

Better Criteria for Better Evaluation? Reforming the International Evaluation Criteria

Noltze, Martin; Harten, Sven

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

Empfohlene Zitierung / Suggested Citation:

Noltze, M., & Harten, S. (2021). *Better Criteria for Better Evaluation? Reforming the International Evaluation Criteria*. (DEval Policy Brief, 1/2021). Bonn: Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval). <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-77815-1>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:

<https://creativecommons.org/licenses/by-nc-nd/4.0>

BETTER CRITERIA FOR BETTER EVALUATION? REFORMING THE INTERNATIONAL EVALUATION CRITERIA

Summary

One of the most important features of an evaluation is measuring the success of an intervention. In evaluations of development cooperation, this assessment has been guided by the internationally established evaluation criteria of the Organisation for Economic Co-operation and Development's Development Assistance Committee (OECD DAC) since the early 1990s. The OECD DAC has now updated the evaluation criteria for the first time with the intention of creating "better criteria for better evaluation" (OECD, 2019). In this process, the organisation enhanced the previous evaluation framework, consisting of the criteria "relevance", "effectiveness", "efficiency", "impact" and "sustainability", with the new criterion "coherence" (see figure).

The starting point for the reform of the evaluation criteria was the principles and goals of the 2030 Agenda for Sustainable Development, which the United Nations adopted in 2015. The main driver of the reform process was the World Bank's independent evaluation unit and, in particular, the German Federal Government, with support from the work of the German Institute for Development Evaluation (DEval). The objective of the reform under the auspices of the OECD DAC evaluation network was to continue developing the evaluation criteria to allow them to support modern and effective development policy even better than before.

This first reform of the DAC criteria is extremely important for the learning and accountability functions of development cooperation evaluations. Both the OECD member states and their partners currently structure their evaluations primarily based on the DAC criteria. In German development cooperation, too, the major implementing organisations determine their success rate at regular intervals using the DAC criteria. The great

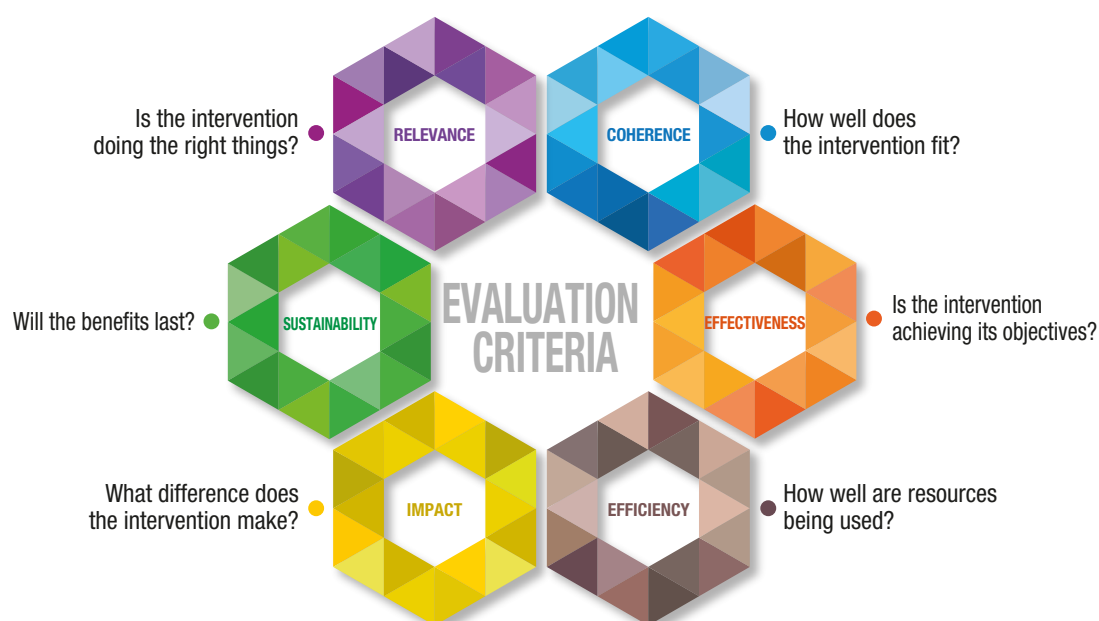
significance of the criteria in development cooperation both accelerated and inhibited the reform process. The global consulting process was sustained by a broad consensus in the development cooperation evaluation community that the evaluation criteria, and therefore development cooperation, should be made "fit for the future". During the consulting process, the stakeholders mentioned the far-reaching acceptance, clarity and universal applicability of the criteria as a foundation worth preserving.

