

15 March 2025

Resolution: Five demands for impact-oriented international cooperation

- Align international cooperation more closely with partners and target groups, strengthen their competencies
- Rely on the tried and tested: best buys save money on evaluation and enable highly effective IC
- Make evaluations more robust and targeted
- Incentives and competencies for the use of evidence
- Transparency for global knowledge transfers

International cooperation¹ (IC) is facing the challenge of reduced budgets. Focusing on impact can help to **achieve more with fewer resources**. In any policy area, numerous hurdles can lessen the impact or sustainability. Only if you know what works and how can you **make good policy**. IC is one of the areas of public policy that is evaluated the most. The available expertise within the Federal Ministry for Economic Cooperation and Development (BMZ), the German Institute for Development Evaluation (DEval), and implementing organisations (such as GIZ and KfW) must now be better utilised. This way, we can work even more effectively and ensure that we continue to learn long-term. Germany currently spends far below average on ensuring the cost-effective use of funds.² German international cooperation must, therefore, continue to expand its **transparency, impact evaluations and proximity to science**.³ With this paper, the BAG Globale Entwicklung (Federal Working Group on Global Development), commits to strengthening impact orientation in German international cooperation.

This paper's concrete proposals for improvement build on Alliance 90/The Greens' long-standing commitment to impact orientation, most recently reiterated in the 2025 Bundestag manifesto as a "requirement of all our international action." The party's policy programme adopted in 2020 also states⁴ that evidence-based approaches and constant dialogue with the scientific community are essential for effective climate and development policy. These demands to strengthen impact orientation are also core demands of the Global Partnership

for Effective Development Cooperation⁵ and the Paris, Accra and Busan agreements for more effective IC, of which Germany is a member.

With this paper, we are updating our decision from 2012⁶ and our discussion status from 2021⁷ and developing concrete proposals to make all expenditure on Official Development Assistance (ODA) within the framework of IC **as effective, efficient, transparent, sustainable and measurable as possible**. Some of our proposals can only be implemented in the long term.

We are convinced that expenditure on impact orientation and (robust) impact measurement are no mere incidental costs. Instead, they are investments that increase the effectiveness of IC in the long term and, therefore, strengthen IC. In the fight against nationalist and populist forces, IC can best be defended if it is based on the principles of science.

1. Partner and target group orientation

Effective international cooperation includes the self-determination of priorities by the partner countries among the sustainable development goals (SDGs) and partner country ownership. The international community agreed on this in the Paris Declaration and enshrined it in Sustainable Development Goal 17 of the 2030 Agenda. This also means that IC is accountable to the partner countries where it is implemented and should support partner countries in using impact data for their decision-making.

We also want to anchor economic and social justice and ecological sustainability more firmly in IC design, implementation and impact evaluation. In concrete terms, this means:

- 1.1 Evaluations should be emancipatory and participatory. As part of a decolonial approach, they should involve local knowledge of the partner organisations and target groups. Evaluation standards from the Global South should be regarded as equal to standards developed in the Global North. This is in the spirit of localisation and **"Nothing about us without us"**.
- 1.2 As part of a decolonial and feminist approach, a more equal approach towards partners is to become the basis for the entire project management process (planning, monitoring and evaluation).⁸ Such an approach is to be further developed in the future.
- 1.3 As part of technical cooperation, it should be systematically examined if measures can be implemented better and more efficiently by local executing agencies (para-state institutions, companies, NGOs) instead of GIZ. If so, the contracts would still be awarded to the local executing agencies by GIZ on behalf of the BMZ.
- 1.4 For tendered evaluations, teams with **evaluators from partner countries** should be favoured. This would promote international collaborations in the medium term and create independent capacities in the long term. Existing evaluations by partners should be utilised as much as possible to relieve the burden on German evaluation units.⁹ In addition, it should be critically scrutinised whether evaluations by local organisations are more cost-efficient and possibly also more effective (local ownership). In the context of decolonial IC, it is also essential to create conditions that

ensure equality in the contractual relationship when local and international evaluator teams work together.

