Evaluation Systems in Development Co-operation 2023

OECD

Evaluation Systems
in Development Co-operation
2023

OECD

BETTER POLICIES FOR BETTER LIVES

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Member countries of the OECD.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Photo credits: Cover © enjoynz, Getty Images

Foreword

Evaluation plays a vital role in improving the effectiveness and impact of development co-operation and humanitarian assistance. By providing high-quality and timely evidence on intervention processes and results, evaluations shed light on what has worked, what has not, why and for whom. Evaluation plays a critical role in organisational learning, shaping future decision making and resource allocations to improve impacts, and strengthens the transparency and accountability of development actors. Evaluative evidence also plays a role in responding to domestic demands to demonstrate the results and value for money of international aid.

However, evaluation cannot be effective without strong systems. In this context, the overall aim of the OECD's Development Assistance Committee (DAC) Network on Development Evaluation (EvalNet) series of evaluation systems studies is to prompt critical thinking around humanitarian and development evaluation, and to help strengthen evaluation systems of EvalNet participants and beyond.

This latest report in the series brings insights to development co-operation providers as they seek to establish or strengthen credible, fit-for-purpose evaluation systems of their own. It provides a snapshot of the core elements and ways of working of development co-operation and humanitarian assistance evaluation systems, drawing on the experiences of EvalNet members and observer organisations. It identifies major trends, covering the role of evaluation, policy and institutional arrangements, evaluation processes and the use of evaluation findings. This study covers EvalNet member countries and observers, focusing on central evaluation units. It provides profiles for DAC member countries with an established evaluation system. The research included a literature review, a survey and key informant interviews.

Chapter 1 provides an overview of the purpose, scope and methodology of the study. Chapter 2 explores the role of evaluation in development co-operation and humanitarian assistance. Chapter 3 outlines the various policy and institutional arrangements used to govern evaluation systems. Chapter 4 reviews the evaluation process, starting with how decisions are made on which evaluations to undertake. Chapter 5 focuses on how evaluation findings are used to improve development co-operation efforts. Profiles of each participant are provided in Annex C.

Acknowledgements

Table of contents

FIGURES

TABLES

BOXES

Follow OECD Publications on:

https://twitter.com/OECD

https://www.facebook.com/theOECD

https://www.linkedin.com/company/organisation-eco-cooperation-development-organisation-cooperation-developpement-eco/

https://www.youtube.com/user/OECDiLibrary

https://www.oecd.org/newsletters/

Abbreviations and acronyms

AIIB	Asian Infrastructure Investment Bank
ADA	Austrian Development Agency
ADB	Asian Development Bank
AFD	Agence Française de Développement (France)
AfDB	African Development Bank
BMZ	Federal Ministry for Economic Cooperation and Development (Germany)
COVID-19	Coronavirus disease 2019
DAC	Development Assistance Committee
DEval	German Institute for Development Evaluation (Germany)
DFAT	Department of Foreign Affairs and Trade (Australia)
EBA	Expert Group for Aid Studies (Sweden)
EvalNet	Development Assistance Committee Network on Development Evaluation
FCDO	Foreign, Commonwealth & Development Office (UK)
FPS Foreign Affairs	Federal Public Service Foreign Affairs (Belgium)
GAC	Global Affairs Canada
GEI	Global Evaluation Initiative
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit (Germany)
GNI	Gross national income
IOB	Policy and Operations Evaluation Department
IsDB	Islamic Development Bank
JICA	Japan International Cooperation Agency
KfW	Kreditanstalt für Wiederaufbau Development Bank (Germany)
KOICA	Korean International Cooperation Agency (Korea)
MEAE	Ministere de l'Europe et des Affaires Étrangères (France)
NORAD	Norwegian Agency for Development Cooperation (Norway)
ODA	Official development assistance
SDC	Swiss Agency for Development and Co-operation (Switzerland)
SEO	Special Evaluation Office (Belgium)
SIDA	Swedish International Development Cooperation Agency (Sweden)
UNDP	United Nations Development Programme

Executive summary

Key findings

The study shows an increase in the number of centralised evaluations undertaken by participating organisations. By contrast, the number of joint evaluations has declined, highlighting room for increased collaboration. On average, most multilateral organisations conduct more evaluations than bilateral organisations, in line with their respective mandates. However, beyond this, there is no consistent, identifiable relationship between the volume of development co-operation provided by an organisation, the age of the organisation and the number of centralised evaluations conducted.

There is strong consistency in the purposes and principles that guide evaluations across organisations, which are largely based on longstanding OECD policy guidance documents, demonstrating their continued relevance. The evaluation principles of *independence* and *usefulness* are the most frequently cited. This is in line with the findings of previous surveys, indicating consensus amongst the global evaluation community on the core elements of evaluation systems, and which principles are most vital for quality.

All organisations noted that evaluation units are structurally and functionally independent, with many reporting to an independent oversight committee. There is also strong consistency in institutional arrangements to govern evaluation, with nearly all organisations having evaluation policies and guidance. This shows the core place for evaluation in the work of development co-operation providers.

The increased focus on the usefulness of evaluations reflects persistent challenges in meeting learning objectives. Participating organisations are taking concerted action to address this, ensuring that usefulness is considered earlier in the evaluation process, engaging more with end-users, ensuring that evaluations reflect policy priorities and that findings are timed to fit in with programme cycles.

While there are clear commonalities in the purposes, principles, and institutional arrangements that govern evaluation systems, there are differences in how participating organisations conduct evaluations, often linked to an organisation's size and resources. For example, larger organisations report doing most evaluation work in-house, while smaller ones oversee external evaluators.

In recent years, the use of virtual data collection methods has increased. While this shift was underway before the COVID-19 pandemic, it was accelerated by the onset of the crisis, which demanded real-time evidence on the success of response and recovery options, and restricted travel by member organisations' staff. Many organisations have noted that they will continue to use the innovations developed during the pandemic.

Finally, limited partner country engagement in evaluations is hindering full ownership of development co-operation. Most organisations note that partner country governments are asked to facilitate visits and data collection, but are not engaged meaningfully in the substantive planning or follow-up of evaluations.

Areas for future action

These findings point to five action points for EvalNet members and evaluators more broadly:

- **Work towards a more holistic understanding of data and evidence within an organisation to support *learning* and *strategic planning*.** While there might be a need to maintain the distinct governance structures and roles of these different functions, there may be value in exploring a common approach to collecting evidence across different units, in support of more systematic and efficient data collection in pursuit of cross-cutting learning objectives. This requires working closely with other parts of the institution, such as statistics, results reporting, research and decentralised evaluation units.

- **Develop a common approach to assessing the resources spent on evaluation activities.** Data collection and analysis for this study brought to light the many different ways in which the amount spent on evaluation by participating organisations is calculated. The objective of collecting this information is to understand the share of spending on various elements (e.g., human resources, communications) and whether overall resourcing is sufficient for the evaluation function to fulfil its roles. Creating a cost benchmark by type of evaluations could also be useful.

- **Consolidate learning on virtual ways of working.** Reports vary on the value of virtual engagement and data collection during the COVID-19 pandemic. While some view virtual methods as important for broadening stakeholder engagement, others perceive them as undermining quality. Despite these mixed reports, nearly all organisations plan to continue using virtual methods. In light of this, it may be useful to explore when virtual methods work best and when in-person evaluation work is necessary, and to develop guidance.

- **Track when and how evaluation findings are used and the overall state of evidence-informed decision making.** A recurring theme in this report is the challenge of meeting learning objectives: many respondents are searching for new and more effective approaches to increase the use of findings. With new ways of stakeholder (both internal and external) engagement and dissemination being tested, the time is right to measure which methods work best, with a view to starting each evaluation with a clear uptake pathway.

- **Revisit questions of ownership and its links to partner country engagement.** Greater engagement with partner countries in evaluations will promote national ownership of development activities and bolster evaluation capacity. This will help meet development co-operation commitments to country ownership and localisation.

1 Background and methodology

This study is the most recent edition in a regular series that aims to prompt critical thinking around development co-operation and humanitarian aid evaluations, with a view to strengthening evaluation systems.

1.1. Background

Evaluation is a process that critically examines an intervention[1] (e.g., a policy, programme or project) to better understand if and how it has delivered its intended results. Accountability and learning are the dual objectives of evaluation. Across the public sector, international development co-operation and humanitarian assistance have regularly been subject to relatively high levels of evaluation. This reflects the unique accountability dynamics of international co-operation, and the high degree of domestic scrutiny associated with this external spending. Evaluation is more critical than ever in the current global co-operation context, characterised by more complex and interlinked humanitarian and development challenges and increasingly scarce development co-operation resources.

Evaluation Systems in Development Co-operation is the most recent in a series of similar reviews and provides an update to reports published in 2010 and 2016 (Box 1.1). The overall aim of this series is to prompt critical thinking to strengthen evaluation systems for members of the Development Assistance Committee (DAC) Network on Development Evaluation (EvalNet) and beyond.

The study provides details on the entities, structures, mandates, policies, procedures, and resources that support the evaluation function in participating development organisations. It includes four substantive chapters organised around the following issues: the role of evaluation in development co-operation, covering the number and types of evaluations conducted; the policy and institutional arrangements put in place to govern evaluation systems; the evaluation process, from deciding on which evaluations to undertake to how findings are disseminated; and the use of evaluation findings to maximise development effectiveness and impact.

The core of the study is a comparative report that presents overall trends in evaluation systems, covering sustainable development and humanitarian assistance, with a focus on common challenges, good practice, and innovative ways of working. This is complemented by individual profiles of all participating organisations, which provide an in-depth look at the policies, principles, and ways of working.

Box 1.1. DAC reviews of evaluation systems: a brief history

The DAC Network on Development Evaluation conducts periodic reviews of its members' evaluation systems to support its efforts to increase the effectiveness of development co-operation policies and programmes by promoting high-quality, independent evaluation and the development of institutional evaluation systems. Back in in 1975, at a time when evaluation policy and practice were still nascent, the DAC published its first study: "Aid Evaluation – The Experience of Members of the Development Assistance Committee and of International Organisations."

The DAC first convened an expert group on evaluation in 1982, which took stock of member evaluation practices. Once formally established in 1983, the OECD DAC Network on Development Evaluation (EvalNet) periodically reviewed how members were conducting evaluations, both in terms of methods and systems. EvalNet meetings often featured peer exchanges and member updates on these topics.

Evaluation Systems in DAC Members' Agencies was published in 2004. This paper discussed some of the challenges members' evaluation systems were facing, drawing mainly on data contained in DAC Peer Reviews between 1996 and 2004, as well as information provided over the years to the Secretariat by Evalnet members.

The next study in this series, ''A comparative Study of Evaluation in Policies and Practices in Development Agencies'' was conducted in 2007. The study was commissioned by France to map and compare evaluation practices across development agencies, primarily to stimulate an internal debate within Agence Française de Développement (AFD) on its evaluation systems during a time of reform.

Nine EvalNet agencies were reviewed as part of this study, and results confirmed that there was indeed variation among the strategies and practices adopted by different evaluation units.

In 2010, *Better Aid: Evaluation in Development Agencies* used an approach that has been used ever since. The report provided a snapshot of management and resourcing in 30 development evaluation agencies. It covered the topical trends in evaluation – how evaluations are moving away from project outputs to assess broader impacts of development assistance, adapting to new aid modalities and cross-cutting issues, and increasingly involving country partners. The report identified areas that required collaborative efforts, such as improving communication and supporting the use of evaluation findings to enhance the contribution of evaluation to achieving development goals.

Next in the series, *Evaluation Systems in Development Co-operation*, was conducted in 2016. The study reviewed the evaluation systems of 37 bilateral and 9 multilateral development agencies, providing insights into the role and management of evaluations. It covers topics such as institutional structures and governance, resourcing and management, and evaluation policies and processes. Additionally, the study reflects on general trends in evaluation systems and includes a comparative analysis with the 2010 study.

1.2. Scope and methodology

This study covers 34 of the 41 EvalNet members and observers[2], representing 51 distinct institutions providing development co-operation and humanitarian assistance (Table 1.1). Throughout the text, "participating organisation" refers to those evaluation units on which data were collected, which includes both EvalNet members and EvalNet observers. For a full list of members and observers see Annex A.

To note, OECD member countries disburse and implement their development co-operation and humanitarian assistance in a variety of ways, often via multiple ministries or agencies. Their evaluation systems, and the institutional set up of central evaluation units, reflect these institutional structures. Several EvalNet members have multiple organisations that conduct relevant evaluations. In these cases, all relevant organisations have been included where possible, and are presented together in the member profile (Annex C). While the study does ask participants to report on the overall number of centralised and decentralised evaluations undertaken in 2021, it does not cover decentralised evaluation systems, which in some cases account for the vast majority of evaluations conducted by a participant, as discussed further in Chapter 2.

The study included a literature review of policy documents provided by members and published information on member and observer evaluation systems in 2021. This was followed by a qualitative survey with a questionnaire sent to participating organisations in 2022. Survey results were reviewed and enhanced through individual semi-structured interviews with 28 heads of evaluation units. Data profiles were prepared by the authors for all participating organisations, which were then validated by relevant organisation staff. It is important to note that this study focuses on *centralised* evaluation units. It does not attempt to provide detailed information on the results, monitoring and evaluation systems, including decentralised evaluations, undertaken by other parts of the participating organisations.

Figure 1.1. Research methods

All data and evidence presented in this report draw on the information collected specifically as part of the report process unless otherwise stated or referenced. Where a comparison is drawn with 2010 or 2016, this refers to data and evidence presented in the 2010 and 2016 iterations of this study (Box 1.1). It is important to note that the datasets from 2010, 2016 and 2022 vary slightly depending on EvalNet participation at the time of the study. Where relevant, the 2022 dataset has been adapted to ensure that comparisons over time are meaningful. Additionally, the data reported in the main report are current as of 15 March 2023. However, as minor revisions to country profiles were received after this date, there might be slight discrepancies in the trends reported in the report and the individual data reported in the profiles.

Table 1.1. Participating organisations

MEMBERS		
Member	**Organisation**	**Short Name**
Australia	Department of Foreign Affairs and Trade	Australia – DFAT
Austria	Federal Ministry for European and International Affairs	Austria – BMEIA
	Austrian Development Agency	Austria – ADA
Belgium	Foreign Affairs, Foreign Trade and Development Cooperation	Belgium – FPS Foreign Affairs
	Belgian Development Agency	Belgium – ENABEL
Canada	Global Affairs Canada	Canada – GAC
Czechia	Ministry of Foreign Affairs	Czechia – MFA
Denmark	Ministry of Foreign Affairs	Denmark – MFA
European Commission	European Commission	EC
Finland	Ministry for Foreign Affairs	Finland – MFA
France	Ministère de l'Europe et des Affaires Étrangères	France – MEAE
	French Development Agency	France – AFD
	Direction Générale du Trésor	France – Treasury
Germany	Federal Ministry for Economic Cooperation and Development	Germany – BMZ
	Deutsche Gesellschaft für Internationale Zusammenarbeit	Germany – GIZ
	German Institute for Development Evaluation	Germany – Deval
	KfW Development Bank	Germany – KfW
Iceland	Ministry for Foreign Affairs	Iceland – MFA
Ireland	Department of Foreign Affairs	Ireland – DFA
Italy	Ministry of Foreign Affairs and International Co-operation	Italy – ESTERI
	Italian Agency for Development Co-operation	Italy – AICS
Japan	Ministry of Foreign Affairs	Japan – MOFA
	Japan International Cooperation Agency	Japan – JICA

MEMBERS		
Member	**Organisation**	**Short Name**
Korea	Korean International Cooperation Agency	Korea – KOICA
	Economic Development Cooperation Fund	Korea – EDCF
Luxembourg	Ministry of Foreign and European Affairs	Luxembourg – MFEA
	Luxembourg Development Co-operation Agency	Luxembourg – LuxDev
Netherlands	Ministry of Foreign Affairs	Netherlands – MFA
New Zealand	Ministry of Foreign Affairs and Trade	New Zealand – MFAT
Norway	Norwegian Agency for Development Cooperation	Norway – Norad
Portugal	Institute for Cooperation and Language (Camões, I.P.)	Portugal –- Camões, I.P.
Spain	Ministry of Foreign Affairs, European Union and Cooperation	Spain – MAEC
Sweden	Swedish International Development Cooperation Agency	Sweden – SIDA
	Expert Group for Aid Studies	Sweden – EBA
Switzerland	Federal Department of Foreign Affairs	Switzerland – SDC
	State Secretariat for Economic Affairs	Switzerland – SECO
United Kingdom	Foreign and Commonwealth Development Office	UK – FCDO
	Independent Commission on Aid Impact	UK – ICAI
United States of America	Department of State	USA – DoS
	United States Agency for International Development	USA – USAID
	Millennium Challenge Corporation	USA – MCC
	Development Finance Corporation	USA -- DFC
OBSERVERS		
Participating organisations		
Organisation	**Short Name**	
African Development Bank	AfDB	
Asian Development Bank	ADB	
European Investment Bank	EIB	
European Bank for Reconstruction and Development	EBRD	
Inter-American Development Bank	IDB	
International Monetary Fund	IMF	
Islamic Development Bank	IsDB	
United Nations Development Programme	UNDP	
World Bank	World Bank	

Notes

[1] For the purposes of this report, references to development evaluation will be understood to mean evaluations of both development co-operation and humanitarian assistance activities, including contributions to multilateral institutions and United Nations (UN) organisations. Several members also have a mandate to evaluate the development effects of non-development specific policies and actions, including trade and foreign policy.

