



Evaluation

Meta-Evaluation of ADA Project and Programme Evaluations

Executive Summary

Imprint

Austrian Development Agency (ADA),
the operational unit of the Austrian Development Cooperation
Zelinkagasse 2, 1010 Vienna, Austria
Phone: +43 (0)1 90399-0
Fax: +43 (0)1 90399-1290
office@ada.gv.at
www.entwicklung.at



Dr. Stefan Silvestrini, Sandra Bähge, M.A.

Saarbrücken, June 2019

This is an independent report. Views and conclusions expressed do not necessarily represent those of the contractors.

Introduction

In its effort to combat poverty, ensure peace and preserve the environment, the Austrian Development Agency (ADA) yearly funds several hundred individual projects and programmes, implemented by partner organizations in its key regions and priority countries. According to ADA's Guidelines for Project and Programme Evaluations "every project or programme must be evaluated once internally within the project or programme cycle" (ADA, 2009). In order to systematically assess the quality and usefulness of these evaluations, ADA mandated the CEval GmbH to conduct a **meta-evaluation** of ADA project and programme evaluations completed between January 2016 and June 2018, the timeframe of the most recent Three-Year Austrian Development Policy (MFA 2017).

Purpose of the evaluation

The purpose of the meta-evaluation was

- to better understand the causal pathways of project and programme evaluations in ADA and what is actually expected and assumed,
- to assess the quality and usefulness of project and programme evaluations at ADA,
- to provide an overall assessment of their strengths and weaknesses, and
- to identify enabling and hindering factors for evaluation quality and use.

By providing recommendations for improving the project and programme evaluation tool, the meta-evaluation should contribute to making project and programme evaluations more effective in increasing ADA's performance.

Methodology

The meta-evaluation was based on a participatory approach and combined inductive and deductive elements as well as a mix of qualitative and quantitative methods of data collection and analysis. 50 evaluations were randomly selected out of the full portfolio of 93 evaluations. Based on a comprehensive quality assessment protocol (with almost 150 unique (sub)criteria), the evaluations' ToR and reports were analysed along the following four analytical dimensions: (1) accuracy, (2) evaluation feasibility and process quality, (3) fairness and (4) usefulness and use.

Apart from the document analysis, primary data was collected from ADA staff and ADA partner staff via two semi-standardized surveys as well as two focus group discussions. This data was used either to rate (sub-) criteria for which the documents did not provide relevant information or for explorative and triangulation purposes.

In terms of data analysis, qualitative content analysis, descriptive statistics, group comparisons and correlation analyses of supposed independent and dependent variables were used. Finally, a structural equation model based on partial least squares analyses was developed to estimate the factors influencing the quality and utility of the evaluations at large.

Key findings

Evaluation quality

(analytical dimensions: (1) accuracy, (2) evaluation feasibility and process quality, (3) fairness)

While it is a positive result that no evaluation was assessed as inadequate (looking at methodology and process), on the other side there were very few evaluations that stood out as being of good or very good quality.

Especially the quality of the evaluations relating to **accuracy and adequacy** shows rooms for improvement:

A critical discussion of the intervention logic was frequently missing, which is noteworthy given that practically all evaluations in the sample looked into the effectiveness of the interventions under scrutiny and given that an assessment of effectiveness usually requires some reference to the intervention logic.

One of the weakest performance relates to the presentation and interpretation of evaluation results. One in four evaluations were considered fully inadequate in relation to how causal inferences were made, and in addition two thirds of the sample proved worthy of improvement. Observed changes were simply assumed to be caused by the intervention without either laying open the effect mechanisms or properly discussing alternative explanations.

It is worrying that the methodology sections in the reports most often only focused on the data collection stage, while hardly addressing the analytical phase.

By contrast, the sample performed considerably better with regard to evaluation **feasibility and process** as well as **fairness**.

The involvement of the evaluation stakeholders and their interest in the evaluation results as well as contributed to a well-assessed evaluation process. Evaluation results and recommendations were mostly needed for project steering, development and/or adaptation, while measuring the project's effectiveness and/or impact apparently played a less important role. It appears that the is rather a means to an end than an actual reason for an evaluation.

The assessment through stakeholders of the competence and credibility of the evaluators was among the issues which delivered the most promising results. However, the findings relating to quality suggest that the evaluators' actual methodological performance is not quite in accordance with those perceptions. Hence, the aforementioned result rather has to be interpreted in the sense that from a stakeholders' perspective, the evaluators' methodological competences are regarded as sufficient.

Apart from the reports as such, the ToR also showed room for improvement particularly with regard to the adequacy of the methodological requirements, the descriptions of the evaluations' objectives and scope, the number of evaluation questions as well as their overall feasibility in terms of time and resources. Most ToR contained far too many evaluation questions to be realistically answered. What is striking against the backdrop of the exigencies formulated in the majority of the ToR, is that almost one third of the evaluations only had a budget of 5.000 Euro or less.

