to as **results-based management** (RBM) (Bergmüller and Quiring, 2019; Caspari, 2012; Harten and Petrova, 2018; High Level Forum, 2005; Holzapfel, 2014; Klingebiel, 2011; OECD, 2019; UNDG, 2011). No comprehensive RBM system has yet been established in bilateral GDC, including within GDC programmes.

The following aspects are key components for an effective orientation towards results:

1. **Definition of objectives and indicators** in accordance with the SMART criteria (**S**[pecific], **M**[easurable], **A**[chievable], **R**[elevant] and **T**[ime-bound]; Lehmann et al., 2020)
2. **Identification of baseline and target values** as part of programme design, with regular **monitoring** of current values during programme implementation
3. Sufficient data available for verification of programme progress and achievement of objectives
4. Close cooperation with development partners by, for example, exchanging data and working together to define objectives (Holzapfel and Römling, 2020)

---

[1] The governmental implementing organisations are the Federal Institute for Geosciences and Natural Resources (BGR), Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), KfW Development Bank (KfW), the Physikalisch-Technische Bundesanstalt (PTB) and Engagement Global. For further information on these implementing organisations, see the evaluation guidelines published by the BMZ: *https://www.bmz.de/resource/blob/92884/08507d1204d093141b5f00bf5cbb8db7/bmz-leitlinien-evaluierung-2021.pdf* .

Cause–effect relationships can be illustrated via a **programme theory**. As programme theory allows for the systematic definition and presentation of objectives at all impact levels, it also serves as a programme management tool (Caspari, 2012; Holzapfel, 2014). Subsequent **monitoring activities** are essential for measuring progress and should provide current data (such as current values) for each programme indicator defined during the design phase.

How should the intended results of GDC programmes be achieved?

The country strategy defines the high-level development policy objective (long-term impact level) that GDC programmes are intended to achieve. However, GDC programme implementation is conducted entirely through individual modules without any additional or **overarching measures at the programme level**. Hence, the underlying assumption is that programme objectives are achieved (exclusively) through outcomes of and **interactions between its modules**. A key prerequisite for achieving objectives is therefore a **consistent and realistic definition of those objectives** at the programme level (BMZ, 2020) that takes into account the individual modules.

It is assumed that every individual module makes a specific and essential contribution to the programme objectives. **Modules may have a joint effect within a certain field of action** – which often includes a module for financial cooperation and one for technical cooperation – or collectively contribute to the achievement of a programme objective common to **several different fields of action**. Some modules may initially also have effects (and results) **in isolation from other modules** before developing joint effects with other fields of action at a higher level (medium-term impacts).

The **degree of interaction can vary greatly between modules**, which are usually implemented and steered by different implementing organisations, from logistical coordination, through joint planning to in-depth intertwining of activities to maximise synergy effects.

Evaluation and evaluability of GDC programmes

Neither GDC programmes nor country strategies are currently subject to systematic and regular evaluation. The module evaluations currently available contain only a few specific questions and little reference to the respective programme. In other words, module evaluations do not yield any significant findings regarding GDC programme relevance and effectiveness. A considerable **evidence gap** results, both **at the programme level** and **between the module and programme levels** (BRH, 2021), and important information is missing regarding the strategic relevance and effectiveness of bilateral development cooperation. Therefore, and with a view to **improving the steering and accountability** of GDC programmes, the Federal Court of Auditors (Bundesrechnungshof, BRH) and the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD) both recommend the introduction of evaluations at the programme level (BRH, 2021; OECD DAC, 2021).

To design and apply an evaluation format at the programme level, it is necessary to first **assess its evaluability**, which means examining whether GDC programmes meet the requirements for an evaluation. The relevant aspects should be consistent with the key components for effective results orientation, as described above, and include performance indicators, a monitoring system and the availability of data that is both adequate and reliable. This report examines the evaluability of programmes and is aimed at providing relevant information for anyone piloting and introducing a standardised evaluation format for GDC programmes.