[2] All 41 EvalNet DAC members and observer organisations participating in EvalNet as of end 2021 were invited to take part in the study. Seven members and observers with more nascent evaluation systems and limited data are not included in the overall trends analysis: the Asian Infrastructure Investment Bank (AIIB), Greece, Hungary, Poland, Slovak Republic and Slovenia. Lithuania joined the DAC at the very end of the study period and is therefore also not included.

2 The role of evaluation

This chapter focuses on the role evaluation plays in ensuring effective and impactful development co-operation. The dual evaluation objectives of accountability and learning remain steadfast. However, persistent challenges in ensuring consistent use of evaluation findings have led to greater focus on the learning objective. This chapter also reviews the overall number of evaluations undertaken by different types of organisations, setting the context for subsequent chapters that explore the driving forces and processes behind evaluations.

2.1. Evaluation purpose

Evaluation plays an essential role in development co-operation. It is a vital tool for maximising the impact of development co-operation efforts. At its core, development evaluation is a process that critically examines a policy, programme or project to better understand whether, and how, it has delivered the intended results. Institutions have put in place unique governance arrangements to ensure evaluation systems provide relevant, objective, and impartial insight into the performance of development efforts, while connecting country-level data and evidence with decision makers, who are often removed from day-to-day implementation and results.

> "The Department uses evaluation to assess its policies, strategies, programmes, projects and other initiatives to generate evidence that provides both accountability for public funding and learning to inform strategic and operational decision making."
>
> Ireland, Department of Foreign Affairs

Evaluation helps answer critical questions about interventions and their results. These questions are captured in the six criteria defined by the OECD-DAC (Figure 2.1). Evaluation questions include: Is the intervention doing the right things? How well does the intervention fit? Is it achieving its objectives? Is it implemented coherently and efficiently? Is it having positive impacts that last? (OECD, 2021[1]).

Figure 2.1. OECD evaluation criteria
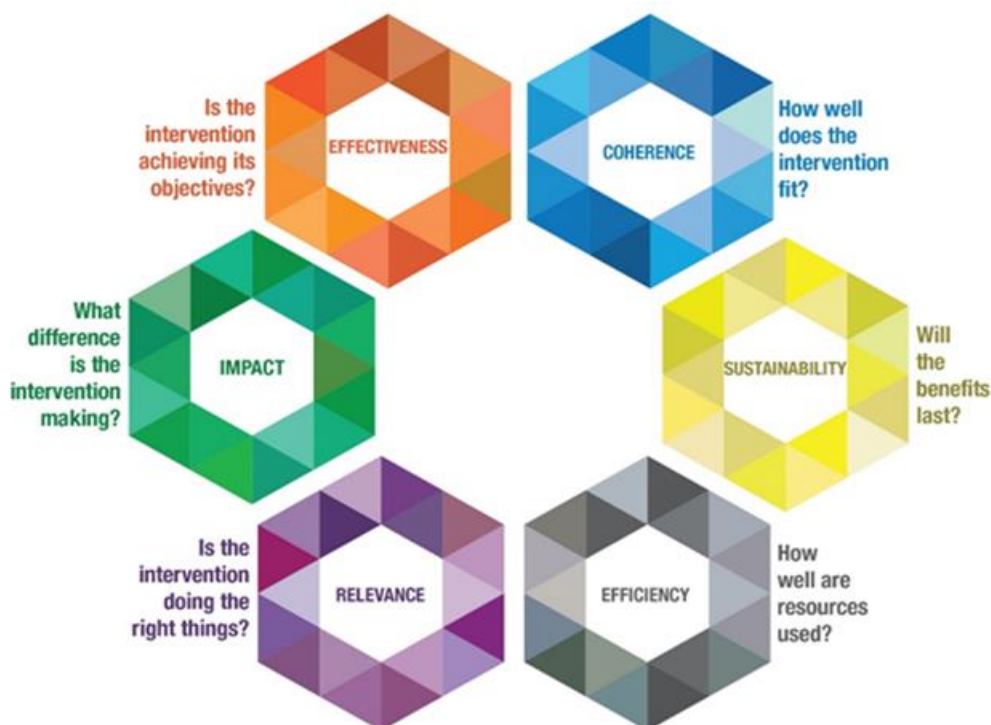
Source: OECD (2021[1]), *Applying Evaluation Criteria Thoughtfully,* OECD Publishing, https://doi.org/10.1787/543e84ed-en

In the current global context, evaluation is more critical than ever. Eight years ago, countries around the world adopted the 2030 Agenda for Sustainable Development. As we pass the halfway point in implementing this ambitious agenda, it is clear that progress has not been fast enough. The global community remains off track to achieve the Sustainable Development Goals (SDGs) by 2030. Persistent poverty and inequality, deadly conflicts, gender inequality and the climate crisis all threaten people. These challenges have been further exacerbated in recent years by the COVID-19 pandemic, growing disaster risks from climate change, rising food and fuel prices, and violent conflicts that have led to record levels of people being displaced.

The development financing landscape puts further stress on development initiatives. Aid budgets face downward pressure in many countries, while development organisations are pushed to demonstrate results and value for money. In this context, it is important to think critically about development co-operation systems and structures to ensure that each development dollar is spent on relevant priorities and has maximum impact.

Accountability and learning are the dual objectives of development evaluation. Accountability in development co-operation, between governments and development partners, as well as towards citizens, civil society and other development stakeholders, is essential for effective development activities and maximising impact. Accountability in development co-operation means ensuring all resources are used efficiently and as intended. It also goes beyond this, ensuring that results are achieved and co-operation delivered in a way that supports inclusive, green and sustainable development.

> "Evaluations contribute to the ongoing streamlining of innovation of the development co-operation programme as a whole by making recommendations for the improvement of future development interventions."
>
> Czechia, Department of Development Cooperation and Humanitarian Aid

Learning is an equally important evaluation objective. Development evaluations provide important data and evidence on why an approach has worked or has not, as well as whether results are contributing to overall development goals and whether they are sustainable. When evaluation findings are used to inform future planning and resourcing decisions, the quality of co-operation is improved. In this vein, evaluation findings, including lessons learned and success factors, are used alongside complementary research to design future policies, programmes and projects.

In line with past findings, participating organisations highlighted that both accountability and learning objectives remain central. The 2016 study noted that organisations aimed to find a balance between accountability and learning (OECD, 2016[2]). Similarly, all respondents in 2022 noted that the purpose of evaluations includes both accountability and learning dimensions. However, interviews confirm an increasing focus on the learning function of evaluation (further details can be found in Chapter 5), as well as on results reporting.

The recent focus on learning reflects challenges in systematically using evaluation findings to inform decision making and strengthen overall effectiveness. During the 2022 data collection process, including written submission and interviews, participants highlighted a strong focus on the use of evaluation findings to drive organisational learning and to support evidence-based decision making. This is not to say

that one objective is more important than the other. Rather, it indicates that more work can be done to facilitate the use of evaluation findings (see Chapter 5).

2.2. Number and type of evaluations

The overall number of centralised and decentralised evaluations undertaken by participating organisations has increased in the past decade.[1] This demonstrates the significant value participating organisations place on evaluative evidence and reflects a growing appetite for evaluations from programme staff and other audiences. The number of evaluations conducted in the 2017-2021 period increased 82 percentage points over the 2011-2016 period (Figure 2.2). Over 2017-2021, the number of evaluations undertaken remained relatively steady (Figure 2.3), despite a slight drop in 2020 – likely due to the COVID-19 pandemic – and a jump in 2021 when travel restrictions were eased in most member countries allowing for evaluation units to catch up on some planned evaluations.

Figure 2.2. Total number of evaluations, 2011-2016 vs 2017-2021
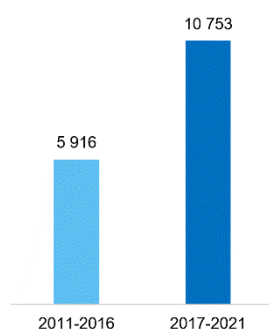
Note: Data on the number of evaluations conducted were reported by 34 members and observers in 2022, and these varied slightly from those reporting in 2016, the two time periods are therefore not perfect comparisons. However, the conclusion of an increase in the total number of evaluations increasing is supported by other data sources including interviews and member submissions to EvalNet.

Figure 2.3. Total number of evaluations, 2017-2021
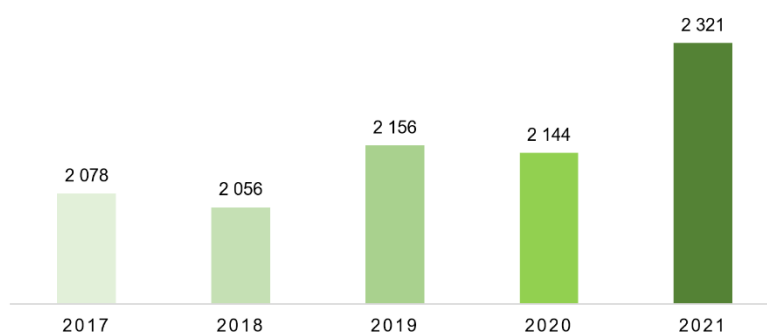
Note: Data on the number of evaluations conducted were reported by 34 members and observers, although the datasets vary slightly between 2011-2016 and 2017-2021. Not all participating centralised evaluation units reported on the number of decentralised evaluations.

Figure 2.4. Total number of centralised evaluations conducted in 2021, bilateral organisations

Organisation	Value
Korea - KOICA	105
Germany - GIZ	85
Germany - KfW	80
Japan - JICA	73
Belgium - ENABEL	17
France - MEAE	15
Luxembourg - LuxDev	13
Netherlands - MFA	12
France - AFD	12
USA - USAID	9
Sweden - EBA	8
Finland - MFA	6
Japan - MOFA	5
Germany - BMZ & DEval	5
Czech Republic - MFA	5
European Commission	4
Switzerland - SDC	4
Luxembourg - MFEA	4
Iceland - MFA	4
Denmark - MFA	4
Canada - GAC	4
Sweden - SIDA	3
Korea - EDCF	3
Italy - ESTERI	3
Ireland - DFA	3
France - DG Treasury	3
Belgium - FPS Foreign Affairs	3
Portugal - Camões, I.P.	2
Norway - NORAD	2
Austria - BMEIA	2
Switzerland - SECO	1
New Zealand - MFAT	0
Austria - ADA	0
Australia - DFAT	0

Note: Data on the number of evaluations conducted were reported by 35 out of 42 bilateral organisations responding to the questionnaire (several members reported data from multiple institutions).

Figure 2.5. Total number of centralised evaluations conducted in 2021, multilateral organisations
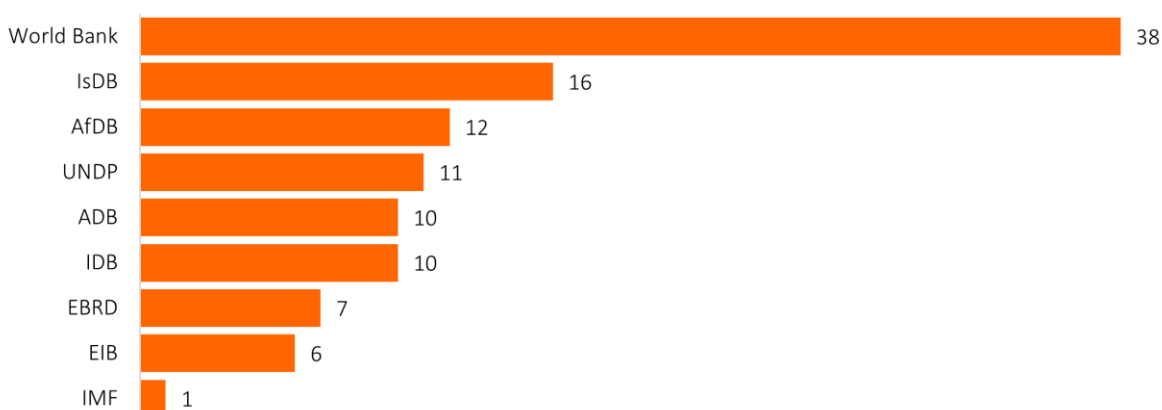
Note: Data on the number of evaluations conducted was reported by 9 out of 9 multilateral organisations that responded to the questionnaire. See Table 1.1 for multilateral acronyms.

The number of evaluations conducted by reporting institutions varies greatly across organisations. Of the responding evaluation units, only a handful conduct more than 50 centralised evaluations (Figure 2.4,Figure 2.5). Conversely, 27 respondents conducted 10 evaluations or fewer (see Annex C). This reflects the significant diversity in the mandates, roles, and resourcing of evaluation units and the size and composition of the portfolios covered, rather than the overall size of the organisation per se. For example, KfW's evaluation unit is mandated to evaluate the programmes and projects that it implements on behalf of BMZ, while DEval has a crosscutting function, covering all German development assistance (Box 2.1). In contrast, in the Netherlands, one evaluation unit is mandated to cover all Dutch development assistance, as well as foreign policy and trade activities that impact partner countries.

The age and size of bilateral organisations do not systematically affect the number of evaluations conducted, which is instead driven by the role evaluation plays within an organisation and the institutional set-up. Analysis was conducted to determine whether the age of the organisation and the amount of official development assistance (ODA) provided (both in terms of total amount and as a share of gross national income – GNI) influences the number of evaluations undertaken. No such relationship was found. This reflects the significant diversity in how development organisations set up and use their evaluation function. For example, some participating organisations conduct many centralised evaluations themselves. Other organisations focus more on decentralised evaluations, with centralised units providing an oversight or quality assurance function, or only conducting a smaller number of strategic evaluations.

Multilateral organisations conduct more evaluations overall than bilateral organisations, and implementing agencies conduct more evaluations than those with a policy-focused role. On average, in 2021, multilateral organisations conducted 123 evaluations (centralised and decentralised), compared to 35 for bilateral organisations. The Asian Infrastructure Investment Bank, which is a relatively new institution, has just begun conducting evaluations and is therefore not included in the overall dataset for this study (Box 2.2). When comparing the number of evaluations undertaken by organisations with different roles (primarily policy, primarily implementation or a dual role), analysis found that implementing agencies conduct more evaluations, on average (Figure 2.6). As can be seen in the individual profiles (Annex C), 2020 – and to a lesser extent 2021 – saw a dip in the number of evaluations conducted, due to pandemic-related challenges including travel restrictions.
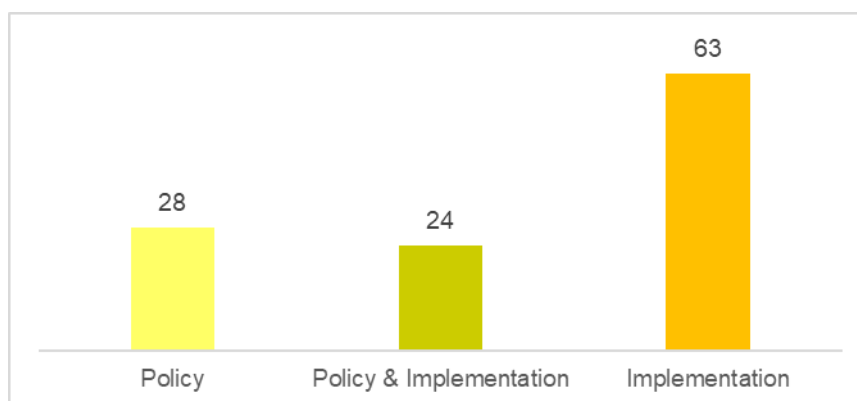
Figure 2.6. Average number of evaluations conducted in 2021, by role

Note: Data on number of evaluations conducted were reported by 44 out of 51 reporting organisations. Not all participating centralised evaluation units reported on the number of decentralised evaluations.

Most organisations use both centralised and decentralised evaluations. Centralised evaluations, undertaken by units in headquarters, often focus on high-level policies, strategies or themes. Decentralised evaluations are undertaken by evaluation units of implementing agencies and by programmatic units or country offices, and often focus on specific sectors, programmes or projects.

There is significant variety in the types and methodologies of evaluations conducted. Across participating organisations, centralised evaluation units conduct many different types of evaluations, including policy/strategy evaluations, sector programme evaluations, country programme evaluations, project evaluations, process evaluations, thematic evaluations, cluster evaluations, impact evaluations; and syntheses or meta-evaluations (Box 2.3).

Box 2.1. Evaluation in Germany

Overall responsibility for development evaluation in Germany falls under the remit of the Federal Ministry for Economic Cooperation and Development (BMZ). Under the direction of BMZ, other key actors in this system are: the German Institute for Development Evaluation (DEval); the German Corporation for International Cooperation (GIZ); and the KfW Development Bank (KfW).

The Evaluation Unit within BMZ provides overall direction to the evaluation system. This includes setting standards and overseeing their implementation, as well as ensuring coherence among evaluations undertaken by all organisations within the system. Since the foundation of DEval in 2012, BMZ's evaluation unit only conducts evaluations in exceptional cases.

DEval is mandated by the German federal government, through BMZ, to conduct evaluations of German development co-operation. It conducts a variety of types of evaluations, contributes to the setting of evaluation standards and conducts capacity building. GIZ and KfW are the two largest implementing agencies of BMZ. Both organisations conduct evaluations related to their work and develop learning processes and products for their own activities.