Having weighed up future orientation against established approaches to the success of development interventions, the OECD decided to reform rather than transform the criteria. The strengths of this approach lie in the compromise between proven and reformed practice. Overall, the reform increases the quality and usefulness of the international evaluation criteria, and therefore improves the future viability of development cooperation interventions.

The reformed OECD evaluation criteria

In October 2017, the Organisation for Economic Co-operation and Development's Development Assistance Committee decided to reform its evaluation criteria using inputs from a global consultation process. While assessing the status quo, the OECD DAC's evaluation network determined that the key strengths of the original criteria were their universal acceptance and prevalence, their conceptual consistency and scope, and their simplicity, clarity and neutrality. Aspects that the network cited as central challenges were their inadequate alignment to the principles of the 2030 Agenda, the poor applicability to strategic and institutional evaluations and the highly linear and not particularly systemic overall design of the criteria.

The OECD DAC evaluation criteria



Source: Presentation according to OECD DAC (2019).

To ensure simplicity and clarity, the OECD DAC has modified the definitions of the individual evaluation criteria and clarified the dividing lines between them. For example, the evaluation canon now includes aspects of the responsiveness and adaptability of development interventions in order to better account for the potential transformative nature of development cooperation. The Federal Government of Germany also supported the reform efforts, having previously addressed the design and practical application of the evaluation criteria as part of a meta-evaluation by DEval on the sustainability of German development cooperation (Noltze et al., 2018).

In an innovative move designed to strengthen the systemic character of increasingly complex development cooperation interventions, the OECD has supplemented the previous five evaluation criteria with the new criterion of coherence. This criterion is used to determine how well an intervention fits within the context of other interventions. It also links the evaluation of an intervention's design in terms of its relevance and its implementation from an effectiveness perspective.

A high-quality and useful evaluation should focus on the purpose of the evaluation and the needs of those it addresses. For this reason, the reformed evaluation criteria now include some basic

principles for use to ensure their practical applicability and usefulness. For example, one update that OECD member states and other stakeholders requested was that the criteria should be contextualised; in other words, those applying the criteria should understand them in the context of the specific evaluation and the intervention to be evaluated and should apply them in dialogue with those who are involved and affected. It is very likely that the definition of these application principles will improve the usefulness and acceptance of the evaluation. At the same time, there is a risk that stakeholders will use the criteria arbitrarily by changing their substance in individual evaluations – for example with regard to the content that they focus on and the weighting of criteria or individual questions. This, in turn, would reduce the desired standardisation and comparability of development cooperation evaluations, which could endanger the harmonisation achieved in development cooperation evaluation to date (see also Faust and Verspohl, 2018).

The report “Better Criteria for Better Evaluation” (OECD 2019) presents definitions and principles for using the revised OECD DAC evaluation network's evaluation criteria. A further OECD guide (“Thoughtful evaluation: Using criteria for better evaluation”) will support practical application of the criteria from the beginning of 2021.

The evaluation criteria in German Development Cooperation

The result of the reform process at OECD level will have far-reaching effects for the evaluation systems of the individual member states. On the initiative of the German Federal Ministry for Economic Cooperation and Development (BMZ), German development cooperation got involved with the reform process at an early stage and looked into the potential effects of the reformed criteria on its own development cooperation evaluation system. Within the framework of a wide-ranging consultation process and with the support of DEval, the BMZ amended its binding guideline (mandatory since 2006). The “Evaluation criteria for German bilateral development cooperation” (BMZ, publication in preparation) have now been binding for government implementing organisations since September 2020 and serve as a guide for non-governmental organisations and DEval.