- 1.5 Where possible, data on the effectiveness of IC projects should be collected in cooperation with local research communities. The utilisation and strengthening of information systems in the partner countries and their use for decision-making should be a cross-sectoral objective of German IC.
- 1.6 In bilateral and multilateral international cooperation, Germany should advocate for impact orientation in partner countries. This should be based on the principles and international agreements described here. International systems for impact measurement should be promoted to counteract data fragmentation.
- 1.7 German IC should take its lead from organisations that have proven to be exceptionally cost-effective¹⁰ and work closely with them to determine which measures can permanently improve the evidence base, efficiency and effectiveness of its own projects.
- 1.8 It should be ensured that the impact of IC focuses on economic and social justice (including equality) and ecological sustainability.
- 1.9 The BMZ should join the 3ie Global Evidence Commitment (as KfW has already done) and thus signal its commitment to evidence-based development cooperation internationally. This would also strengthen the exchange with other evidence-based donors.

2. Making better use of evidence

Evaluation is not an end in itself but must always serve the decision-making process. Unfortunately, not all available findings are always utilised at the funding allocation and project planning level. It is true that new IC projects are elaborately prepared in highly formalised processes. However, there is a lack of prior consideration of possible approaches based on current scientific findings and past internal evaluations, both in the selection process at the BMZ and in the implementing organisations.

It can thus occur that a project is implemented even if existing studies already indicate that the selected approaches have little or no impact. This is because concrete technical planning in bilateral IC usually only takes place once funding has already been assigned to a project in a specific sector in a specific partner country, which is then rarely revised. During project implementation, there is often great pressure to spend funds within a particular calendar year. This creates incentives that can counteract effectiveness and efficiency. This means:

- 2.1 As called for by the German Supreme Audit Institution (Bundesrechnungshof) in 2023,¹¹ **evaluation findings** must be systematically used to improve existing programmes and design new ones. This should apply not only to the BMZ, but to all ministries: All ministries should adhere to the same requirements for impact measurement and evidence.

- 2.2 To ensure that measures always correspond to the current state of knowledge regarding their expected effectiveness, ministries should utilise existing **public knowledge** such as best buy lists of the most efficient interventions or systematic reviews. This saves costs as it eliminates the need for in-house evaluations. Examining best buy options should be mandatory in project and programme planning processes.
- 2.3 In particular, the review of scientific findings and previous evaluations should be enshrined more systematically than before when **awarding contracts**. The effectiveness of comparable project activities should be compared at impact level,¹² for example through systematic reviews.
- 2.4 In project applications, proof of supporting robust evidence on the effectiveness of the planned activities should be mandatory. Particularly robust and decision-relevant findings - e.g. from systematic reviews and meta-evaluations - should be strongly favoured. This systematic evaluation should use globally available evidence repositories¹³ and AI tools to minimise additional workload.¹⁴

If a project application for a project of over 1 million euros cannot refer to existing evidence due to the innovative nature of the project, this should be justified and the justification published transparently (insofar as diplomatically possible). Corresponding evidence should then be generated in the course of the project. In such cases, a robust evaluation design should be used to generate evidence for future applications.

Where possible, project applications should be justified against an adequate benchmark. One suitable benchmark is direct monetary transfers to the project's target group. They have been widely researched and are almost universally applicable. A "best in class" approach, which compares best buys in specific sectors with the proposed project, could also achieve immense progress in effectiveness. In using evidence for project planning, experience should be drawn from international partners such as USAID and Norad.¹⁵

- 2.5 Projects should **be able to adapt** their planning much more to changing conditions. Ever tighter bureaucratic requirements for planning and cost management prevent projects from being able to carry out impact-orientated interventions. For this reason, project application procedures and strategic and cost-related operational planning need to be made more flexible to meet the requirements of agile management in the dynamic contexts of partner countries and institutions.
- 2.6 The information obtained during a project must be used for steering and **for the development of follow-up projects**. It is only worth collecting evidence if it can be used for these two purposes. When in doubt, fewer evaluations should be carried out, but in the right places, instead of spending a lot of money on evaluations that later offer no added value. GlZ's approach of evaluating projects mostly unsystematically and not based on the expected value of the knowledge must, therefore, be questioned. The scientific rigour of evaluations should be assessed using the Maryland Scale, with a preference for approaches with a valid control group.¹⁶