Box 2.2. Evaluation at the Asian Infrastructure Investment Bank

The Asian Infrastructure Investment Bank (AIIB) was established in 2016 and is still in the process of fully establishing its learning and evaluation function. It is therefore not included in the overall dataset for this study.

The Complaints-Resolution, Evaluation and Integrity Unit was envisioned in the Bank's Articles of Agreement. A Terms of Reference (ToR) was included in AIIB's 2019 Oversight Mechanism and a Learning and Evaluation Policy was approved in 2021. The policy recognised the importance of building a fit-for-purpose approach and a corporate and learning culture to support institutional performance, continuous improvement, and credibility for this young Bank — all aspects of accountability. A small professional staff implements the Learning and Evaluation Policy, working with task-specific consultants, consistent with AIIB's portfolio size and values.

The Learning and Evaluation team reports to the Board quarterly on annual workplan implementation, providing Members and Bank staff with valuable lessons from its activities and those of peer multilateral development banks. For example, it conducts Early Learning Assessments annually to identify findings, lessons and evaluability in selected on-going Bank financing. Going forward, it will undertake Project Learning Reviews (independent post-evaluations) of completed Bank financing. The Policy and Strategy Committee of the Board discusses each assessment and review. Similarly, peer independent evaluation departments are invited to discuss their evaluation findings and recommendations with AIIB staff in quarterly Practitioner Dialogues. The unit is also an EvalNet observer and a member of the COVID-19 Global Evaluation Coalition.

As was found in 2016, the links between centralised evaluation units and colleagues conducting decentralised evaluations are often weak. This is usually because there is no comprehensive evaluation workplan that includes all evaluations undertaken by the organisation. In some cases, there is no full accounting of all decentralised evaluations undertaken. However, some centralised units, for example in FCDO, do provide guidance, advice and quality assurance support to other parts of the organisation, allowing for a more complete understanding of the evaluation portfolio across departments.

> "A large majority of the decentralised evaluations are deemed to be useful. However, insights from decentralised evaluations are not systematically exploited for organisational learning but [instead] remain on the individual level."
>
> Finland, Development Evaluation Unit

About half of evaluations conducted in 2021 were decentralised, though due to limited reporting this is likely an underestimate**.** Of the total 2 321 evaluations reported by participating organisations, 760 (33 %) were conducted or commissioned by centralised evaluation units. Some 1 032 decentralised evaluations were conducted (44 %). However, as noted, this study focuses on centralised evaluation units in headquarters. While data were collected on both centralised and decentralised evaluations, detailed information on evaluations undertaken by specific programmatic units, country offices or projects was not specifically included in the data collection and therefore is likely undercounted in these figures.

Box 2.3. Evaluation types

In June 2023, the OECD published the second edition of the *Glossary of Key Terms in Evaluation and Results-Based Management*. This glossary provides definitions for the common types of evaluations:

- **Cluster evaluation:** An evaluation of a set of related activities or interventions, either similar interventions in different locations or a cluster of complementary components of an overall initiative.

- **Country programme evaluation:** Evaluation of one or more institution's or partner's portfolio of interventions in a specific country, including the strategy behind them in a specific period of time.

- **Impact evaluation**: An evaluation that assesses the degree to which the intervention meets its higher-level goals and identifies the causal effects of the intervention. Impact evaluations may use experimental, quasi-experimental and non-experimental approaches.

- **Meta-evaluation**: The term is used for evaluations designed to synthesise findings from a series of evaluations. It can also be used to denote the assessment of an evaluation to judge its quality or scrutinise the performance of the evaluators. **Process evaluation**: An evaluation of the internal dynamics of implementing organisations, their policy instruments, their service delivery mechanisms, their management practices and the linkages among these.

- **Programme evaluation**: Evaluation of a set of interventions, combined to attain specific global, regional, country, or sector development objectives.

- **Project evaluation**: Evaluation of an individual intervention designed to achieve specific objectives within specified resources and implementation schedules, often within the framework of a broader programme, examining its relevance, coherence, effectiveness, efficiency, impact and sustainability.

- **Sector programme evaluation**: Evaluation of a cluster of interventions within one country or across countries, all of which contribute to the achievement of a specific goal.

- **Thematic evaluation**: Evaluation of a selection of interventions, all of which address a specific sustainable development priority or topic, that cuts across countries, regions, and sectors.

Source: OECD (2023[3]) Glossary of Key Terms in Evaluation and Results-based Management for Sustainable Development (Second Edition) https://doi.org/10.1787/632da462-en-fr-es

References

OECD (2023), *Glossary of Key Terms in Evaluation and Results-based Management for Sustainable Development (Second Edition)*, OECD Publishing, Paris, https://doi.org/10.1787/632da462-en-fr-es.   [3]

OECD (2021), *Applying Evaluation Criteria Thoughtfully*, OECD Publishing, Paris, https://doi.org/10.1787/543e84ed-en.   [1]

OECD (2016), *Evaluation Systems in Development Co-operation: 2016 Review*, OECD Publishing, Paris, https://doi.org/10.1787/9789264262065-en.   [2]

Notes

[1] Participants were asked to report on the total number of evaluations conducted by their organisations in 2021 – both centralised and decentralised. However, kindly note that not all participating centralised evaluation units reported on the number of decentralised evaluations, and several have provided only an estimate.

3 Policy and institutional arrangements

The policy and institutional arrangements that govern evaluation systems are important for how they support the intended role of evaluations. There is strong coherence between the principles that inform evaluations and the kinds of policies and guidance documents that steer and support high-quality evaluation processes. However, participating organisations diverge in the human and financial resources allocated for evaluation.

3.1. Principles and policies

Integrating evaluation systems into organisational architecture supports evaluation objectives. As noted in Chapter 2, accountability and learning are the dual objectives of evaluation, with an overall aim of supporting better development co-operation. The policy and institutional arrangements surrounding evaluation systems can influence whether these objectives are achieved.

Commitment to the OECD DAC principles that guide development evaluations remains steadfast. As part of its mandate to increase the volume and effectiveness of development co-operation, and with an understanding of the importance of evaluation for enhancing the quality of development efforts, the DAC adopted five *Principles for Evaluation of Development Assistance* in 1991 (OECD, 1991[1]). These are: impartiality, independence, credibility, usefulness and partnership. These principles lay out a definition of evaluation, including the core criteria of relevance, effectiveness, efficiency, impact and sustainable – to which coherence was added in 2019. As seen in previous studies, the criteria are very widely used, and included in every reporting evaluation unit's policy or guidelines.

Today, independence and usefulness are the principles most often cited by participating organisations. All DAC evaluation principles continue to be used by participating evaluation units to inform their evaluation activities. These principles are referenced in the majority of DAC Network on Development Evaluation (EvalNet) participant evaluation policies or other normative documents. However, not all principles are given equal weight. The principles of independence and usefulness (utility) were both mentioned 28 times when participants were asked which principles guide their evaluations (Figure 3.1). These were followed by credibility (23 mentions), impartiality (19 mentions), partnership (19 mentions) and transparency (18 mentions). In addition to these principles, many organisations also now include principles related to value for money, ethics, and human rights.

Figure 3.1. Main guiding evaluation principles as reported by EvalNet participants

Note: the size of the words reflects the number of times they were mentioned.

Utility is a major driving force of evaluation decisions, continuing the trend towards learning. While all five principles are referred to by participating organisations, many emphasise utility or usefulness as a priority consideration today (see Box 3.1 for an example). Utility requires that evaluations respond to continually evolving organisation priorities and needs. In this vein, the timeliness of evaluations is of critical importance, and is an additional sub-principle that is necessary to ensure utility (further details can be found in Chapter 5). The importance of timeliness was a key lesson learned during the COVID-19 pandemic (see Chapter 4).

Box 3.1. Ensuring the utility of the United Kingdom's evaluations

When asked about what principles inform their evaluation activities, the United Kingdom (UK) responded that evaluations must be useful. They defined "useful" as including the following: responds to organisational priorities; fills evidence gaps; and provides opportunities for influencing change. In order to meet these criteria, an evaluation must ask well-defined questions that are feasible to answer; be timed in order to inform internal and external decision making; engage internal and external stakeholders in a participatory approach; and include actionable recommendations. In short, the use of evaluation findings must be central throughout the design and implementation of an evaluation. Other principles that guide evaluation in the UK are credibility, robustness, proportionality, safety and ethics.

Evaluation principles are codified in evaluation policies. The cornerstone of evaluation systems is the organisational evaluation policy, which sets out the purpose and principles that guide evaluation. Some elements are common to many organisations' evaluation policies: evaluation purpose; evaluation focus (what will be evaluated and how evaluation decisions will be made); institutional set-up of evaluation units; roles and responsibilities; evaluation process (discussed in Chapter 4); evaluation quality standards; evaluation dissemination; and evaluation follow up.

In addition to codifying the *Principles for Evaluation of Development Assistance,* evaluation policies reflect the *Quality Standards for Development Evaluation* (OECD, 2010[2]). Adopted in 2010 as part of the DAC Guidelines and Reference Series, the Quality Standards outline key pillars needed to ensure high-quality evaluation processes and products. They have been systematically adopted and integrated into participating organisations' systems through their policies and other normative documents. The Quality Standards are intended to guide practitioners across the various stages of the evaluation process, including in the creation of policy documents.

The number of organisations with an evaluation policy in place has increased. The high number of policies in place highlights the importance given by development organisations to evaluation as a core function. The importance of evaluation systems is also emphasised by the fact that its existence is a soft criterion for DAC membership. In 2021, all but four organisations indicated having an evaluation policy, with one in the process of being designed. Those that do not have a policy often have other internal documents that guide the evaluation function. There has been a steady increase in institutionalised policies for evaluation units since 2010 (Figure 3.2).

Figure 3.2. Proportion of organisations with evaluation policies, 2010-2021
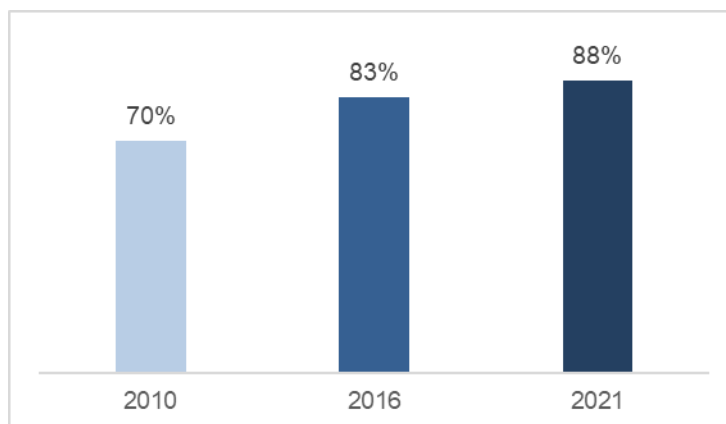

Note: Data on evaluation policies were reported by 49 out of 51 organisations that responded to the questionnaire.

Evaluation policies are often complemented by concrete evaluation guidance. Nearly all organisations have developed comprehensive guidance to ensure consistently high-quality evaluations. Organisations that do not have evaluation policies typically provide evaluation guidance that outlines evaluation standards and processes, and in some cases also includes practical tools (e.g., checklists and templates) for use by evaluation commissioners, managers, and evaluators. Organisational guidance reflects global evaluation norms, including the *Quality Standards for Development Evaluation* (OECD, 2010[2]), the *Big Book on Good Practice Standards* (ECG, 2012[3]) and the *Norms and Standards for Evaluation* (UNEG, 2016[4]). These guidelines provide flexible advice that can be adapted depending on evaluation objectives.

Significant effort has been made – and is ongoing – to mainstream human rights, gender, and inclusion into evaluation practices (Figure 3.3). EvalNet meetings often raise these topics and members have reported that there is growing political priority and pressure for evaluation units to better address these topics. The joint work of EvalNet and DAC Network on Gender Equality (GenderNet) in 2020-2021, and EvalNet's new publication on human rights and gender equality (Box 3.2), reflect this growing interest. Of the 47 organisations that indicated that they have evaluation guidance, 44 (90 %) report that their guidance addresses leave no one behind and inclusion issues, and 45 (92 %) address gender (Box 3.2).

Figure 3.3. Inclusion of cross-cutting issues in evaluation guidance
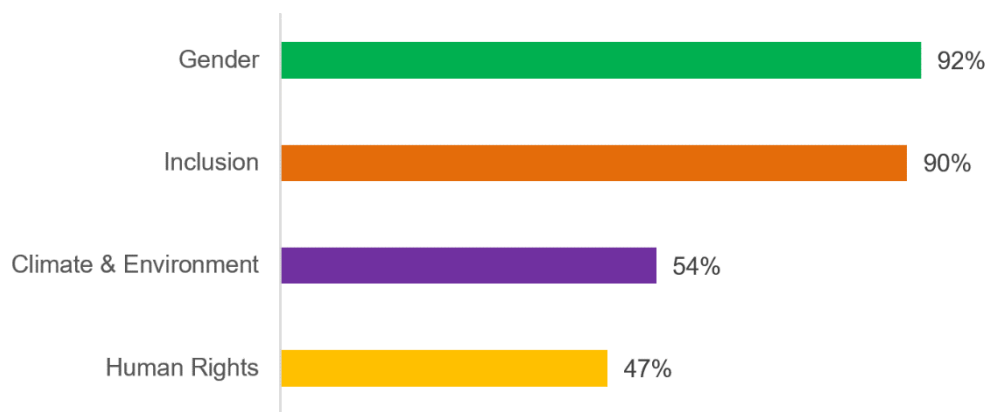
Note: Data on cross-cutting issues were reported by 47 out of 51 participating organisations.

Climate and environment are also now frequently being considered as cross-cutting issues. However, while specific guidance is often provided on mainstreaming gender and inclusion into evaluations, climate and environment is often only considered in a broad way (Figure 3.3). More specific guidance is needed on how these areas can be meaningfully incorporated into evaluations. This reflects ongoing work in the global evaluation community to strengthen monitoring and evaluation of climate efforts, as well as work by EvalNet members to develop their own approaches and strategies.[1] Finland has prepared "Practical tips for assessing cross-cutting objectives in evaluations" (Ministry of Foreign Affairs of Finland, n.d.[5]), which provide evaluation managers and practitioners with resources and tools for integrating cross-cutting objectives, such as climate resilience, low emission development, environment and biodiversity, into evaluations. The document presents key considerations for each stage of the evaluation process, providing examples of evaluation questions.

While consistency and rigour are vital, there is also a recognition that evaluation practices may need to be adjusted to specific contexts. Participating organisations often noted the importance of flexibility in implementing evaluation guidance, as different areas of work and country contexts may call for adapted approaches. In this vein, there has been a heightened focus on the design stage of evaluations to ensure that the decided approach will respond to specific needs and will support evaluation objectives (further details in Chapter 4).

While there is strong coherence in the purpose and principles that guide evaluations, how evaluation units are set up, and their scope to achieve these, varies across organisations. In general, the role of a centralised evaluation unit is to develop evaluation policy, guidance and standards, and to oversee the evaluation function, including providing some level of quality assurance during the evaluation process. Beyond these core functions, there is a degree of variation in how evaluation units function – with some conducting or commissioning evaluations by external consultants, and others primarily supporting evaluations conducted by other parts of the institution.

The inclusion of learning and knowledge management as a core function of evaluation units is increasing. Considering the challenges of using evaluation findings (see Chapter 5), there is increasing focus on communications and knowledge management in evaluation units to facilitate learning. A growing number of evaluation units now include dedicated experts in knowledge management and communication to turn evaluation findings into digestible and actionable knowledge products and to organise learning events.

Box 3.2. Guidance to ensure inclusive evaluations

Mainstreaming gender and inclusion issues in development evaluations is critical for identifying and understanding the extent to which gender and other intersecting identity factors were considered in the design, development and implementation of development policies, programmes and projects, and what the outcome of those considerations were for various groups.

New OECD guidance on evaluating human rights and gender

The OECD (2023[6]) publication Applying a Human Rights and Gender Equality Lens to the OECD Evaluation Criteria, provides in-depth practical guidance for evaluators, evaluation managers, and programme staff in applying a human rights and gender equality lens to the six OECD evaluation criteria (Figure 2.1). It aims to support evaluators and evaluation managers in the design, management and delivery of credible and useful evaluations that assess whether and how interventions contribute to realising human rights and gender equality – be they interventions with explicit human rights objectives or not. It also provides broader guidance to programme staff in applying the six criteria with a human

---

rights and gender equality lens at the outset of an intervention and addresses the main considerations and challenges in doing so.

Gender mainstreaming in Canada

Canada's Feminist Evaluation Framework ensures that feminist principles are central throughout evaluations. Taking this framework forward, the Government of Canada has published guidance on gender-based analysis. Integrating Gender-Based Analysis Plus into Evaluation: A Primer (2019[7]), outlines considerations and methodological approaches, along with practical examples, for how to include gender impacts throughout the evaluation cycle. The guidance is complemented by the *Gender equality and empowerment measurement tool* (2022[8]), an innovative data collection system that incorporates feminist elements of participation, inclusivity, intersectionality, and empowerment.

Disability inclusion in Australia

Australia has developed the *Disability inclusion in the DFAT Development Program: Good practice note* to provide advice on how to engage people with disabilities and their representative organisations throughout policy and programming design, implementation, and monitoring and evaluation. The overall aim of this advice is to identify and address barriers to inclusion of people with disabilities in development efforts.