Methodological procedure

The DEval evaluation team assessed the existing BMZ guidelines and specific programme documents for the three case studies selected: Egypt, Mozambique and Myanmar. Those GDC programmes were chosen for their relevance to the main focal area of German bilateral development cooperation – sustainable economic development. An online survey of BMZ country officers was also conducted and interviews were held with BMZ and implementation organisation staff involved in or in charge of the case study programmes,

along with experts and officers from other relevant departments, especially evaluation units. A literature review helped in the classification and validation of the findings.

The data analysis followed the grounded theory research approach, hence the assessment process was characterised by constant interplay between deduction and induction. New findings were systematically reviewed during data collection and analysis to determine the generalisability of the results.

Essential results

GDC programmes only partially fulfil their purpose as set out in the GVR. GDC programmes serve as an envelope encompassing the high-level objectives of country strategies and the operational activities of programme modules. However, strategic planning and steering is currently limited at the programme level, with most such activities being conducted at the module level.

Demand for evidence at the GDC-programme level

The demand for programme evaluation in general and the resulting evidence specifically regarding programme progress is high and covers a wide range of topics.

Actors at the BMZ and the implementation organisations expect substantial benefits from conducting such evaluations. It is considered particularly important to **learn more about the definition of objectives at the programme level and the interconnections between modules**. GDC programme evaluations can be performed at various points in time, for instance: ex ante, mid-programme or at the end of a programme (final or ex-post evaluations). They could either be performed as reviews, following a streamlined format, or as fully fledged strategic evaluations. The majority of survey respondents and interviewees prefer **evaluations during programme implementation** (as mid-programme evaluations) to allow the evaluation results to feed into the implementation of the rest of the programme.

Two main options exist for the evaluation of GDC programmes:

- **Higher-level strategic evaluations**: dealing with overarching strategic thematic or regional questions based on multiple programmes to generate evidence for strategic decision-making, institutional learning and accountability
- **Evaluations of individual programmes**: focusing on one programme and aimed at providing specific results at the operational level, thus informing programme management and steering while improving accountability

Results orientation and evaluability of GDC programmes

The intended results at the programme level should reflect the development policy objectives stipulated in the country strategy document (impact level). In practice, **some variation occurs in the way and the degree of detail to which programme objectives are defined**. That is due to imprecise specifications and differing standards expected by those responsible. Quality assurance is described as challenging due to the necessary technical and methodical expertise. Moreover, **programme objective expectations are very heterogeneous**. For instance, disagreements arise over whether objectives should generally be formulated in such a way that they can be achieved through interaction between the modules within a specified period or whether they should be understood and defined as a long-term contribution to development policy changes in the partner country. However, an **aligned understanding of GDC programme objectives** is crucial for **monitoring and steering activities**. The current practice of formulating intended results is thus found to be insufficient for adequately evaluating GDC programme objectives.

The **definition of indicators** is regarded as a very challenging process, partly due to the previously described weaknesses in defining objectives and partly because of a lack of data. Data for measuring indicators is generally to be provided by the development partners in each country. However, such data is often not available when needed for reporting, at least not at the required aggregation or quality levels. Furthermore, the programme objective indicators used do not always meet the SMART criteria. The case studies

also show that the necessary data is not provided consistently, thus confirming the finding that the **evaluability of the GDC programmes** reviewed varies and is only partially ensured.

According to the guidelines, programme theory specifies the intended results of a GDC programme and should be developed early in the design phase. However, there is currently insufficient documentation regarding the **impact assumptions and, especially, the foreseen intertwining of modules (interconnections)**. Consequently, it is not always apparent how exactly the modules should take **effect at the programme level**. Overall, barely any interaction in the sense of close operational intertwining of modules is found in practice.

Moreover, the three case studies reviewed show that the partner strategy documents are used to define the thematic priorities and focus of GDC programmes. The development challenges in the partner country and the related consequences for certain target groups are also described. However, it is not always clear whether the intended changes for the target groups are realistic and achievable within the foreseen programme duration.

Although external factor **assumptions and risks** are generally highlighted and discussed, less information is available on the potential unintended (positive or negative) effects of GDC programmes on, for example, the target groups. Consequently, it is hard to find information on anticipated **strategies and measures for detecting and reducing risks** during programme implementation. As a result, it is also hard to assess and evaluate those aspects in retrospect.