- 2.7 It must be possible to **improve projects**, for example, by reallocating funds. An evaluation that certifies that a project has had little impact but identifies plausible opportunities for improvement must not jeopardise follow-up funding. Otherwise, there are strong incentives to embellish evaluation reports or not publish them at all.¹⁷
- Impact monitoring should be used consistently to learn from experience and dynamically adapt projects as they progress. This requires more manager attention, competences and resources on the part of the IC institutions. It also requires that impact monitoring be steered together with the partners towards jointly planned goals.
- 2.8 In the interests of legitimacy, external communication, and ethical imperatives and because of the limited resources of the federal budget, the focus of IC should be on those sectors where German IC can achieve the most with limited resources.
- 2.9 The prioritisation of funds must be based even more consistently on the cost-effectiveness of various sectors and measures. Particular attention should be paid to **preventive approaches** that can avoid future crises and their high humanitarian and financial costs, as well as measures for particularly **neglected and vulnerable groups** that are not reached by other actors. Frameworks such as the 'Scale-Tractability-Neglectedness' model or best buy lists can serve as orientation.
- 2.10 Evidence-based prioritisation does not mean abandoning entire sectors but rather identifying and strengthening the most demonstrably effective approaches within each sector. The BMZ should, therefore, introduce a systematic procedure for assessing cross-sectoral cost-effectiveness and reserve a **fixed budget share** for projects with proven cost-effectiveness.

3. Targeting where to collect evidence

Evaluations should be designed in such a way that they credibly measure the impact of IC on people and the environment and show potential for improvement rather than simply documenting project implementation and the use of funds. Additionally, evidence should be measured more precisely, as called for by the Supreme Audit Institution in 2023.¹⁸ Ex-post evaluations and before-and-after comparisons measure causal effects (as opposed to developments that would have occurred anyway) only to a limited extent.

By contrast, robust evaluations and impact analyses over longer time horizons help to make development cooperation effective and targeted. We are convinced that expenditure on (robust) impact measurement is no mere side cost. Instead, it is a direct investment into the long-term effectiveness of international cooperation, which contributes to the legitimacy of German development cooperation. This requires that impact assessments be deployed in a targeted manner and used explicitly to improve IC.

Data and analyses collected this way can be a global public good from which other stakeholders can benefit. Where robust evaluation is difficult or disproportionately expensive,

Germany should continue to utilise the full range of evaluation tools. The method must always fit the question. In concrete terms, this means:

- 3.1 Where appropriate, robust evaluation methods should be used. This includes randomised controlled trials and other methods (qualitative and quantitative) that are based on the most robust possible methodology, e.g. according to the Maryland Scale¹⁹. The choice of methods must be justified according to transparent criteria.
- 3.2 Where robust impact evaluations are not possible, adequate other methods should be used. In addition, where there is a risk of unintended side effects, mixed methods and qualitative anthropological and sociological studies should be considered. The collection of suitable baseline data on impact indicators (at impact level, not output level) should also be expanded.
- 3.3 **Evidence competence** should be mandatory for senior hierarchical levels in German IC, as it was in the past at the British DFID.
- 3.4 The increased implementation of robust impact evaluations requires in-depth knowledge in this area. This is to be achieved through **training and further education** for employees of implementing organisations, ministries and other relevant IC institutions.
- 3.5 The **RIE Funding Programme**²⁰ of DEval should be provided with sufficient funds in a significantly enlarged further round in order to carry out a large number of such impact evaluations and to promote the development of adapted methodological approaches. German implementing organisations and non-governmental and church organisations should be closely involved in this process.

The aim is to close knowledge gaps and strategically strengthen local capacities and competencies.

- 3.6 In cooperation with the French development agency AFD, the BMZ should expand the **Fund for Innovation in Development** (FID) into a pillar of European development policy and provide it with sufficient funding of at least €20 million per year.²¹ It is a financing instrument for developing and scaling innovative and exceptionally effective projects.
- 3.7 German international cooperation should focus more on A/B testing - a standard process in the digital economy that is already being used successfully by other ministries, such as the Federal Foreign Office.²² Together with the implementing organisations, the BMZ should set up an A/B testing pilot programme that initially systematically tests different implementation variants in three priority areas. Successfully tested approaches can then be quickly rolled out across the board.
- 3.8 To guarantee **incentive compatibility**, ministries must ensure that comprehensive impact evaluations and their results are never to the detriment of project management and employees.
- 3.9 Ministries should actively demand impact orientation as part of projects from the outset. The promotion of robust evidence should be ensured by setting up a fund for

robust impact evaluations, from which project teams can apply for funding for evaluations independently of the project budget.