Sources: OECD (2023[6]), *Applying a Human Rights and Gender Equality Lens to the OECD Evaluation Criteria*, https://doi.org/10.1787/9aaf2f98-en; Treasury Board of Canada Secretariat (2019[7]), *Integrating Gender-Based Analysis Plus into Evaluation: A Primer,* https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/evaluation-government-canada/gba-primer.html#H-04; Government of Canada (2022[8]), *Gender equality and empowerment measurement tool,* https://www.international.gc.ca/world-monde/funding-financement/introduction_gender_emt-outil_renforcement_epf.aspx?lang=eng; DFAT (2021[9]), *Disability inclusion in the DFAT Development Program: Good practice note*, https://www.dfat.gov.au/sites/default/files/disability-inclusive-development-guidance-note.pdf.

3.2. Governance and independence

Evaluation units have implemented multi-pronged strategies to maintain their independence. Independence is essential for ensuring that evaluations are credible and impartial. All organisations report that their evaluations are free from undue influence and biases, including political or organisational pressure, and are independent in their function from the processes associated with policy making, delivery, and the management of development assistance. EvalNet members have a nuanced and comprehensive understanding of independence, which is operationalised across the network through one or more of the following ways of working.

Structural independence

All EvalNet organisations reported that their organisation's evaluation function is organisationally separate from the line function responsible for development assistance management and delivery. This status helps to establish the evaluation unit's independence in the choice, implementation and dissemination of its evaluations. Nearly all Evaluation Departments commission and manage evaluations, making decisions about what to evaluate, as well as evaluation design, questions, and methodological choices.

The independence of evaluation units' activities is reinforced through direct reporting lines to decision makers, such as a senior board of directors. Typically, the Head of Evaluation reports directly to the Secretary-General or the senior board members of the agency, rather than to Programme or Policy Directors. This helps prevent any compromise of the evaluation process and its results and safeguards

independence. In many cases there are also staff-related limitations related to serving as head of evaluation.

Functional independence

EvalNet members reported having their own budget, staff and workplan, which are not overseen or controlled by the organisation's management. In many cases, evaluations are funded from a separate, stand-alone budget line independent of programme or operational budgets.

To further safeguard the independent functioning of the evaluation units and shield them from outside influence, the appointment and removal of the head of the evaluation unit are overseen by an independent jury consisting of members of the senior board of directors and advisory committees. The head of evaluation has authority over the implementation of the work plan, the conduct and content of evaluations and their publication, budget utilisation, and staff appointments and management.

Approximately half of the organisations have an independent advisory committee in place. The purpose of these bodies varies across organisations, but often includes guiding the design of evaluation policies, appointing the head of the evaluation unit, providing input on the evaluation priorities and workplans, and making recommendations to management based on evaluation findings.

Behavioural independence

All EvalNet members reported on their methods for producing high-quality and unbiased reports. Evaluation units generally have the sole responsibility for final evaluation reports. While the board, management and staff may comment and fact check reports, they cannot impose unfounded changes or alter the conclusions. Members work to ensure that evaluators – with evaluation unit staff or consultants – are free to conduct the evaluation work without interference and to express their opinions freely. Any possible conflicts of interest are addressed openly and honestly.

EvalNet members use oversight committees and multistakeholder reference groups to provide independent guidance, scrutinise evaluation design and process, and ensure quality. Independent quality assurance takes place at key points in the evaluation cycle, for example by asking external evaluation experts to review terms of reference and final reports. The French Development Agency (AFD) establishes multi-stakeholder reference groups for each evaluation consisting of diverse external and internal stakeholders. These include parliamentarians, representatives of civil society, researchers, and organisations that directly benefit from the policy being evaluated, and the group is presided over by an independent qualified chairperson. The group is responsible for reviewing all the reports produced by external evaluators and facilitating discussions during meetings. Their perspectives shape the evaluation report and recommendations.

External evaluation specialists are either outsourced or engaged in some capacity to provide a more objective and independent assessment. All participating units engage independent consultants or firms in some way, most often through a competitive public procurement procedure, to conduct evaluations either autonomously or in collaboration with the evaluation unit. External consultant involvement is believed to strengthen the impartial and independent nature of evaluations. When consultants are hired locally (in developing countries), this contributes to the understanding of the context, as well as fostering local ownership of the evaluation. However, the relationship between external consultants and independence is not always straightforward, and hiring consultants is not in and of itself a guarantee of independence or objectivity. Members have long debated how best to involve consultants, as well as their role in ensuring impartiality and limiting bias. They recognise that the impartiality and objectivity of external evaluators can be compromised by information asymmetries and their desire to secure the contract (Picciotto, 2018[10]).

It is best practice to have formal guidelines in place to avoid potential conflicts of interest. Many interviewees described this was the case in their institutions. Conflict of interest safeguards ensure that evaluators' judgement is not influenced by current, recent, or past professional or personal relationships and factors. These safeguards also prevent the appearance of bias or a lack of objectivity (ECG, 2012[3]). For example, the Asian Development Bank (ADB) has specific procedures in place to protect against a range of conflicts of interest that could weaken the objectivity and integrity of its evaluations (ADB, 2008[11]). These include protections against official, professional, financial and familial conflicts.

3.3. Resourcing

Human and financial resources dedicated to evaluation vary significantly. Multilateral organisations often have larger centralised evaluation units, averaging 38 full-time equivalent (FTE) staff. Bilateral organisations have, on average, 10 FTE staff within the evaluation unit, though with significant variation between organisations. Of the 38 bilateral organisations that reported their staff numbers, only 9 (24 %) have 10 or more staff members. In two organisations, there are no staff fully dedicated exclusively to the evaluation function.

> "Use of resources is also an important attribute of independence – if you don't have enough resources, how can you really be independent?"
>
> United Nations Development Programme, Independent Evaluation Office

Some organisations use staff to conduct evaluations, while others rely on external evaluators, as described above. Larger evaluation units more often conduct evaluations directly, using external evaluators only when specific skills or knowledge are required. Conversely, there are also evaluation units that do not conduct evaluations themselves but rather commission external consultants, with staff playing a management and oversight role.

Gender balance in evaluation units is improving. Across participating organisations, 60 % of evaluation staff are women – a slight decrease from the 62 % in 2016. In leadership roles, widely viewed as a more significant indicator of gender parity, 25 out of 46 (54 %) evaluation units of reporting organisations are led by a woman, up from 46 % in 2016.

Accounting for financial resources dedicated to evaluation is challenging. As part of the data collection process for this study, participating organisations were asked to report on the total amount spent on development and humanitarian evaluations in 2021, in local currency[2] (Table 3.1). Annex B provides an overview of resources for each member. However, evaluation units' varied mandates and institutional set ups mean that not all organisations are able to report on evaluation costs in the same way. A calculation of evaluation spending may include staff costs, including professional and administrative staff; contractual services – most often consultant evaluators; regular operating costs, which are sometimes included as part of an organisation-wide budget and not calculated for a specific department or unit; travel costs; and communications, including costs for publications and knowledge sharing events.

Whilst a robust comparison over time is not possible because there is not yet an agreed methodology for counting these costs, **overall trends in spending from 2010 to 2015 and from 2015 to 2021 suggest modest increases, with an increase of 15% overall from 2010 to 2021**. During this

period, total ODA (which can provide a rough measure of the increasing portfolio of development finance evaluation units are mandated to evaluate) increased by 35%. From 2015 to 2023, multilateral spending increased more than for bilateral units (Table 3.1). Annexes B and C provide a more nuanced picture for each individual member country, with some members – like Australia – increasing their evaluation budgets relative to ODA provided, and others decreasing spending on evaluation.

Table 3.1. Financial resources for evaluation, 2010 to 2021

	Average 2010	Average 2015	Average 2021	% Change 2010-2021
All organisations	4.08 million EUR	4.10 million EUR	4.68 million EUR	+15%
Bilateral organisations	1.92 million EUR	2.76 million EUR	2.95 million EUR	+54%
Multilateral organisations	7.76 million EUR	9.96 million EUR	11.00 million EUR	+42%
Total ODA provided	129.81 billion EUR	156.19 billion EUR	175.62 billion EUR	+35%

Note: Data on financial resources were reported by 42 out of 51 participating organisations. It is important to note that these figures are approximations only, as there is not yet a common methodology for accounting for evaluation spending. Reported figures were converted to euros, as it is the most common currency used by members.

References

ADB (2008), *Review of the Independence and Effectiveness of the Operations Evaluation Department*, https://www.adb.org/documents/review-independence-and-effectiveness-operations-evaluation-department. [11]

DFAT (2021), *Disability inclusion in the DFAT Development Program: Good practice note*, https://www.dfat.gov.au/sites/default/files/disability-inclusive-development-guidance-note.pdf. [9]

ECG (2012), *Big Book on Evaluation Good Practice Standards*, Evaluation Cooperation Group, https://www.ecgnet.org/document/ecg-big-book-good-practice-standards. [3]

Government of Canada (2022), *Gender equality and empowerment measurement tool*, https://www.international.gc.ca/world-monde/funding-financement/introduction_gender_emt-outil_renforcement_epf.aspx?lang=eng. [8]

Ministry of Foreign Affairs of Finland (n.d.), *Practical tips for addressing cross-cutting objectives in evaluations*, https://um.fi/documents/384998/0/Practical_tips_Cross-cutting_objectives_Climate_change_environment.pdf/169bbeb0-47f0-59a0-1c90-0bd7cb3752a0?t=1648783016982. [5]

OECD (2023), *Applying a Human Rights and Gender Equality Lens to the OECD Evaluation Criteria*, Best Practices in Development Co-operation, OECD Publishing, Paris, https://doi.org/10.1787/9aaf2f98-en. [6]

OECD (2010), *Quality Standards for Development Evaluation*, OECD Publishing, https://doi.org/10.1787/9789264083905-en. [2]

OECD (1991), *DAC Principles for Evaluation of Development Assistance*, OECD Publishing, https://www.oecd.org/dac/evaluation/41029845.pdf. [1]

Picciotto, R. (2018), *The Logic of Evaluation Independence and Its Relevance to International Financial Institutions (IMF)*, https://ieo.imf.org/-/media/IEO/Files/Publications/Books/first-decade/first-decade-chapter-5.ashx. [10]

Treasury Board of Canada Secretariat (2019), *Integrating Gender-Based Analysis Plus into Evaluation: A Primer*, https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/evaluation-government-canada/gba-primer.html#H-04. [7]

UNEG (2016), *Norms and Standards for Evaluation*, United Nations Evaluation Group, http://www.unevaluation.org/document/detail/1914. [4]

Notes

[1] There has been strong demand from the members of the OECD-DAC Results Community for guidance to support partner countries to implement and monitor climate adaptation efforts. In response, the Results Community Secretariat is preparing guidance on how to design results frameworks for climate adaptation, as well as how to support partner country monitoring and evaluation efforts in this area. The guidance process kicked off with a dialogue in October 2022, co-hosted by the United Nations Framework Convention on Climate Change (UNFCC) Adaptation Committee.

[2] Reported figures were converted to euros, as it is the most common currency used by members.

4 Evaluation processes

This chapter focuses on the evaluation process itself, from decisions on which evaluations to undertake to how findings are disseminated. Data show that organisations are putting increased emphasis on the early stages of the evaluation process to ensure high-quality end products that respond to evolving organisational priorities. The chapter also explores how the COVID-19 pandemic prompted innovative ways of working. Finally, it analyses how organisations collaborate with partner countries on evaluations – an area where there is room for improvement.

4.1. Evaluation plans

Planning for evaluation is important to ensure evaluations reflect organisational priorities. An evaluation plan or work programme sets out the evaluations that will be conducted over a specific timeframe. These plans may also include supporting activities, such as training, capacity building, knowledge sharing and communications (OECD, 1991[1]). Our survey shows that 94 % of evaluation units prepare evaluation plans (46 out of 49 reporting organisations), and all but four reporting evaluation units publish their plans – providing useful transparency about their work.

Several members report adjusting their ways of working to ensure that the evaluation topics chosen are aligned with organisational priorities. For example, Norway and Germany both hold consultative meetings with senior management and staff in their respective ministries to identify topics of interest. In line with the findings of previous studies in this series, evaluation plans generally reflect the organisations' strategic plans and priorities (OECD, 2010[2]; 2016[3]). The following issues were most frequently cited by survey respondents as guiding evaluation decisions: (i) issues of thematic interest; (ii) relevance of an initiative or area to the overall portfolio of an organisation; (iii) the level of ODA committed to an initiative or area; (iv) timeliness related to ongoing planning or review processes; (v) gaps in evidence; (vi) evaluation resources and capacity available; and (vii) possible collaboration with partners.

Using a consultative approach to design evaluation plans ensures they are useful. Many interviewees report that in designing evaluation plans, consultations are organised with management, programme and operational staff and advisory boards, as applicable. The aim is to identify relevant evaluations that will provide timely evidence on priority areas. This approach also helps to strengthen buy-in across the organisation, which will help ensure the findings are used. This is a shift from past practice, where the emphasis was on safeguarding evaluation plans from outside influence and programme staff were often kept at arm's length to avoid bias in the selection of evaluation topics. Today, many EvalNet participants report taking a more consultative approach, further demonstrating the increased emphasis on using and learning from evaluations.

Based on responses, evaluation programming and planning seems to be an area where EvalNet participants' practices have matured and are generally functioning well. However, the data do not allow us to draw conclusions on how successful they are overall in addressing the most relevant priorities (either for their own organisations or collectively for the DAC). A more comprehensive approach to develop and implement ongoing learning agendas across development partners for particular topics – such as one sustainable development goal – is one useful way to ensure evidence is generated systematically to address decision making priorities.

The time period covered by evaluation plans has increased since 2016, reflecting the greater focus on strategic evaluations integrated with organisational priorities. In 2016, the majority of evaluation work plans covered a one to two-year timeframe. This has increased to an average of three years in 2021 (Figure 4.1).

Figure 4.1. Duration of evaluation work plans, 2016 vs 2021
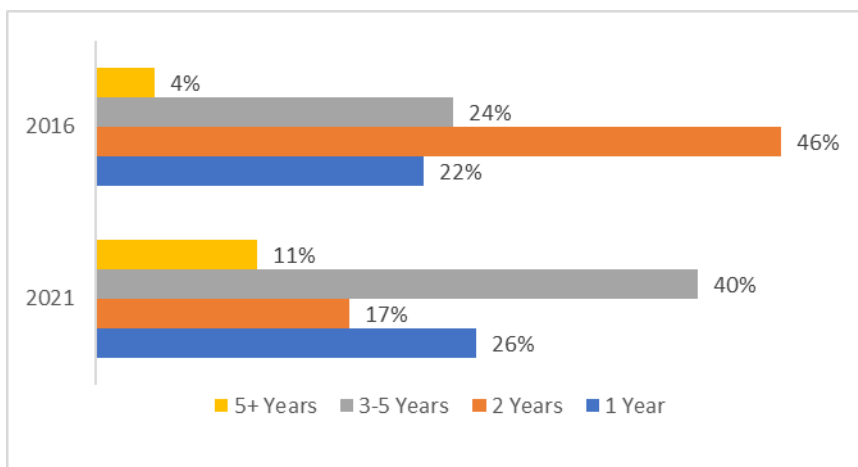
Note: Data on evaluation plans were reported by 44 out of 51 participating organisations.

4.2. Evaluation design

Each evaluation requires a focused design process to ensure it will produce robust and credible findings and meet organisational needs. There is no one-size-fits-all approach to evaluation. Rather, evaluation units draw on various approaches and methodologies depending on the specific context and objectives of an evaluation. As noted in Chapter 3, global and organisation-specific guidance supports evaluators in designing evaluations that are fit for purpose. This guidance is complemented by evaluability assessments (Box 4.1 gives an example from Australia), which are a key tool for enhancing the quality, feasibility and cost-effectiveness of evaluations.

Strong terms of reference (ToRs) are a key component of high-quality evaluations. Dialogue with the intended users, such as while developing the ToR, can also ensure evaluations are relevant and answer priority questions. During the survey and in EvalNet discussions, participating organisations highlighted that high-quality ToRs help set the stage for robust evaluations. These ToRs often include the evaluation objectives, approach/methodology, evaluation questions, quality assurance mechanisms, and roles and responsibilities. As with the evaluation planning process, increased effort is being made to consult relevant staff in the preparation of ToRs, seeking their input on evaluation questions that will provide the most useful evidence.

> **Box 4.1. Evaluability assessments in Austrian development co-operation**
>
> In 2022, Austria developed a Guidance Document on Evaluability Assessments to support the planning, implementation, and use of evaluability assessments in Austrian development co-operation. The document provides practical guidance and tools for assessing evaluability along four dimensions: (i) the quality of the intervention design; (ii) the availability of relevant data; (iii) the perceived usefulness of the evaluation by relevant stakeholders; and (iv) its feasibility in the temporal, geographical and institutional context. It also describes practical aspects of evaluability assessment and outlines individual steps in the implementation process. The guidance can be applied for assessing the evaluability of a range of interventions – strategies, policies, programmes, projects – and at different points of time in an intervention cycle: during the design phase, immediately after approval, prior to commissioning an evaluation, or as the first step of an evaluation.
>
> Source: Austrian Development Agency (2022[4]), Evaluability Assessments in Austrian development cooperation: Guidance Document, https://www.entwicklung.at/fileadmin/user_upload/Dokumente/Evaluierung/GL_for_Evaluability_Assessments.pdf.