Regarding organisational structures, the essential steps and **responsibilities within GDC programme design and implementation** are set out in various guidelines. However, as any unclear aspects can lead to inconsistency and sometimes even a loss of quality, **further clarification and specification is needed** in some places for the implementation of individual procedural steps. That applies to the clarification of roles not only between the BMZ and its implementing organisations but also among the implementing organisations themselves. That is particularly important for improving operational programme management and quality assurance.

Throughout the whole GDC programme planning and implementation process, close **exchange takes place with various implementing partners** at the module level. Long-term GDC objectives and commitments are discussed with the government in a respective partner country as part of regular government negotiations and consultations. However, GDC programmes merely serve as **internal planning and steering tools**; they are not usually subject to discussions with actors in the partner countries.

Much in the same way as in the programme design phase, programme implementation and **corresponding reporting** grapples with the **verification of impact assumptions** (especially in terms of detecting the joint impact achieved by modules).

Programme reporting activities also frequently suffer from a lack of both **primary and secondary data**. A GDC programme collects little data itself, as no resources are specifically assigned to that task. The three case studies draw on secondary data (studies and reports), to varying degrees. Issues appear with the provision of data by partner institutions and consequently cause **data gaps**, which make the monitoring of indicators and the overall reporting of programme progress difficult. Generally, only limited exchange occurs with partner organisations in the given countries regarding the necessary data, its quality and when it is needed. In the case studies reviewed, programme-level monitoring is insufficient for providing up-to-date information on programme objective indicators. In other words, the **lack of data** presents a huge **challenge for measuring** the progress of GDC programmes. Given the previously mentioned weakness regarding the defining of objectives and indicators, the prerequisite for effective programme **monitoring** is rarely met under the current procedural approaches.

Although GDC programme reports in the implementation phase contain information on **factors influencing** the achievement of programme objectives, **internal risks or challenges** were identified in all cases. One practical example of that is provided by the implementation of individual modules relating to either financial or technical cooperation measures. Despite being planned to run simultaneously, those modules are

ruled by different procedures and timelines, thus they generate the risk of implementation gaps or delays. Such internal risk factors are not systematically or sufficiently displayed in current reporting formats.

The deficiencies identified in GDC programme design continue to pose implementation issues, especially in terms of monitoring and reporting. Assessing the progress of GDC programmes over the course of their implementation is thus a huge challenge, which can have detrimental effects on the BMZ's steering capability.

Recommendations

GDC programme objectives and added value

Recommendation 1: The BMZ should more clearly define the objectives and purposes of programmes and their associated steering tasks to enhance transparency and clarity for everyone involved and to increase programme benefits for strategic, results-oriented steering regarding thematic priorities.

Implementation guidance 1.1: The BMZ should work in coordination with the implementing organisations to examine whether the requirements and standards for defining objectives, purposes and coherence should vary depending on the context or fields of action.

Implementation guidance 1.2: More information should be provided on how partner interests are accounted for, especially when the implementing organisations work with various partner institutions potentially pursuing different interests and priorities. Partners' contributions should be described more clearly to better reflect their status as key actors for the successful implementation of GDC programmes.

Programme results orientation and evaluability

Recommendation 2: The BMZ and implementing organisations should improve the results orientation and evaluability of programmes to provide the BMZ with steering-relevant information and allow for sound verification and evaluation of GDC programmes.

Improvements are needed in the following areas, especially in the design phase:

- **Coherent definition of GDC programme objectives at the various impact levels**
- **Appropriate and transparent presentation of the envisaged intertwining and interconnections between modules**
- **Definition of suitable programme objective indicators or, where appropriate, target values**
- **Data availability**

Implementation guidance 2.1: To ensure the systematic embedding of results orientation and evaluability in future planning and steering, BMZ and implementing organisation evaluation units should be involved in developing the corresponding processes and procedures.

Recommendations 3 to 5 expand on the identified weaknesses and possible improvements regarding the individual points listed above.

Heterogeneous formulation of GDC programme objectives

Recommendation 3: The BMZ should further improve the guidelines and quality standards for formulating GDC programme objectives and ensure that those guidelines are implemented.