- 3.10 Scientists and experts, **especially from the Global South**, should be better involved in evaluations to ensure their quality, exploit synergies between IC and research, and strengthen capacities in partner countries and organisations.

4. Enshrining scientific rigour institutionally

Scientific findings from all disciplines are underrepresented in the planning of German IC. There is a lack of structures and incentives that guarantee a direct exchange between politics and science. Often, there are already published, robust impact evaluations of specific approaches in international IC, but these are not sufficiently incorporated into planning.

In view of the diversity of the SDGs and the large number of government services and interest groups involved in IC, conflicts of objectives in the implementation and evaluation of IC measures are unavoidable. We no longer want to see government action - including in IC - as the sum of individual departmental decisions but as a coherent and binding whole geared towards sustainability, peace and human rights. In concrete terms, this means:

- 4.1 Following the Norwegian example and the motto "Fakta har makta" (facts have power),²³ German development cooperation should commit itself word and deed to the principles of evidence-based impact orientation in the sense of the greatest possible cost-effectiveness.
- 4.2 The **evaluation unit** at the BMZ should be systematically strengthened.
- 4.3 For the further development of German evaluation and learning practice, **DEval** must be **independent** of the BMZ, as only an independent institute can critically and transparently evaluate complex development cooperation. DEval should not be managed by and accountable to the BMZ. Alternative options should be explored, including making DEval an independent federal authority, a public-law institution (AöR) or a registered association (e.V.).
- 4.4 DEval should be able to assess all of the German government's international cooperation measures, including those outside the BMZ's direct remit. The Federal Government should make greater use of cross-departmental evaluations of international cooperation instruments and programmes and provide the necessary funding.
- 4.5 Cooperation with leading scientific research institutions such as J-PAL, 3ie, ISDC, C4ED, IPA and comparable organisations should be intensified and institutionalised to guarantee the analysis of global evidence and its utilisation for German IC. This would ensure the best possible utilisation of globally available evidence and the strategic identification of knowledge gaps and opportunities to increase effectiveness.

- 4.6 The BMZ should set up an AI-supported **evidence chatbot** or hotline that gives project managers round-the-clock access to systematic reviews and best practices.²⁴ Such an infrastructure would significantly facilitate and accelerate the use of robust evidence. This would build on existing approaches in the BMZ for an AI search engine using retrieval-augmented generation.
- 4.7 Mechanisms and incentives should be created that strengthen effectiveness, impact orientation and impact measurement. At the same time, there should be less pressure to spend or commit funds for planned measures within specific time frames simply because they would otherwise lapse or the amount of future funds for the country/sector portfolio in question would decrease.

5. Making IC and evaluations transparent

We want to increase the transparency of all IC measures. Data and experience from IC measures (such as the BMZ Transparenzportal) should be valorised as a public common good. Enshrining transparency can also ensure that standards are maintained in the future, even if a federal government is in office that is rather hostile to IC. In concrete terms, this means:

- 5.1 Module-related, country-related, portfolio-related, and programme-related documents of governmental IC, including final reports and country strategies, should be publicly available as much as possible. This should enable partner countries, the public and the scientific community to independently assess all IC measures. Final reports should be published in a language widely understood in the partner countries and in English. Confidential parts must be removed to protect partners and target groups.
- 5.2 The **AI tools** that contribute internally to better evidence utilisation (see point 4.6) should be publicly accessible as far as possible in the BMZ Transparenzportal to significantly simplify the systematic use of findings.
- 5.3 To promote the scientific investigation of effective governmental IC, evaluations should be published taking into account scientific ethical standards. In addition, cost-efficiency estimates should be published for all completed projects.
- 5.4 German evaluations should be published in the 3ie Development Evidence Portal on a mandatory basis.
- 5.5 At the same time, the standard of the International Aid Transparency Initiative (IATI) should continue to be implemented.²⁵

¹ By international cooperation, we mean measures by all the government services that have an impact on the implementation of the 2030 Agenda for Sustainable Development in countries of the Global South. This includes, in particular, development, foreign, defence, environmental, agricultural, economic and trade policy. The paper is primarily formulated from a federal policy perspective. The paper is also intended as a stimulus for non-state actors.