4.3. Conducting evaluations

Challenges cited in conducting evaluations include lack of resourcing, weak staff capacity and technical capacity of evaluators, lack of availability of data, and low support from senior management for a learning culture. While larger organisations reported fewer issues around resourcing and capacity, the quality of external evaluators and lack of data were common challenges raised across evaluation units.

Smaller bilateral organisations face the most challenges with external evaluators. As noted in Chapter 3, multilateral organisations often have larger centralised evaluation units than bilateral organisations. Organisations with fewer staff, understandably, rely more on external consultants in conducting evaluations. However, it can be challenging to find consultants with the right combination of relevant expertise to conduct high-quality evaluations.

Competitive procurement processes help select good external evaluators, but are not always sufficient to ensure high quality work. Efforts are made to set strong selection criteria, in line with well developed ToRs, in order to find consultants who will be able to deliver on the objectives of the evaluation. Often, a multi-stakeholder reference group is also established to support the evaluation and to select a qualified team. It is usually chaired by an official from the organisation but who is not directly engaged in the project to be evaluated.

The quality of evaluations is linked to strong results systems. Evaluation units regularly highlighted the lack of reliable and timely data as a major challenge in conducting evaluations. Results systems set the stage for better monitoring and data collection, which in turn allows for better evaluations (UNDG, 2011[5]). If results systems do not support the collection of necessary data, evaluations must gather more primary data or rely on external sources to manage data gaps, both of which require additional time and resources.

Evaluation units increasingly engage results teams. Participants report that given the importance of results data to inform evaluation, they are deepening collaboration with their organisation's results teams. As described by interviewees, this often involves senior evaluation officials providing insights on results strategies and frameworks, as well as participating in internal working groups dedicated to strengthening results measurement systems and tools.

4.4. Evaluation in the context of COVID-19

The COVID-19 pandemic had a significant impact on how evaluations were conducted. On-site data collection and in-person engagement with stakeholders were no longer possible in many contexts. It was particularly difficult to work with local partners who did not have the capacity or resources to engage virtually. Because of this, many evaluations were interrupted, postponed, or cancelled.

New ways of working were explored to ensure the critical evaluation function was fulfilled during the crisis. This included working more with local evaluators, virtual data collection, secondary data analysis, and increased collaboration with partners to overcome challenges. Such ways of working were being tested before the onset of the COVID-19 pandemic, but their use was accelerated because of the crisis. During EvalNet meetings, most members report that they will continue to use virtual engagement and data collection methods.

> "The pandemic has complicated the collection of data, but the situation has also provided an opportunity for valuable testing and the development of methodological approaches to data collection, including the use of machine learning and satellite data."
>
> Norway, Department for Evaluation

More than three-quarters of participating organisations relied on local consultants to conduct evaluations during the pandemic. With their staff unable to travel, evaluation units engaged local evaluators to support in-country data collection and stakeholder engagement. Local evaluation organisations – such as the Mongolian Evaluation Association – reported increased demand and growing momentum in their national evaluation capacity development due to the travel restrictions. However, many organisations reported challenges in identifying evaluators with the right expertise to deliver high-quality work. Some also raised concerns that virtual methods and reliance on local partners could reduce the quality of evaluations. As noted, identifying suitable, qualified consultants – whether or local or international – is a widespread challenge for evaluation units.

> "We – the Mongolia Evaluation Association – are really a product of COVID, I would say, because before COVID in Mongolia, it was mostly international evaluators flying in, and we acted only as local data collectors or working in the shadows. However due to flight restrictions, they were not able to come, and we took on a bigger role."
>
> National partner from the Mongolian Evaluation Association

Virtual engagement and data collection were key for conducting evaluations during the pandemic. Remote assessment guidelines, along with training for evaluators, helped support the switch to online interviews and virtual data collection. In addition to ensuring evaluations could continue despite the pandemic, in some cases this also helped to gain broader participation in evaluation processes. For example, the ADB's evaluation team working on its real-time pandemic response evaluation was able to

interview more than two dozen ministers of finance (virtually), many more than would normally be the case. Participating organisations noted, however, that there is still value in conducting in-person evaluations, and that the quality of engagement is better in-person. In short, virtual interviews have an advantage of quantity over quality.

Big data was used to supplement evaluation-specific data collection in the constrained COVID-19 context. Big data refers to data characterised by high volume, velocity, and variety (OECD, 2021[6]). While virtual data collection was used to support many evaluations during the pandemic, participating organisations also reported the increased use of big data to complement and validate findings. Examples of big data sources used include country-owned administrative data, satellite imagery, and data on weather patterns and other characteristics of the natural environment (Box 4.2).

The COVID-19 pandemic also highlighted the critical importance of collecting real-time evidence to strengthen crisis response. During the pandemic there was an urgent need for evidence on which response and recovery measures were working and which did not, in order to scale up successful approaches. Evaluation units therefore shifted their focus from lengthier ex-post evaluation processes to generating early evidence and producing brief knowledge products to guide ongoing efforts. There was also a huge increase in rapid syntheses early in the pandemic, drawing lessons from past evaluative work to quickly inform the ongoing response, as discussed during EvalNet's June 2020 meeting.

Participation in the COVID-19 Global Evaluation Coalition supported more co-ordinated efforts and sharing of data and evidence. The COVID-19 Global Evaluation Coalition was established by EvalNet members, partner countries and members of the UN Evaluation Group (UNEG) and the Evaluation Co-operation Group (ECG) of the multilateral development banks in 2020 to bring stakeholders together to share lessons and ensure that the global humanitarian and development communities are able to deliver results, even during times of crisis (COVID-19 Global Evaluation Coalition, 2020[7]). The COVID-19 Global Evaluation Coalition developed a new way of working[1] to produce five rapid "Lessons from Evaluation" briefs (2023[8]), which were submitted to meetings of development ministers.

Box 4.2. Innovations in evaluation sparked by the COVID-19 pandemic

Satellite imagery in Kenya

The African Development Bank's Independent Development Evaluation unit adapted its data collection techniques during the COVID-19 pandemic. For example, during the evaluation of the Last Mile Electricity Connectivity Project (LMCP) in Kenya in 2021, it drew on satellite images to assess progress.

Artificial intelligence for development analytics

UNDP's Independent Evaluation Office, in partnership with the United Nations International Commuting Centre and Amazon Web Services, developed a cloud-based tool to review and extract key information from evaluation reports and other documents. The Artificial Intelligence for Development Analytics (AIDA) tool uses artificial intelligence and machine learning to search for specific information within more than 6 000 evaluation reports. AIDA identifies lessons from evaluations and displays them on a dashboard organised by country, sector, theme, modality, and timeframe.

Crisis response repositories

ADB's Independent Evaluation Department launched an online repository of evaluation evidence related to crisis response and recovery, ensuring information on challenges and solutions were readily available to inform COVID-19 response. Similarly, the European Commission developed the EvalCrisis web platform, which hosts a collection of resources on conducting evaluations in times of crisis. The platform includes guidance documents, training webinars, blogs and podcasts.

Joint evaluations

Working through the COVID-19 Global Evaluation Coalition, the United Nations High Commissioner for Refugees (UNHCR), the Governments of Colombia, Finland and Uganda, and the Active Learning Network for Accountability and Performance (ALNAP), conducted a joint evaluation focused on how host states, development organisations, and non-state actors ensured the protection of refugee rights during the pandemic.

Real-time evaluation

The Islamic Development Bank's Independent Evaluation Department undertook a real-time evaluation of the Bank's COVID-19 Strategic Preparedness and Response Program. The evaluation used an amended version of the OECD evaluation criteria involving five different lenses: relevance, coherence, agility and responsiveness, effectiveness, and efficiency. The data collection methodologies used included virtual interviews, remote focus discussions, and online surveys. The evaluation provided timely feedback that guided operational improvements through mid-course adjustments as the programme moved into the next phase of the crisis response. It used evidence and drew lessons gleaned from the experience to identify opportunities for improvement, presented lessons learned, and offered guidance on the way forward.

Source: ADB (2020[9]), *COVID-19 Response Resources*, https://www.adb.org/site/evaluation/evaluations/covid-19-response; IEO, World Bank (2020[10]), *Evaluation in the Age of Big Data: Prospects and Challenges for Independent Evaluation*, https://ieg.worldbankgroup.org/event/datascience-and-evaluation.Taylor, G. et al. (2022[11]), *Joint Evaluation of the Protection of Rights of Refugees during the COVID-19 Pandemic*, https://www.covid19-evaluation-coalition.org/documents/Final%20Report%20%20Refugees%20Rights%20During%20The%20Covid%2019%20.pdf.

4.5. Quality assurance

A focus on quality assurance is apparent throughout the full evaluation process. Quality assurance is not only a final step in the evaluation process. Evaluation units systematically integrate quality assurance checks at each stage of the evaluation process, including defining objectives, deciding methodology and approach, designing ToRs, procuring consultants and drafting and reviewing inception and final reports (Box 4.3). Several participants report recently conducting reviews of evaluation quality, including France, Japan, Norway, Sweden and ADB.

Independent entities are used to assess the quality of evaluation-related documents and products. All but eight members and observers have some kind of independent oversight entity, such as advisory bodies or steering committees, to provide regular quality checks. These entities often include representatives of relevant policy or programme teams, subject-matter specialists and external advisors with expertise in development evaluations. Independent entities review and provide input on context, methodological approach, data accuracy and analysis, overall findings and recommendations.

Each evaluation unit has its own quality procedures in place, and all but a handful require an independent quality assessment before approval and publication of evaluation reports. These quality assessments use the *Quality Standards for Development Evaluation* (OECD, 2010[12]) to determine the quality and credibility of an evaluation.

> **Box 4.3. Quality assurance in Sweden's evaluations**
>
> Sweden's Expert Group for Aid Studies' (EBA) quality assurance process rests on widespread recognition of the properties required of a good evaluation, including utility, accuracy, and feasibility. However, the EBA recognises that it isn't enough for an evaluation to only possess these qualities – it must also use probing questions to bring new knowledge to the field.
>
> EBA therefore places an emphasis on factual questions when assessing quality and aims to avoid methodological reductionism: the questions come first, and the choice of scientific method comes second. The quality assurance process must be designed such that the properties that constitute quality can be tailored and adapted to different contexts. The quality assurance process is guided by the Expert Group, with the process involving a number of factors to ensure the quality of an EBA report at each stage (Table 4.1). The focus is on the elements that the Expert Group and the secretariat are able to influence.
>
> **Table 4.1. EBA's quality assurance process**
>
> | 1. Assessments of and decisions on study proposals | 2. The process of reference group meetings | 3. Expert Group's decision on publication |
> |---|---|---|
> | • Processes for evaluating proposals received
• Assessment of the author's skills and experience
• The secretariat and the working groups' early dialogue with the authors and target groups
• The secretariat's documentation for the Expert Group | • Guidelines for the reference group
• Composition and participation
• The Chair and the secretariat's contribution, input, intervention
• Dynamics and working method
• Communication between reference group and authors
• Reading and input from the programme manager | • Quality of final review statements from reference group members
• Summary from the Chair of the reference group to the Expert Group
• Reading by members
• Dynamics and dialogue in the Expert Group
• Communication from the Expert Group to the Chair of the reference group and the secretariat |

4.6. Working jointly

Fewer joint evaluations have been undertaken since 2010. Joint evaluations increase efficiency, improve credibility, reduce the burden on partner country governments and implementing partners, and create more opportunities for mutual learning (OECD, 2006[13]). Despite these well-known benefits, the share of joint evaluations undertaken by participating organisations has declined significantly since 2010, when nearly a quarter of all evaluations were joint (Figure 4.2).

Of the 2 321 evaluations undertaken by participating organisations in 2021, only 77 were conducted jointly, a slight increase from 2016. UNDP conducted the most (63 joint evaluations), followed by Australia's DFAT (5), France's MEAE (3), Norway's Norad (2), Canada's GAC (1), Germany's BMZ/DEval (1), Sweden's Sida (1) and Switzerland's SDC (1). The remaining 36 organisations (out of 44 reporting organisations) reported that no joint evaluations took place in 2021. Meanwhile, partners in the United Nations Evaluation Group report that they are seeing a huge increase in the proportion of joint work in their portfolios, reflecting in part the ONE UN reforms (DAC, 2021[14]).

Figure 4.2. Share of joint evaluations undertaken by participating organisations, 2010 to 2021

Note: Data on joint evaluations were reported by 44 out of 51 participating organisations.

The general move away from joint evaluations and sharing evaluation plans is often attributed to limited staff resources. Collaborating across organisations, including participating in global initiatives such as EvalNet, requires an investment of time and effort. Participating organisations have noted that it is too challenging to engage in this way in the context of limited human resource capacity. This reflects broader trends in the development space, with many co-operation providers facing the need to respond to national priorities and a reprioritisation of the aid effectiveness agenda.

4.7. Partner country engagement

Partner country engagement in centralised evaluations remains low, despite commitments in this area. Partner country engagement in development co-operation, including in development evaluations, is important for ensuring local ownership of development activities and fostering mutual accountability. Partner country engagement in evaluations is also a way to strengthen local evaluation capacity (OECD, 2010[12]). The survey finds that partner country engagement is particularly low when preparing evaluation work plans, including deciding which evaluations to conduct within a specific timeframe (Figure 4.3). This may reflect the increasing attention paid to ensuring evaluations are aligned with organisational or domestic priorities.

Partner countries play a key role in stakeholder engagement and data collection. Over three-quarters of participating organisations (34 out of 44) reported that partner countries were always or regularly engaged in country-level data collection efforts, although the degree of engagement varies from active participation to facilitating connections.

Figure 4.3. Partner country engagement in evaluations

Note: Data on partner country engagement were reported by 41 out of 51 participating organisations.

4.8. Capacity strengthening

Evaluation units signal strong commitment to improving evaluation capacity across organisations and in partner countries. Given the importance of evaluation to strengthen accountability for results and inform future development efforts, significant emphasis is placed on strengthening the capacity of evaluation units, as well as of programme teams and partner country stakeholders. This support often draws on the standards, guidance and tools provided through EvalNet (OECD, 2022[15]).

Many reporting centralised evaluation units engage in capacity strengthening and quality assurance for decentralised evaluations. Centralised evaluation units are often tasked with providing technical support to staff working on decentralised evaluations, who are more often subject-matter experts than trained evaluators. This technical support includes training, advisory support during evaluation design and implementation, and providing quality assurance checks. Internal capacity strengthening often consists of training academies and online courses.

EvalNet participants are providing increased capacity support to partner countries through training programmes and knowledge exchange events (Box 4.4). Approximately three-quarters of participating organisations reported that they provide evaluation capacity building to partner country actors, including governments, evaluation societies, civil society, and individual evaluators. This is an increase from 2016, when half of respondents indicated that they provide support to partner country stakeholders.

Box 4.4. Member examples of building evaluation capacity

Training and courses

The Asian Development Bank (ADB), along with the Asia-Pacific Finance and Development Institute, finances the Shanghai International Program for Development Evaluation Training which trains government officials from ADB member countries to manage or undertake evaluations of projects funded by international finance institutions.

The Special Evaluation Office (SEO) within Belgium's FPS Foreign Affairs is a leader in providing training to country-level partners. In partnership with the EGMONT Institute, the SEO organises evaluation training within partner countries. The SEO also funds an annual two-week intensive evaluation course, organised by the University of Antwerp, that brings together evaluation professionals from partner countries.

EvalNet members, including Germany, Switzerland and the Netherlands, support the International Program for Development Evaluation Training (IPDET). IPDET provides training to professionals in developing countries who commission, manage, practise or use evaluations.

Knowledge exchange

The African Development Bank (AfDB) is part of *Twende Mbele*, in partnership with CLEAR Anglophone Africa and six African governments (Benin, Ghana, Kenya, Niger, South Africa and Uganda). *Twende Mbele* is a network that facilitates collaboration in developing and implementing monitoring and evaluation systems.

Japan's Ministry of Foreign Affairs regularly hosts ODA Evaluation Workshops which bring together development partners and partner countries from the Asia-Pacific region for knowledge exchange and peer learning on development evaluation.

Source: SHIPDET (2022[16]), About SHIPDET, https://afdi.snai.edu/SAG_html.aspx?type=SHIPDET&h=5&l=English; University of Antwerp (2021[17]), *Strengthening National Monitoring and Evaluation Capacities*, https://www.uantwerpen.be/en/study/programmes/all-programmes/monitoring-and-evaluation-capacities/.

EvalNet members played a key role in establishing the Global Evaluation Initiative (GEI). GEI was established in 2020 to respond to high partner country demand for support for strengthening monitoring and evaluation systems. At the time, many EvalNet members and observer organisations, as well as regional and global evaluation initiatives, were providing individual support to their partner countries in this area. The GEI was established by the UNDP and the World Bank to consolidate these efforts, pooling resources for evaluation capacity building and leveraging the comparative advantages of individual organisations (UNDP, 2016[18]). GEI supports partner country governments in their efforts to improve national evaluation capacity, providing diagnostics, training, technical assistance and other support to local evaluation entities and professionals (GEI, 2022[19]). Today nearly half of participating organisations report that most of their evaluation capacity building in partner countries is channelled through GEI.