Implementation guidance 3.1: The objectives should be formulated as a long-term contribution to development policy changes (long-term impact level) or specifically and measurably for a clearly defined timeframe (rather than at the outcome or medium-term impact level). When feasible, specific and measurable objectives should be formulated. If any hindering framework conditions require adjustments, the rationale should be discussed and specified in the planning documents of the given GDC programme.

Implementation guidance 3.2: If the content of modules cannot be sufficiently integrated into a GDC programme and they do not make any measurable contributions to the programme objective or programme objective indicators, those modules should not be presented as part of that GDC programme.

Implementation guidance 3.3: The guidelines and procedural information should provide clearer guidelines prior to programme planning. Questions regarding whether the objectives are measurable and specific are currently only addressed during programme design and at the discretion of those involved.

Recommendation 4: Before GDC programmes are launched, the BMZ should assign contact people with sufficient methodical skills to ensure the quality of the formulated objectives and indicators. Quality assurance for the objectives should also include the revision of the indicators and be performed by either the BMZ or the implementing organisations.

Implementation guidance 4.1: Specific responsibilities, financial resources and time requirements should be defined and made available for quality assurance purposes.

Implementation guidance 4.2: If implementing organisations are to perform the quality assurance of objectives and indicators, they should also be more involved in defining the GDC programme objectives. Objectives are currently often formulated at the country-strategy level, with little involvement of the implementing organisations.

Improved data availability

Recommendation 5: Suitable monitoring should be established at the programme level. The implementing organisations should clearly indicate during GDC programme design to what extent the data required for monitoring is available from the partners, secondary literature and modules, and what additional data needs to be collected at the programme level. The BMZ should provide corresponding resources.

Implementation guidance 5.1: During programme design and finalisation of objectives, the following measures should be implemented:

- The BMZ and implementing organisations should agree on what data is required for monitoring GDC programme progress. On that basis, the implementing organisations should develop a suitable monitoring system
- Partner data should be used wherever possible
- Supplementary targeted data collection measures should be adopted to determine baseline and monitoring data, where necessary. The BMZ should provide the required resources. It is important that the implementing organisations provide for the necessary resources during GDC programme design

Implementation guidance 5.2: The implementing organisations should assign qualified staff with sufficient resources for the monitoring of each GDC programme.

Demand for evidence and benefits of GDC programme evaluations

Recommendation 6: To ensure maximum added value and the most benefit from GDC programme evaluations, the BMZ needs to clarify the following issues when planning and designing the evaluation format:

- **Who will use the GDC programme evaluation results?**
- **What should the results be used for (accountability, steering/redirecting ongoing programmes, learning across different programmes)?**

- **When and how frequently should GDC programme evaluations be conducted to fulfil the intended purpose (for each programme; a random or a criteria-based sample)?**
- **How can partners be involved in the evaluations and in what form can results be shared with them?**
- **Which of the specified issues and questions should GDC programme evaluations address and which should they omit? Can other formats be used to answer those omitted questions?**

Implementation guidance 6.1: Depending on the demand for evidence, GDC programme evaluations need to be conducted at various times and for different target groups. If the results are to be used by programme managers from the BMZ and the implementing organisations for planning and steering, the evaluations need to be conducted and completed during programme delivery, wherever possible. A suitable time would be during the regular reflection phase after three years of programme implementation or a little later if the results are to be used to inform the planning process for a follow-on programme. In that case, GDC programme evaluations would be conducted as mid-programme evaluations or reviews, depending on the programme duration and scope. Strategic programme evaluations investigating multiple programmes in relation to overarching issues could be performed as an alternative or in addition. Such evaluations would place the emphasis on learning across different programmes.

Implementation guidance 6.2: With a view to the sequencing of programme evaluations, it is necessary not only to clarify whether evaluations should generally be used as a source of information for steering and reflecting on programmes but also whether a mid-programme evaluation should be conducted as part of every GDC programme. A targeted sample could be evaluated for accountability purposes and to answer higher-level and strategic questions. Using geographical and thematic clusters would be another option.

Implementation guidance 6.3: The GDC programme evaluation format should be subject to regular review and enhancement. Particular reflection is needed on whether further questions can be answered in future or whether additional data collection within programme evaluations can be reduced if the necessary data is made sufficiently available through improved monitoring and/or reporting.