² While Germany spends just 0.5 per cent of the BMZ budget on research, evaluation and qualification in IC, the expenditure of comparable actors such as USAID (internal benchmark 1-3 per cent) and Norad (internal benchmark 8-10 per cent) is significantly higher, allowing for medium-term cost savings and far offsetting gains in effectiveness. The return on investment of evidence-based policymaking is between 21 and 37 per cent (Abadie et al. 2023, <https://arxiv.org/pdf/2306.13681>)

³ Rigorose Wirkungsevaluierung in der deutschen Entwicklungszusammenarbeit, DEval Policy Brief, 2019. www.deval.org/files/content/Dateien/Evaluierung/PolicyBriefs/2019/DEval_PB_Wirkungsevaluierung.pdf

⁴ Policy programme, Chapter 8, <https://www.gruene.de/artikel/das-neue-grundsatzprogramm>

⁵ <https://www.effectivecooperation.org/>

⁶ https://gruene-bag-globale-entwicklung.de/userspace/BV/bag_nord-sued/Dokumente/120304_BAG_Nord-Sued_Beschluss_Ergebnisorientierung.pdf

⁷ https://gruene-bag-globale-entwicklung.de/userspace/BV/bag_nord-sued/Dokumente/Diskussionspapier_Wirkungsorientierung_-_Webseite.pdf

⁸ See BAG Globale Entwicklung position paper for reform of development cooperation of 22 September 2024, demand 1b, 'Strengthening monitoring and evaluation instruments on the part of the global South' https://gruene-bag-globale-entwicklung.de/userspace/BV/bag_nord-sued/uploads/Beschluss_Reformansaetze_fuer_eine_zukunftsorientierte_internationale_Zusammenarbeit.pdf

⁹ cf. ibid. requirement 2e: 'In [Financial and Technical Cooperation], even greater emphasis must be placed on the use of local and regional expertise, local executing agencies and specialists (short-term experts). In this way, local ownership can be strengthened and FC and TC can be organised more efficiently.'

¹⁰ These include multilateral initiatives such as Gavi and the Global Fund, as well as bilateral agencies such as the Swedish SIDA, see QuODA 2021 <https://www.cgdev.org/quoda-2021>

¹¹ <https://www.bundesrechnungshof.de/SharedDocs/Downloads/DE/Berichte/2023/88-evaluierung-entwicklungszusammenarbeit-volltext.pdf>

¹² For a definition of the various levels of impact, see <http://www.oecd.org/dac/results-development/what-are-results.htm>

¹³ The 3ie Development Evidence Portal serves as a good resource here:

<https://developmentevidence.3ieimpact.org/>

¹⁴ DIME Artificial Intelligence. The ImpactAI system, developed by the World Bank Group, can serve as a basic system for significantly accelerating and improving the use of global evidence.

<https://www.worldbank.org/en/about/unit/unit-dec/impactevaluation/ai>

¹⁵ USAID (October 2024) Cost-Effectiveness Position Paper,

<https://web.archive.org/web/20241129043136/https://www.usaid.gov/policy/cost-effectiveness>

¹⁶ <https://whatworksgrowth.org/resource-library/the-maryland-scientific-methods-scale-sms/>

¹⁷ The evaluation organisation GiveWell can serve as a model for a healthy culture of learning from errors:

<https://www.givewell.org/about/our-mistakes>

¹⁸ https://www.bundesrechnungshof.de/SharedDocs/Downloads/DE/Berichte/2023/88-evaluierung-entwicklungszusammenarbeit-volltext.pdf?__blob=publicationFile&v=2

¹⁹ <https://whatworksgrowth.org/resource-library/the-maryland-scientific-methods-scale-sms/>

²⁰ <https://rie.deval.org/de/rie-foerderprogramm/das-rie-foerderprogramm/foerderprogramm>

²¹ <https://fundinnovation.dev/en>

²² In this approach, two or more variants of a measure are tested in parallel in order to quickly and cost-effectively determine which version achieves the best effect.

²³ https://www.norad.no/contentassets/edb6d72917a84b70bf1bed3283586efc/4.2024_evaluation-of-norads-use-of-knowledge-in-portfolio-management.pdf

²⁴ The Norwegian Development Learning Lab (<https://www.devlearnlab.no/>) and the British Evidence Helpdesk (<https://www.ids.ac.uk/programme-and-centre/knowledge-evidence-and-learning-for-development-k4d/>) can serve as models for this.

²⁵ <https://iatistandard.org/en/>