References

ADB (2020), *COVID-19 Response Resources*, https://www.adb.org/site/evaluation/evaluations/covid-19-response (accessed on 24 August 2022).

[9]

Austrian Development Agency (2022), *Evaluability Assessments in Austrian development cooperation: Guidance Document*, https://www.entwicklung.at/fileadmin/user_upload/Dokumente/Evaluierung/GL_for_Evaluability_Assessments.pdf.

[4]

COVID-19 Global Evaluation Coalition (2023), *Lessons from Evaluation*, https://www.covid19-evaluation-coalition.org/lessons-from-evaluation/ (accessed on 2023).

[8]

COVID-19 Global Evaluation Coalition (2020), *COVID-19 Global Evaluation Coalition*, https://www.covid19-evaluation-coalition.org/about/ (accessed on 24 August 2022).

[7]

DAC (2021), *Summary Record of the 25th Meeting of the DAC Network on Development Evaluation*, Development Advisory Committee, OECD, Paris, https://one.oecd.org/document/DCD/DAC/EV/M(2020)1/En/pdf.

[14]

GEI (2022), *Global Evaluation Initiative*, https://www.globalevaluationinitiative.org/who-we-are (accessed on 30 August 2022).

[19]

IEO, World Bank (2020), *Evaluation in the Age of Big Data: Prospects and Challenges for Independent Evaluation*, https://ieg.worldbankgroup.org/event/datascience-and-evaluation (accessed on 24 August 2022).

[10]

OECD (2022), *Renewal of the Mandate of Network on Development Evaluation (EvalNet)*, OECD, https://one.oecd.org/official-document/DCD/DAC(2022)12/REV1/en.

[15]

OECD (2021), *Development Co-operation Report 2021: Shaping a Just Digital Transformation*, OECD Publishing, Paris, https://doi.org/10.1787/ce08832f-en.

[6]

OECD (2016), *Evaluation Systems in Development Co-operation: 2016 Review*, OECD Publishing, Paris, https://doi.org/10.1787/9789264262065-en.

[3]

OECD (2010), *Evaluation in Development Agencies*, Better Aid, OECD Publishing, Paris, https://doi.org/10.1787/9789264094857-en.

[2]

OECD (2010), *Quality Standards for Development Evaluation*, OECD Publishing, https://doi.org/10.1787/9789264083905-en.

[12]

OECD (2006), "Guidance for Managing Joint Evaluations", *OECD Papers*, https://doi.org/10.1787/oecd_papers-v6-art6-en.

[13]

OECD (1991), *DAC Principles for Evaluation of Development Assistance*, OECD Publishing, https://www.oecd.org/dac/evaluation/41029845.pdf.

[1]

SHIPDET (2022), *SHIPDET: About*, SHIPDET, https://afdi.snai.edu/SAG_html.aspx?type=SHIPDET&h=5&l=English (accessed on 30 August 2022).

[16]

Taylor, G. et al. (2022), *Joint Evaluation of the Protection of Rights of Refugees during the COVID-19 pandemic*, UNHCR, Geneva, https://www.covid19-evaluation-coalition.org/documents/Final%20Report%20-%20Refugees%20Rights%20During%20The%20Covid%2019%20.pdf (accessed on 25 August 2022).   [11]

UNDG (2011), *Results-based Management Handbook: Harmonizing RBM concepts and approaches for improved development results at country level*, United Nations Development Group, https://unsdg.un.org/sites/default/files/UNDG-RBM-Handbook-2012.pdf.   [5]

UNDP (2016), *Evaluation and Independence - Existing evaluation policies and new approaches*, United Nations Development Programme, New York, http://web.undp.org/evaluation/evaluations/documents/Independence_of_Evaluation.pdf.   [18]

University of Antwerp (2021), *Strengthening National Monitoring and Evaluation Capacities*, https://www.uantwerpen.be/en/study/programmes/all-programmes/monitoring-and-evaluation-capacities/ (accessed on 30 August 2022).   [17]

Notes

[1] A first draft was produced by one member, then validated by other units with relevant work to add supporting evidence and triangulate findings. It was then published as a joint product.

5 Using evaluation evidence

Ensuring that evaluation findings are used has been a consistent challenge, which is reflected in the increased focus on usefulness as a guiding principle of evaluation. To increase use, organisations are engaging across teams to ensure buy-in; aligning with policy, programming and project cycles; and preparing more diverse products to share evaluation evidence in a way that responds to the needs of target audiences and facilitates action.

5.1. Using evaluation findings

Ensuring that evaluation findings are used has been a persistent challenge for evaluation units. As noted in Chapter 2, accountability and learning are the dual objectives of development evaluation. Reporting suggests that while stakeholders are satisfied with the extent to which evaluations meet accountability objectives, it is more challenging to make sure evaluation findings are used to inform future policy, planning, and resourcing decisions to enhance effectiveness.

Use of findings has been viewed largely as an ex-post activity. In the past, participating organisations have prioritised accountability aspects of evaluations, focusing on the success or failure of policies, programmes, and projects. Given this prioritisation, promoting the use of findings has been left until after evaluations are complete, rather than considering learning objectives in the early stages of the evaluation process (Serrat, 2009[1]; Cracknell, 2000[2]).

> "There is learning in accountability and accountability in learning."
>
> Netherlands, Policy and Operations Evaluation Department

The nature of challenges in using evaluation findings has changed. The 2010 study reported that DAC Peer Reviews and dialogue at the DAC Network on Development Evaluation (EvalNet) meetings often cited accessibility and use of evidence as the primary challenges for evaluation processes, often related to the limited ability to share findings widely (OECD, 2010[3]). These accessibility challenges have been addressed, with the majority of participating organisations publishing evaluation reports and summaries online and disseminating them through various platforms where findings can be discussed. However, new challenges have emerged:

- **Tension between providing timely evidence while ensuring quality.** Timely evaluation findings are essential if they are to be useful. Participating organisations have to juggle providing quick evidence that can inform ongoing efforts with ensuring high-quality and credible findings. This challenge came to the fore during the COVID-19 pandemic, when organisations were working in new ways and in a new context and when real-time evidence was needed on which response and recovery measures were working (see Section 4.4).

- **Absorption capacity affects the ability of target audiences to act on evaluation findings.** Unlike in the past, there is now an overwhelming wealth of data and evidence available to policy makers and programme/project managers, including internal evaluation findings from EvalNet and beyond. In a resource-constrained development environment and with the high volume of information produced, development professionals are not able to fully review and incorporate all evaluation findings.

Evaluation units are making concerted efforts to address these uptake-related challenges. Key measures include:

- **Putting renewed focus on identifying learning objectives early in the evaluation cycle and increasing emphasis on how evaluation findings are presented and disseminated.** This includes consulting on evaluation needs, drafting Terms of Reference (ToRs) that place equal emphasis on learning objectives, highlighting the areas in which findings can be used to inform decision making, and considering how findings will be presented and disseminated.

"If evaluation is to fulfil its role as a decision-making tool, we need to improve ownership of its conclusions."

France, Agence Française de Développement, Department for Evaluation and Learning

- **Aligning evaluations with overarching organisational objectives and ensuring timely release of findings.** In practice, this means choosing to conduct evaluations in time for their findings to inform upcoming policy decisions. For example, AFDB's 2019 evaluation of its performance on gender mainstreaming was timed to feed into international negotiations on gender equality and women's empowerment objectives (IDEV, 2020[4]). And Denmark has assessed its overall portfolio of climate finance commitments (DANIDA, 2021[5]). In the same vein, evaluations can be undertaken in line with programme and project cycles. When it is not possible to finalise evaluation findings in time for a specific deadline, evaluation units increasingly share preliminary findings with relevant teams to provide timely inputs into key decisions.

- **Engaging early with staff across the organisation to increase relevance and ownership.** About half of the reporting units systematically engage with programme staff, including colleagues working in thematic areas, during the evaluation planning stages to identify relevant evaluation topics. Not only does this help to ensure the utility of evaluation findings, it also ensures broad organisational buy-in to evaluation processes, promotes ownership and encourages the use of evaluation results.

- **Conducting more cross cutting analyses and syntheses.** For example, AFD's 2023 Evaluations Report aims to raise awareness of its evaluation work among a wider audience, and answer crucial questions around key thematic areas: essential services in Africa, climate and biodiversity, and access to education and training (AFD, 2023[6]).

Figure 5.1 lists the key elements for supporting the use of findings based on information reported by participants.

Figure 5.1. Evaluation utility checklist

- [ ] Align to organisation strategic priorities
- [ ] Engage with senior management and subject-matter experts
- [ ] Design ToRs that incorporate learning objectives
- [ ] Prepare a communication strategy during the evaluation planning phase
- [ ] Share findings with key stakeholders on an ongoing basis
- [ ] Formulate concrete and actionable recommendations
- [ ] Draft reports in non-technical language
- [ ] Use a variety of presentation and dissemination methods

Note: Based on challenges in using evaluation findings and the mitigation efforts participating organisations are using to address them.

How findings are presented and disseminated influences their use. The 2016 study highlighted the need for more focus on how evaluation findings are prepared and shared, and suggested ensuring that the design of knowledge products is user-centric (OECD, 2016[7]). Participating organisations have taken this recommendation forward, exploring new ways to present evaluation findings alongside publishing full evaluation reports. They are also looking at new ways to share information that meet the varied needs of different audiences, ensuring key points are clear and easily absorbed. Some organisations have gone so far as to employ staff dedicated to knowledge sharing and communication in the evaluation unit.

Growing diversity in evaluation knowledge products reflects an effort to respond to the needs of diverse audiences. In line with past commitments to transparency and accountability, nearly all participating organisations – 92 % (45 out of 49) – post evaluation reports on publicly-accessible websites. Those that do not post full evaluation reports at least post executive summaries. However, increasingly, participating organisations complement full length reports with other messaging, in an effort to better connect with internal and external audiences and facilitate use of results. Understanding the differing interests and needs of different audiences, evaluation units routinely publish executive summaries, press releases, policy briefs, infographics, data visualisations, blogs, videos, and podcasts to present evaluation findings (Figure 5.2). Compared to previous studies, we see more members (now the vast majority) diversifying their knowledge products.

Figure 5.2. Evaluation knowledge products
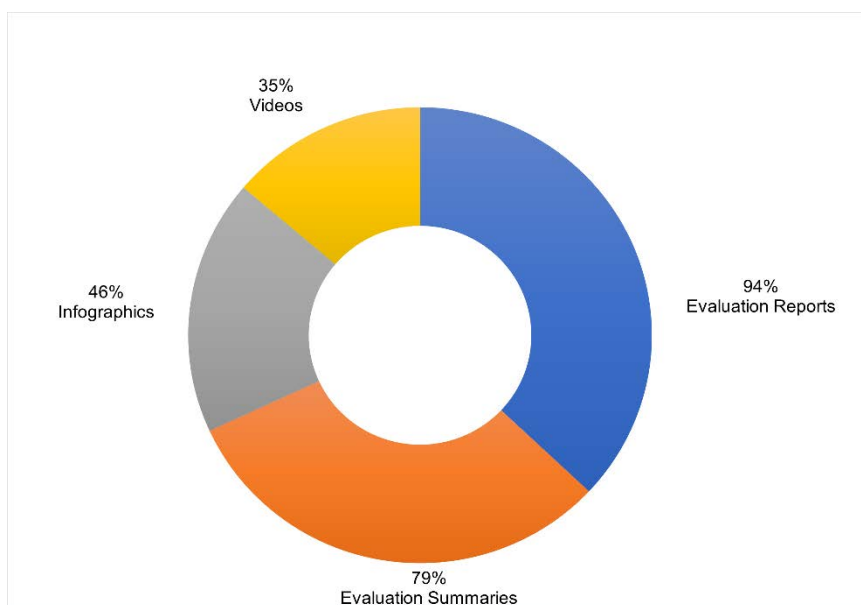
Note: Data on dissemination methods were reported by 48 out of 51 participating organisations.

Dissemination to internal audiences often extends beyond simple sharing of knowledge products to an increased focus on interactive learning. While evaluation units continue to share evaluation reports and other knowledge products across their organisations, there is more emphasis on interactive methods of engagement (Box 5.1). For example, 64 % (30 out of 47) of reporting organisations noted that they plan internal events (e.g., launch events, webinars, brown bag lunches etc.), involving evaluators and thematic units to discuss findings, recommendations and possible ways forward.

Social media is widely used for communicating evaluation findings to external audiences. In 2016, 51 % of participating organisations mentioned using digital platforms for sharing evaluation findings. In

2021, 73 % (29 out of 40) report regularly using social media to promote knowledge products and highlight key findings.

Box 5.1. Inventive ways to strengthen learning from evaluations

Austria's annual learning conference

The Austrian Development Agency (ADA) organises a yearly workshop that serves as a retreat for its staff and that is attended by the Director General for Development Cooperation within the Ministry of Foreign Affairs. The objective of this event is to facilitate mutual reflection, knowledge exchange, and learning on a variety of questions related to global development co-operation.

A dedicated learning function in Canada

Global Affairs Canada's (GAC) evaluation unit is made up of two divisions: the Evaluation Division and the Evaluation Services and Learning Division. The latter is primarily focused on learning from evaluation. Its role involves conducting complementary research and translating evaluation findings into actionable knowledge that informs decision making.

France's participatory approach

Within France's Ministry of European and Foreign Affairs (MEAE), the directorate responsible for international development co-operation organises a series of presentations and workshops with senior-level officials and relevant departments to discuss findings and provide support in responding to evaluation recommendations. This consultative approach helps to ensure buy-in to recommendations, thereby increasing their implementation.

5.2. Management response systems

Management response ensures that evaluations contribute to organisational accountability and learning. It serves as a key tool in ensuring the use of evaluation findings and establishing crucial oversight over organisational efforts. It is considered good practice for all stakeholders targeted by an evaluation, and therefore responsible for implementing evaluation recommendations, to be involved in responding to the findings and planning follow-up actions.

Nearly all organisations have a management response process/system in place. Of the participating organisations, 93 % (43 out of 46) reported that a management response to an evaluation is an absolute requirement. Generally, a management response process consists of: (i) a formal position from management that states whether a recommendation is accepted, partially accepted or rejected; and (ii) an action plan that outlines specific measures to be taken to implement recommendations, including clear roles and responsibilities and concrete timelines.

Management responses are increasingly tracked, highlighting their importance. The existence of a management response does not ensure the use of evidence for decision making; members are increasingly developing new elements to track, and support follow up. In 2010, 54 % (20 out of 37) of participating organisations reported having a system in place to track implementation of management response actions. In 2016, this figure increased to 78 % (36 out of 46). This has jumped again, with 91 % (42 out of 46) of participating organisations in 2021 indicating that they have a system in place to track implementation of management response action plans.

Several units are also pursuing ongoing research to investigate the quality of both their recommendations and the resulting actions, towards the overall goal of improving effectiveness. For example, the ADB's

Independent Evaluation Department's 2020 assessment indicates that only 75% of actions in response to recommendations were either fully or largely implemented (ADB, 2020[8]). They report that shortfalls in implementation are generally due to a lack of alignment between actions and recommendations, and that management actions are not specific about intended outcomes—both of which can be improved through better engagement between management and IED at the action plan formulation stage.

> "The impact of the Independent Evaluation Department's (IED) recommendations depends on the quality of management's action plans."
>
> Asian Development Bank, Independent Evaluation Department

Most often, it is the evaluation unit that is responsible for tracking whether management implements actions in response to evaluation recommendations. In 91 % of cases (39 out of 43), the evaluation unit plays a role in management response tracking and follow-up (Box 5.2). In a small number of cases, the relevant policy or programme unit or organisation management is responsible for tracking. There is growing transparency in tracking of management responses, including the UK's Independent Commission for Aid Impact (ICAI) and the Netherlands' Policy and Operations Evaluation Department (IOB), which report to their respective parliaments and require formal responses from government to their evaluations. Publishing these government responses and reporting on implementation progress support public accountability.

Periodic monitoring reporting is also used to assess implementation progress. Some participating organisations track how the target audience responds to the evaluation findings and recommendations, and reports to management or independent oversight committees on follow up actions taken. These progress reports ensure accountability, while also providing a space for dialogue and learning.

> "The goal of learning requires cooperation between those initiating the evaluations and the intended users."
>
> Norway, Department for Evaluation

Box 5.2. Tracking implementation of management responses

The Management Action Record

The Management Action Record is an online system used by many agencies (such as the World Bank Group's IEG, the African Development Bank's IDEV, the Asian Development Bank's IED, etc.) to support the implementation of management response action plans. The unit responsible for tracking is able to insert all the agreed actions and tag the unit responsible for implementation. As progress is made, the system is updated, providing up-to-date information that is accessible to all relevant staff. This tool facilities periodic follow up on the adoption of the evaluation agency's recommendation by the organisation.

Canada's annual monitoring report

Canada's evaluation function monitors and tracks the implementation status of all active management responses and action plans on an annual basis, and reports back to the Performance Measurement and Evaluation Committee. This annual monitoring helps to understand the impact of evaluations on organisational learning, departmental programming and decision making, as well as accountability. It also provides an opportunity to address potential gaps or changes needed in planned actions; capture lessons and better grasp how evaluations contribute to broader learning and whether they are drivers of change; and support the department in implementing any intended change.