It is worth highlighting here that BMZ had already laid out “coherence” along with “complementarity” and “coordination” as an additional criterion for evaluation at country-programme level and for joint evaluations in 2006. Due to the low number of programme-level evaluations, however, there is still little practical experience of evaluating coherence.

With the reformed evaluation criteria, in future the international community will measure the success of development cooperation based on its contribution to implementing the principles of the 2030 Agenda’s Sustainable Development Goals (SDGs). The principles include universality, shared responsibility and accountability, synergy between economic, environmental and social development, and inclusiveness. In view of the strong significance of the 2030 Agenda for development policy, German development cooperation has supplemented the new DAC evaluation criteria with further explicit evaluation questions alongside the Agenda principles. For example, in the future they will summarise and evaluate key contributions to implementing the 2030 Agenda together.

The evaluation criteria in DEval evaluations

DEval attaches great importance to the DAC criteria in its evaluation work. Every DEval evaluation examines the applicability of the six evaluation criteria against the backdrop of the evaluation object and research interest in question and is transparent about the reasons for excluding any criteria. Application of the coherence criterion is a particular priority for DEval’s strategic evaluations, for example to ensure the coherence of the various policy fields of the German Federal Government. In addition to the environmental, economic and social dimensions of sustainability, DEval is concerned with the political aspects of sustainable development. Another key component of DEval’s work in the evaluation capacity development (ECD) field is the distribution and regionally adapted use of the criteria.

Summary

The OECD DAC evaluation criteria are among the most important public goods in international development cooperation. Both the OECD member states and their international partners structure their evaluation work based on the reference framework of the now six evaluation criteria. Due to their learning and incentive function, the evaluation criteria also have an impact on policy fields outside development cooperation.

By reforming the evaluation criteria, the OECD is contributing to adapting development cooperation evaluations to future challenges and therefore to designing development cooperation interventions to be fit for the future. In particular the integration of additional principles and perspectives of international cooperation that arise from the 2030 Agenda underline the contemporary nature of the otherwise timeless evaluation criteria. However, the success of the criteria will also depend on the extent to which the individual OECD member states and multilateral organisations succeed in mainstreaming the principles and updated design into their own evaluation systems. In this context, the supplementary OECD guide to applying the evaluation criteria is a welcome addition. Furthermore, the OECD has the opportunity to facilitate the exchange of experience, particularly during the implementation phase, and to provide the member states with learning examples via their existing knowledge platforms. At the same time, the team needs to increase efforts to integrate the partner countries into sharing

experiences and knowledge and to support them in applying the criteria by strengthening their evaluation capacities.

In the interests of increasing acceptance and usefulness, it makes sense to adopt the purpose- and context-oriented application principles. A potential risk is the possible low level of standardisation, which may make it difficult to compare the success evaluations between different interventions and over

time. Against this background, BMZ should support the use of the criteria in German development cooperation by offering further details using and weighting individual assessment dimensions and evaluation questions. An already planned pilot phase for introducing the reformed DAC evaluation criteria into the German development cooperation evaluation system will provide a suitable framework for this.

References

BMZ (publication in preparation), *“Evaluation criteria for German bilateral development cooperation”*, Bonn/Berlin.

Faust, J. and I. Verspohl (2018), *“The OECD-DAC Evaluation Criteria: To Reform or To Transform?”*, DEval Policy Brief 9/2018, Bonn.

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

OECD (2019), *“Better Criteria for Better Evaluation. Revised Evaluation Criteria Definitions and Principles for Use”*, Organisation for Economic Co-operation and Development, Paris.



Dr Martin Noltze

Team Leader



Dr Sven Harten

Head of Competence Centre for Evaluation Methodology / Deputy Director

The German Institute for Development Evaluation (DEval) is mandated by the German Federal Ministry for Economic Cooperation and Development (BMZ) to independently analyse and assess German development interventions. Evaluation reports contribute to the transparency of development results and provide policy-makers with evidence and lessons learned, based on which they can shape and improve their development policies.