Evaluations were implemented impartially and the assessment of the evaluation objects which considered both their strengths and weaknesses, was mostly fair. There are evaluation reports which can be considered good practice examples on how to present a balanced and fair assessment

Evaluation utility

(Analytical dimension (4) usefulness and use)

Concerning the utility of the evaluations in the sample, the meta-analysis concluded that overall ADA project and programme evaluations displayed a satisfactory performance. In that respect, it has to be acknowledged that with regards to various aspects the assessments made by the project staff (mostly of partner organisations) to a certain degree differed from those of ADA staff at headquarters. The nature and underlying reasons for those differing perspectives are certainly worth revisiting in the follow-up to this meta-evaluation

Most of the evaluation reports were logically structured and complete, which is a relevant factor for their actual usefulness.

The evaluations were furthermore implemented in a timely manner, which is a crucial prerequisite for the actual use that can be made of its results.

Based on the assessments of their primary users, the meta-evaluation could also show that most of the sampled evaluations were actually useful in terms of further developing interventions and improving intervention approaches, and that they were actually used for those purposes.

However, the validity and usefulness of recommendations was found to be an aspect worth of improvement, especially in regard to directing recommendations to particular actors and prioritizing them.

The completeness of the reports' annexes was also an issue mostly due to the failure to incorporate documents which are relevant in terms of transparency, especially the ToR, data collection instruments and consulted documents.

Only half of reports highlighted important messages and the occasional absence of conclusions sections was identified as an aspect to improve on.

While evaluations are being used, and there is a strong focus of ADA staff on using evaluations as learning tools and the desire to use evaluations more strategically instead of doing them routinely (according to ADA evaluation guidelines 2009 every project or programme must be evaluated (internally) at least once during the project or programme cycle).

Influencing factors

A number of factors have been identified as **influencing factors** behind the results presented above.

In particular, the quality of the evaluations was found to be affected by the quality of the ToR, a comprehensive description of the evaluation methodology, a thorough understanding of the evaluation object, the ability and willingness of the stakeholders to take part in its planning and implementation, the design and maturity of the evaluation object, the introduction of 'external' expertise, the consideration of unintended effects and external factors, and the flexibility of the commissioners to adapt the evaluation design if necessary.

Actually finding qualified evaluators who possess all required technical, local and evaluation expertise, and who are furthermore available for the required timeframe was highlighted as one of the major challenges involved in the evaluation process.

With regard to the usefulness of the evaluations, the meta-evaluation team found that almost all factors that have repercussions on an evaluation's quality also affect its usefulness.

The transparency and credibility of the evaluation results, a clear purpose of the evaluation, the suitability of the evaluation results for external communication, policy dialogue etc., a thorough discussion of lessons learnt, successes and failures, an adequate discussion and dissemination of the

evaluation results, the consideration of external factors, the demonstration of new insights, experiences from other comparable interventions and regions, and the integration of consulting measures in the evaluation for future M&E activities were identified as relevant in that respect.

Recommendations

The meta-evaluation concluded that there are several areas in which there is still room to improve ADA project and programme evaluations. Based on this, the evaluation team developed the following **recommendations**:

General recommendations relating to the meta-evaluation use (directed at ADA Evaluation unit):

1. **Share the results of this meta-evaluation** and invite feedback from various stakeholder groups (esp. ADA project and partner staff); in particular, try to find out the reasons for the somewhat differing perspectives especially in relation to the usefulness and use of evaluations; take this discussion as a starting point to learn more about the various stakeholders' expectations regarding evaluations and in order to improve on the more specific recommendations outlined below.
2. This meta-evaluation was the first meta-evaluation conducted on behalf of ADA. In order to systematically enhance the quality and usefulness of ADA project and programme evaluations, this tool should become a **strategic element of the ADA evaluation portfolio**.

Specific recommendations relating to the project and programme evaluation tool (directed at ADA evaluation unit and other ADA and partner staff):

3. Revisit **ADA evaluation guidelines**, most importantly, the requirement to do an evaluation in case of each and every ADA funded evaluation should be seriously discussed with regard to both the restricted resources available for ADA evaluations and the questionable utility of obligatory evaluations.
4. When planning an evaluation, determine the available evaluation budget upfront. On that basis, assess whether and under what conditions an evaluation is realistically feasible in terms of financial resources, and what can be expected from it considering the available resources.
5. Ensure that **adequate ToR** are drafted, which include the clear determination of evaluation purpose, questions as well as a realistic scope that corresponds to all the former as well as to the available resources, i.e. time and financial budget.
6. Pay thorough attention to a proper inception phase of any planned evaluation in order to ensure that not only data collection but also evaluation design and data analysis rest upon a methodologically sound basis. Furthermore, make sure evaluation stakeholders (especially commissioners) are sufficiently informed about the aforementioned issues in order to enable a smooth implementation process.
7. Establish an organisation wide mechanism or a platform for **institutional learning**, that is, a mechanism to share and communicate evaluation results across and within ADA units.
8. Conduct a needs assessment on the level of ADA and partner staff, and on that basis design appropriate measures **to enhance evaluation capacity**.