References

ADB (2020), *2020 Annual Evaluation Review: ADB's Project Level Self-Evaluation System | Asian Development Bank*, Asian Development Bank, https://www.adb.org/documents/2020-annual-evaluation-review-adb-s-project-level-self-evaluation-system. [8]

AFD (2023), *Evaluations report 2023 | AFD - Agence Française de Développement*, Agence Française de Développement, https://www.afd.fr/en/ressources/evaluations-report-2023. [6]

Cracknell, B. (2000), "Evaluating Development Aid: Issues, Problems, and Solutions", *Public Administration and Development*, Vol. 21/5, pp. 429-429, https://doi.org/10.1002/pad.190. [2]

DANIDA (2021), *Evaluation of Danish Support for Climate Change Adaptation in Developing Countries*, https://um.dk/en/danida/results/eval/eval_reports/evaluation-of-danish-support-for-climate-change-adaptation-in-developing-countries. [5]

IDEV (2020), *Evaluation Synthesis of Gender Mainstreaming at the AfDB, Summary report*, Independent Development Evaluation (IDEV) African Development Bank Group, https://idev.afdb.org/sites/default/files/documents/files/Evaluation%20Synthesis%20of%20Gender%20Mainstreaming%20at%20the%20AfDB%20-%20Summary%20Report.pdf. [4]

OECD (2016), *Evaluation Systems in Development Co-operation: 2016 Review*, OECD Publishing, Paris, https://doi.org/10.1787/9789264262065-en. [7]

OECD (2010), *Evaluation in Development Agencies*, Better Aid, OECD Publishing, Paris, https://doi.org/10.1787/9789264094857-en. [3]

Serrat, O. (2009), *Learning from Evaluation*, Asian Development Bank, Manila, https://www.adb.org/sites/default/files/publication/27604/learning-evaluation.pdf (accessed on 1 September 2022). [1]

Annex A. Member and Observer Organisations of the OECD DAC Network on Development Evaluation

Table A A.1. List of EvalNet Member and Observer Organisations

Member	Organisation	Unit	Short Name
Australia	Department of Foreign Affairs and Trade	Development Evaluation and Assurance Section	Australia – DFAT
Austria	Federal Ministry for European and International Affairs	Unit for Evaluation and Quality Management	Austria – BMEIA
Austria	Austrian Development Agency	Evaluation Unit	Austria – ADA
Belgium	Foreign Affairs, Foreign Trade and Development Cooperation	Special Evaluation Office	Belgium – FPS Foreign Affairs
Belgium	Belgian Development Agency	Evaluation Office	Belgium – ENABEL
Canada	Global Affairs Canada	Evaluation and Results Bureau	Canada – GAC
Czechia	Ministry of Foreign Affairs	Department of Development Cooperation and Humanitarian Aid	Czechia – MFA
Denmark	Ministry of Foreign Affairs	Department for Evaluation, Learning and Quality	Denmark – MFA
European Union	European Commission	Performance, Results and Evaluation	EC
Finland	Ministry for Foreign Affairs	Development Evaluation Unit	Finland – MFA
France	Ministere de l'Europe et des Affaires Etrangeres	Pôle de l'Evaluation et de la Performance	France – MEAE
France	Agence Francaise de Developpement	Departement Evaluation et Apprentissage	France – AFD
France	Direction Generale du Tresor	Unite d'Evaluation des Activites de Developpement	France – Treasury
Germany	Federal Ministry for Economic Cooperation and Development	Evaluation Unit	Germany – BMZ
Germany	Deutsche Gesellschaft für Internationale Zusammenarbeit	Corporate Evaluation Unit	Germany – GIZ
Germany	German Institute for Development Evaluation	N/A	Germany – Deval
Germany	KfW Development Bank	Financial Co-operation Evaluation Department	Germany – KfW
Iceland	Ministry for Foreign Affairs	Independent Evaluation Unit	Iceland – MFA
Ireland	Department of Foreign Affairs	Evaluation and Audit Unit	Ireland – DFA
Italy	Ministry of Foreign Affairs and International Co-operation	Directorate General for Development Co-operation	Italy – ESTERI
Italy	Italian Agency for Development Co-operation	N/A	Italy – AICS
Japan	Ministry of Foreign Affairs	ODA Evaluation Division	Japan – MOFA
Japan	Japan International Cooperation Agency	Evaluation Department	Japan – JICA
Korea	Korean International Cooperation Agency	Evaluation Office	Korea – KOICA
Korea	Economic Development Cooperation Fund	Evaluation & Operation Review Department	Korea – EDCF
Luxembourg	Ministry of Foreign and European Affairs	N/A	Luxembourg – MFEA
Luxembourg	Luxembourg Development Co-operation Agency	N/A	Luxembourg – LuxDev

Member	Organisation	Unit	Short Name
The Netherlands	Ministry of Foreign Affairs	Policy and Operations Evaluation Department	The Netherlands - MFA
New Zealand	Ministry of Foreign Affairs and Trade	Monitoring, Evaluation, Research and Learning Unit	New Zealand - MFAT
Norway	Norwegian Agency for Development Cooperation	Department for Evaluation	Norway - Norad
Portugal	Institute for Cooperation and Language-Camões, I.P.	Evaluation and Audit Unit (Gabinete de Avaliação e Auditoria)	Portugal - Camões, I.P.
Spain	Ministry of Foreign Affairs, European Union and Cooperation	Development Policy Evaluation and Knowledge Management Division	Spain - MAEC
Sweden	Swedish International Development Cooperation Agency	Evaluation Unit	Sweden - SIDA
Sweden	Expert Group for Aid Studies	N/A	Sweden - EBA
Switzerland	Federal Department of Foreign Affairs	N/A	Switzerland - SDC
Switzerland	State Secretariat for Economic Affairs	Quality and Resources Unit	Switzerland - SECO
United Kingdom	Foreign and Commonwealth Development Office	Evaluation Unit	UK - FCDO
United Kingdom	Independent Commission on Aid Impact	N/A	UK - ICAI
United States	United States of America Department of State	Office of Foreign Assistance and Resources	USA - State
United States	United States Agency for International Development	Bureau for Policy Planning and Learning Office	USA - USAID
United States	Millennium Challenge Corporation	N/A	USA - MCC

OBSERVERS

Organisation	Unit	Short Name
African Development Bank	Independent Development Evaluation Department	AfDB
Asian Development Bank	Independent Evaluation Department	ADB
European Investment Bank	Evaluation Division	EIB
European Bank for Reconstruction and Development	Evaluation Department	EBRD
Inter-American Development Bank	Office of Evaluation and Oversight	IDB
International Monetary Fund	Independent Evaluation Office	IMF
Islamic Development Bank	Independent Evaluation Department	IsDB
United Nations Development Programme	Independent Evaluation Office	UNDP
World Bank	Independent Evaluation Group	World Bank

Annex B. Evaluation resourcing

The following table shows approximate spending on evaluation by participating organisations in 2021 (converted using average 2022 conversion rates from local currencies). However, as noted, there is not yet an agreed methodology for accounting development spending. As such, these figures should not be used for direct comparison.

Table A B.1. Estimated spending on evaluation by participating organisations in 2021

Country		Country		Organisation	
Australia	EUR 2 764 922	Ireland	EUR 650 000	African Development Bank	EUR 5 807 776
Austria	EUR 735 000	Italy	EUR 495 709	Asian Development Bank	EUR 13 310 275
Belgium	EUR 1 010 000	Korea	EUR 1 707 235	European Investment Bank	*Not reported*
Canada	EUR 2 959 829	Luxembourg	EUR 638 325	European Bank for Reconstruction and Development	EUR 3 290 000
Czechia	EUR 93 832	The Netherlands	EUR 5 200 000	Inter-American Development Bank	EUR 7 701 761
Denmark	EUR 2 688 645	New Zealand	EUR 509 263	International Monetary Fund	EUR 6 611 559
European Union	EUR 2 100 000	Norway	EUR 1 588 820	United Nations Development Programme	EUR 25 594 904
Finland	EUR 1 690 000	Portugal	EUR 105 993	World Bank	EUR 33 236 548
France	EUR 3 977 000	Sweden	EUR 6 371 148		
Germany	EUR 21 400 000	Switzerland	EUR 566 833		
Iceland	EUR 215 617	United States	EUR 52 700 000		

Annex C. Profiles of OECD DAC Network on Development Evaluation Members

This section provides descriptions of the individual evaluation systems for OECD DAC Network on Development Evaluation (EvalNet) members that responded to the data collection questionnaire in 2022.

Kindly note:

Human resources – in terms of full time equivalent (FTE) staff – and their reported gender identification are represented by the icons below.

1 FTE female staff	1 FTE male staff	In cases where staff work part time the figure is indicated in a lighter colour (and will sum to more than the total)

Financial resources are shown in Euros (converted using average 2022 conversion rates from local currencies).

Australia

Within the Australian Government Department of Foreign Affairs and Trade (DFAT), the Development Evaluation and Assurance Section of the Development Effectiveness and Enabling Division is responsible for international development evaluation activities. DFAT has the overall responsibility for development co-operation policy and provision.

The Director of the Development Evaluation and Assurance Section reports to the Assistant Secretary, Development Risk, Implementation and Evaluation Branch within the Development Effectiveness and Enabling Division. DFAT has an independent oversight committee – the Aid Governance Board.

Purpose

DFAT highlights that evaluations provide evidence and lessons to underpin development spending, support an organisational culture focused on innovation and continuous learning, and strengthen programme management while reinforcing public confidence in government by credibly demonstrating the achievements of the development programme.

Principles

DFAT evaluations are informed by the following principles: utility, prioritisation, independence and quality.

✓ **EVALUATION POLICY**	✓ **EVALUATION PLAN**	✓ **EVALUATION GUIDANCE**	✓ **MANAGEMENT RESPONSE**
Australia's DFAT **Evaluation Policy**, approved in 2020, guides evaluation efforts.	Australia's DFAT has an **annual development evaluation plan**.	Australia's DFAT guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment**.	Australia's DFAT **requires** a **management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Austria's DFAT conducted **42** evaluations, consisting of **37** decentralised and **5** joint evaluations. Between 2017 and 2021, it completed **191** evaluations.

2017	2018	2019	2020	2021
41	45	50	13	42
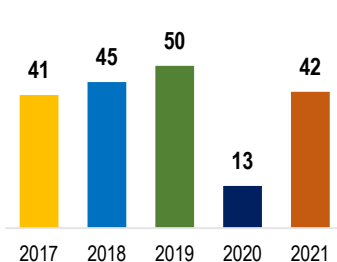
SPENDING

The total amount spent on development and humanitarian evaluations for Australia's DFAT has **increased** from 2016 to 2021.

2016	2021
€ 1 151 597	€ 2 764 922
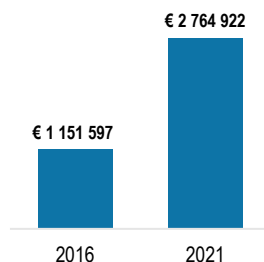
STAFFING

In 2021, Australia's DFAT evaluation unit consisted of **5** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Australia's DFAT evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ✅ Deciding to undertake specific evaluations
- ✅ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ✅ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Austria

The Federal Ministry for European and International Affairs (BMEIA) and the Austrian Development Agency (ADA) share responsibility for strategically evaluating Austrian development co-operation and humanitarian assistance. In 2018, BMEIA established a Unit for Evaluation and Quality Management within the Department for Three Year Programme, Target Group Support and Evaluation of its Directorate General for Development Cooperation. ADA has a separate Evaluation Unit situated within the Executive Unit of Evaluation and Statistics which reports directly to the Managing Director of ADA.

ADA's Evaluation Unit serves as the centre of competence for development evaluation within Austrian development co-operation (ADC). Together with BMEIA's Evaluation Unit, it develops and implements the ADC strategic evaluation plan in line with national and international standards. In addition to strategic evaluations, the ADA Evaluation Unit also manages meta-evaluations and impact studies and provides technical guidance and support to programme and project evaluations which are commissioned by project managers or implementing partners.

Purpose

In Austria, evaluations fulfil a learning, steering, and accountability/communication function. This means that evaluations (i) support institutional learning and ongoing improvement of development co-operation; (ii) contribute to evidence-informed decision making; and (iii) inform accountability and communication to partners and the public.

Principles

Austria's evaluations are informed by the following principles: independence, impartiality, credibility, transparency, utility, feasibility, fairness, accuracy, participation, and partnership.

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	MANAGEMENT RESPONSE
Austria's **Evaluation Policy**, approved in 2019, guides evaluation efforts.	Austria has a **two-year evaluation plan**.	BMEIA and ADA guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment**.	*Data not available*

Austrian Federal Ministry for European and International Affairs (BMEIA)

NUMBER OF EVALUATIONS

In 2021, Austria's BMEIA conducted **2** centralised evaluations. Between 2017 and 2021, it completed **11** evaluations.
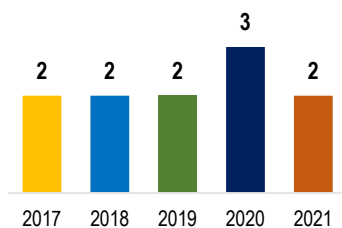
2017	2018	2019	2020	2021
2	2	2	3	2

SPENDING

The total amount spent on development and humanitarian evaluations for Austria's BMEIA is not available.
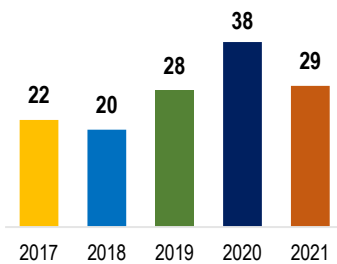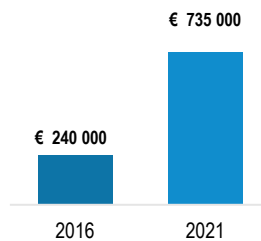
STAFFING

In 2021, Austria's BMEIA evaluation unit consisted of **2** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Austria's BMEIA evaluations in the following ways:

❌ Setting evaluation strategy or work programme

❌ Deciding to undertake specific evaluations

✅ Developing evaluation terms of reference

✅ Facilitating/participating in data collection/field visits

❌ Tracking/implementing follow up actions

LEGEND:
❌ NO
✅ OCCASIONALLY
✅ REGULARLY
⭐ ALWAYS

Austrian Development Agency (ADA)

NUMBER OF EVALUATIONS

In 2021, Austria's ADA conducted **29** centralised evaluations. Between 2017 and 2021, it completed **137** evaluations:

Year	Value
2017	22
2018	20
2019	28
2020	38
2021	29

SPENDING

The total amount spent on development and humanitarian evaluations for Austria's ADA has increased from 2016 to 2021.

Year	Value
2016	€ 240 000
2021	€ 735 000

STAFFING

In 2021, Austria's ADA evaluation unit consisted of **3.5** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Austria's ADA evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ✅ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ❌ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Belgium

The Directorate-General for Development Cooperation and Humanitarian Aid (DGD) is integrated into the Federal Public Service Foreign Affairs, Foreign Trade and Development Cooperation (FPS Foreign Affairs). DGD looks after various aspects of development co-operation (governmental, multilaterals, non-governmental, private sector, etc.) and humanitarian aid. Within FPS Foreign Affairs, the Special Evaluation Office (SEO) leads evaluations of activities funded by Belgian development co-operation.

The Head of the SEO – the Special Evaluator – is under the administrative authority of the President of the Management Committee of FPS Foreign Affairs, and reports to parliament annually. A committee made up of major stakeholders of Belgium development co-operation provides oversight.

The Belgium Development Agency (ENABEL) implements the government's development co-operation priorities. ENABEL has its own Evaluation Office that is responsible for conducting strategic evaluations of its programmes. The Evaluation Office reports directly to ENABEL's management board. In addition, since 2021, the Operational Unit for Planning-Accountability-Learning-Monitoring has been responsible for organising and managing the evaluation of co-operation programmes and projects.

Purpose

In FPS Foreign Affairs, evaluations support decision making, drawing out lessons to improve Belgian development co-operation, while also ensuring proper accountability for the use of public funds. In ENABEL, evaluations of projects and programmes aim to gauge the performance of the agency's work, as well as to generate evidence for learning within each intervention, and overall, within the organisation.

Principles

FPS Foreign Affairs and ENABEL draw on the OECD DAC Principles for Evaluation of Development Assistance of impartiality, independence, credibility, usefulness, and participation.

Belgium Federal Public Service Foreign Affairs, Foreign Trade and Development Cooperation (FPS Foreign Affairs)

| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Belgium's **Evaluation Policy**, approved in 2014, guides evaluation efforts.	Belgium's FPS Foreign Affairs develops a **two-year evaluation plan**.	Belgium's FPS Foreign Affairs has **published evaluation guidance.**	Belgium's FPS Foreign Affairs r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Belgium's FPS Foreign Affairs conducted **3** centralised evaluations. Between 2017 and 2021, it completed **11** evaluations.
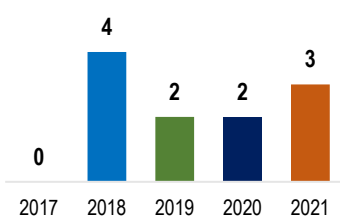
4
2
2
3
0
2017 2018 2019 2020 2021

SPENDING

The total amount spent on development and humanitarian evaluations for Belgium's FPS Foreign Affairs is only available for 2021.

€ 750 000
2016 2021
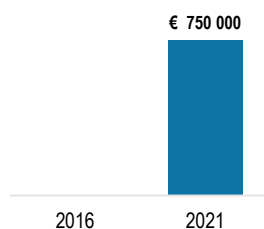
STAFFING

In 2021, Belgium's FPS Foreign Affairs evaluation unit consisted of **6** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Belgium's FPS Foreign Affairs evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ❌ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ❌ Tracking/implementing follow up actions

LEGEND:
❌ NO
🟡 OCCASIONALLY
✅ REGULARLY
⭐ ALWAYS

The Belgium Development Agency (ENABEL)

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Belgium's **Evaluation Policy**, approved in 2014, guides evaluation efforts.	Belgium's ENABEL develops a **two-year evaluation plan**.	Belgium's ENABEL has **published evaluation guidance.**	Belgium's ENABEL **requires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Belgium's ENABEL conducted **17** centralised evaluations. Between 2017 and 2021, it completed **147** evaluations.
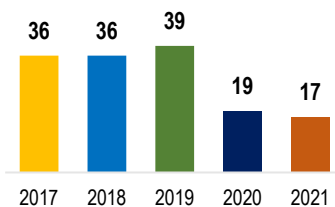
Year	Value
2017	36
2018	36
2019	39
2020	19
2021	17

SPENDING

The total amount spent on development and humanitarian evaluations for Belgium's ENABEL is only available for 2021.

€ 260 000

Year	Value
2016	
2021	€ 260 000
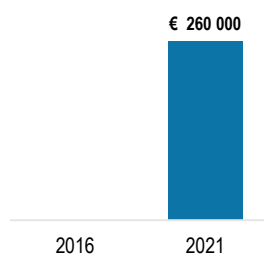
STAFFING

In 2021, Belgium's ENABEL evaluation unit consisted of **2** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Belgium's ENABEL evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ❌ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ✅ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Canada

Global Affairs Canada (GAC) leads Canada's development co-operation efforts. Within GAC, the Evaluation and Results Bureau (PRD) is responsible for evaluating international development co-operation, humanitarian assistance and other foreign policy activities, as applicable.

The Director General of PRD reports to the Assistant Deputy Minister of the Strategic Policy Branch on administrative matters and to the Deputy Minister of International Affairs on functional matters. GAC has an independent oversight committee – the Performance Measurement and Evaluation Committee (PMEC).

Purpose

GAC highlights that evaluations support evidence-based decision making and continuous policy and programme improvement, while contributing to accountability obligations by keeping citizens informed on results achieved and resources used to achieve them.

Principles

GAC evaluations are informed by the following principles: do no harm, gender equality, feminist evaluation approaches, independence, neutrality, international humanitarian assistance principles, principles for engaging in fragile states, human rights principles, and learning from evaluation.

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Canada's GAC **Evaluation Policy**, approved in 2016, guides evaluation efforts.	Canada's GAC has **five-year evaluation plan.**	Canada's GAC guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Canada's GAC r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Canada's GAC conducted **7** evaluations.
Between 2017 and 2021, it completed **46** evaluations:

11 (2017), 7 (2018), 8 (2019), 13 (2020), 7 (2021)
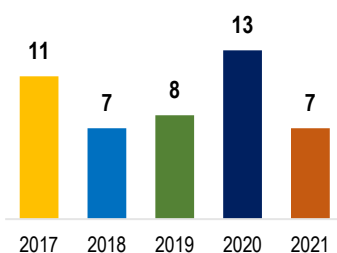
SPENDING

The total amount spent on development and humanitarian evaluations for Canada's GAC has **increased** from 2016 to 2021.

€ 1 300 100 (2016), € 2 959 829 (2021)
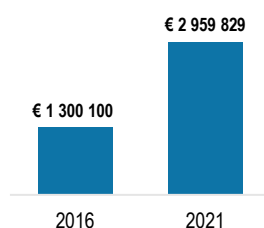
STAFFING

In 2021, Canada's GAC evaluation unit consisted of **41** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Canada's GAC evaluations in the following ways:

❌ Setting evaluation strategy or work programme
❌ Deciding to undertake specific evaluations
❌ Developing evaluation terms of reference
❌ Facilitating/participating in data collection/field visits
❌ Tracking/implementing follow up actions

LEGEND:
❌ NO
🟡 OCCASIONALLY
🟢 REGULARLY
⭐ ALWAYS

Czechia

In Czechia, the Ministry of Foreign Affairs (MFA) leads, oversees, and coordinates the delivery of the country's ODA. The Czech Development Agency is a state organisation under the authority of the MFA, and implements bilateral country development initiatives.

Within the MFA, the evaluation function of development co-operation is embedded in the Department of Development Cooperation and Humanitarian Aid (DCD). The DCD reports on its evaluation activities to, and consults with, the Council of Czech Development Co-operation – an interdepartmental co-ordination body. In addition, the DCD holds a dialogue with the Czech Evaluation Society, which associates evaluators and evaluating companies, conducts training and workshops, and disseminates the code of conduct for evaluators.

Purpose

MFA highlights that evaluations provide a systematic and objective assessment of the success of development interventions in terms of relevance, coherence, effectiveness, efficiency, impact and sustainability. Evaluations also provide recommendations for improving future development interventions and contributing to innovation.

Principles

MFA evaluations are informed by the following principles: sustainability and common benefit, balanced learning, and accountability goals.

| ✕ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Czechia **does not have an Evaluation Policy** but is currently considering adopting one.	Czechia has an **annual evaluation plan**.	Czechia's guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment**.	Czechia r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations**.

NUMBER OF EVALUATIONS

In 2021, Czechia's MFA conducted **5** centralised evaluations. Between 2017 and 2021, it completed **23** evaluations.

Year	Number
2017	4
2018	5
2019	5
2020	4
2021	5
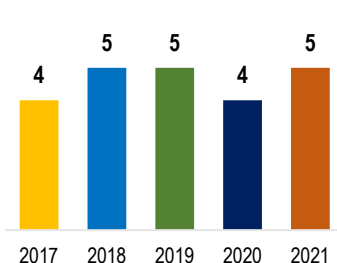
SPENDING

The total amount spent on development and humanitarian evaluations for Czechia's MFA has **increased** from 2016 to 2021.

Year	Amount
2016	€ 85 000
2021	€ 93 832
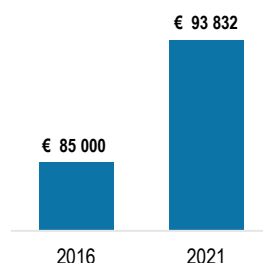
STAFFING

In 2021, Czechia's MFA evaluation unit consisted of **1** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Czechia's MFA evaluations in the following ways:

- ✓ (REGULARLY) Setting evaluation strategy or work programme
- ✓ (OCCASIONALLY) Deciding to undertake specific evaluations
- ✓ (OCCASIONALLY) Developing evaluation terms of reference
- ★ (ALWAYS) Facilitating/participating in data collection/field visits
- ★ (ALWAYS) Tracking/implementing follow up actions

LEGEND:
- ✕ NO
- ✓ OCCASIONALLY
- ✓ REGULARLY
- ★ ALWAYS

Denmark

Under the Ministry of Foreign Affairs, the Danish International Development Agency (Danida) is responsible for providing development co-operation. Danida's Department for Evaluation, Learning and Quality designs and manages development evaluations.

The head of the evaluation unit reports to the Minister of Foreign Affairs through the State Secretary for Development Policy. Both the Danida Programme Committee and the Council for Development Policy share the independent oversight role.

Purpose

The MFA highlights that evaluations support learning with a view to improving the quality and results of development co-operation, while contributing to accountability through reporting and communicating results to stakeholders, including both the public in Denmark and within partner countries.

Principles

MFA evaluations are informed by the following principles: independence, transparency, quality, utility, ethics, partnership, and capacity development and participation.

| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Denmark's **Evaluation Policy for Danish Development Cooperation**, approved in 2020, guides evaluation efforts.	Denmark's Danida develops a **two-year rolling evaluation plan.**	Denmark's guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Denmark's Danida r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Denmark's Danida conducted **9** evaluations. Between 2017 and 2021, it completed **29** evaluations.

5 — 2017
8 — 2018
6 — 2019
6 — 2020
4 — 2021
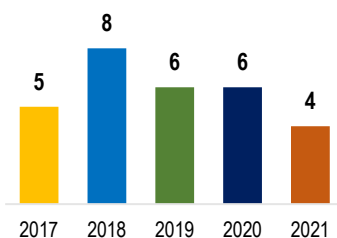
SPENDING

The total amount spent on development and humanitarian evaluations for Denmark's Danida has **decreased** from 2016 to 2021.

€ 3 000 000 — 2016
€ 2 688 645 — 2021
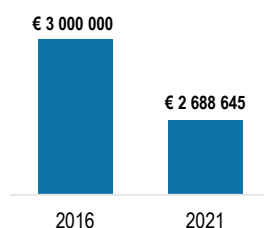
STAFFING

In 2021, Denmark's Danida evaluation unit consisted of **4** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Denmark's Danida evaluations in the following ways:

- ✓ Setting evaluation strategy or work programme
- ✓ Deciding to undertake specific evaluations
- ✓ Developing evaluation terms of reference
- ✓ Facilitating/participating in data collection/field visits
- ✓ Tracking/implementing follow up actions

LEGEND:
- ✗ NO
- ✓ OCCASIONALLY
- ✓ REGULARLY
- ★ ALWAYS

European Commission

Within the European Union (EU) institutions, the European Commission's unit Performance, Results and Evaluation; Internal Communication, Knowledge Management and Collaborative Methods in Directorate General International Partnerships (DG INTPA) is responsible for evaluating the EU's development co-operation. The Head of Monitoring and Evaluations reports to the Head of Unit, who then reports to the relevant Director at DG INTPA.

Purpose

The European Commission (EC) highlights that evaluations generate knowledge, facilitate evidence-based decision making, and policy and programme improvement, while contributing to accountability obligations.

Principles

EC evaluations are informed by the following principles: transparency, independence, ethics, quality, and utility for EC interventions.

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
The European Commission's **Evaluation Matters: The Evaluation Policy for European Union Development Co-operation**, approved in 2014, guides evaluation efforts.	The European Commission has a **three-year evaluation plan.**	The European Commission's guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	The European Commission's r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, the European Commission conducted **4** centralised evaluations. Between 2017 and 2021, it completed **28** evaluations.

7 — 2017
6 — 2018
5 — 2019
6 — 2020
4 — 2021
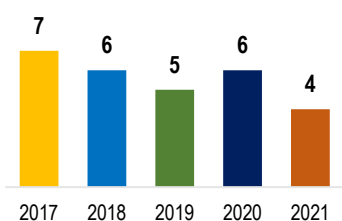
SPENDING

The total amount spent on development and humanitarian evaluations for the European Commission has decreased from 2016 to 2021..

€ 5 000 000 — 2016
€ 2 100 000 — 2021
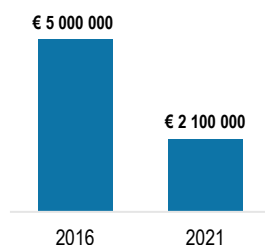
STAFFING

In 2021, the European Commission evaluation unit consisted of **13** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in European Commission evaluations in the following ways:

- Setting evaluation strategy or work programme (OCCASIONALLY)
- Deciding to undertake specific evaluations (REGULARLY)
- Developing evaluation terms of reference (OCCASIONALLY)
- Facilitating/participating in data collection/field visits (REGULARLY)
- Tracking/implementing follow up actions (REGULARLY)

LEGEND:
NO
OCCASIONALLY
REGULARLY
ALWAYS

Finland

Within the Ministry of Foreign Affairs (MFA), the Department for Development Policy is responsible for Finland's international development policy, development co-operation policy, and development and humanitarian financing. Different departments and units in the MFA manage development co-operation. The Development Evaluation unit is responsible for evaluating development co-operation, humanitarian assistance, and other foreign policy activities, as applicable.

The Development Evaluation unit reports directly to the Under-Secretary of State responsible for development policy and development co-operation. The MFA does not have a committee that provides oversight of evaluation.

Purpose

The MFA highlights that evaluations ensure high-quality development co-operation by contributing to learning and accountability through independent and impartial evidence-based knowledge. It also highlights their role for transparency and openness of development policies and for decision making related to the use of development and humanitarian funding.

Principles

MFA evaluations are informed by the following principles: utility, independence, impartiality, credibility, honesty, integrity, accountability, participatory approaches, partnerships, and working jointly.

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Finland's **Development Evaluation Norm,** approved in 2015, guides evaluation efforts.	Finland has a **three-year evaluation plan.**	Finland's guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Finland r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Finland's MFA conducted **25** evaluations.
Between 2017 and 2021, it completed **150** evaluations.

33 35 32 25 25
2017 2018 2019 2020 2021
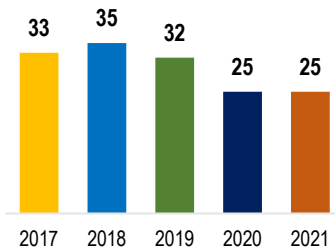
SPENDING

The total amount spent on development and humanitarian evaluations in Finland's MFA has **decreased** from 2016 to 2021.

€ 2 000 000
€ 1 690 000
2016 2021
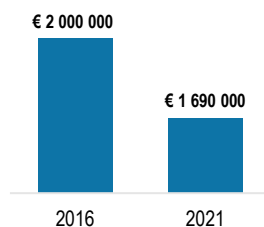
STAFFING

In 2021, Finland's MFA evaluation unit consisted of **5** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Finland's MFA evaluations in the following ways:

❌ Setting evaluation strategy or work programme
✅ Deciding to undertake specific evaluations
✅ Developing evaluation terms of reference
✅ Facilitating/participating in data collection/field visits
❌ Tracking/implementing follow up actions

LEGEND:
❌ NO
🟡 OCCASIONALLY
✅ REGULARLY
⭐ ALWAYS

France

French development co-operation is provided by a variety of government entities, including the *Ministere de l'Europe et des Affaires Étrangères* (MEAE), *Direction Générale du Trésor* (DG Treasury), and *Agence Française de Développement* (AFD).

Within MEAE, the evaluation unit is situated in the Delegation of Programmes and Operators (DPO) of the Directorate General for Globalisation (DGM). It carries out strategic evaluations of MEAE development initiatives, as well as providing support to other government entities undertaking evaluations.

Within DG Treasury, the Development Evaluation Unit (UEAD) is attached to the Multilateral Affairs and Development Department. UEAD is responsible for evaluating development co-operation provided by DG Treasury. Within AFD, the evaluation unit is part of the Innovation, Strategy and Research Department. It is responsible for evaluating development co-operation provided by AFD.

In 2021, a commission was established within the Court of Auditors to conduct evaluations of development co-operation activities undertaken though other French ministries.

Purpose

The MEAE evaluation highlights that evaluation is a tool to increase accountability, support decision making, and promote continuous learning.

DG Treasury highlights that evaluations improve decision making and serve the dual purposes of accountability and learning.

AFD highlights that evaluations support accountability, effectiveness, and continuous learning.

Principles

All French evaluations are informed by the OECD DAC Principles for Evaluation of Development Assistance: impartiality, independence, credibility, usefulness, and participation.

Ministère de l'Europe et des Affaires Étrangères (MEAE)

| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
France's MEAE **Development Evaluation Policy** guides evaluation efforts.	France's MEAE has a **three-year evaluation plan.**	France's MEAE guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	France's MEAE r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, France's MEAE conducted **98** evaluations, including **15** centralised, **83** decentralised and **3** joint evaluations. Between 2017 and 2021, it completed **250** evaluations.
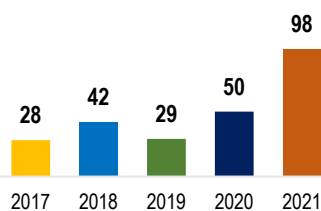
28 2017
42 2018
29 2019
50 2020
98 2021

SPENDING

The total amount spent on development and humanitarian evaluations for France's MEAE has **increased** from 2016 to 2021.

€ 450 000 — 2016
€ 1 430 000 — 2021
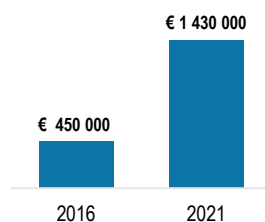
STAFFING

In 2021, France's MEAE evaluation unit consisted of **3.5** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in France's MEAE evaluations in the following ways:

❌ Setting evaluation strategy or work programme
❌ Deciding to undertake specific evaluations
✅ Developing evaluation terms of reference
⭐ Facilitating/participating in data collection/field visits
✅ Tracking/implementing follow up actions

LEGEND:
❌ NO
✅ OCCASIONALLY
✅ REGULARLY
⭐ ALWAYS

Direction Générale du Trésor (DG Treasury)

| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
France's DG Treasury **Development Evaluation Policy** approved in 2017, guides evaluation efforts.	France's DG Treasury has a **three-year evaluation plan.**	France's DG Treasury guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	France's DG Treasury r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, France's DG Treasury conducted **3** centralised evaluations. Between 2017 and 2021, it completed **11** evaluations.
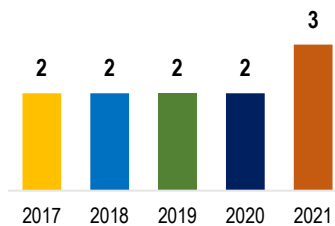
SPENDING

The total amount spent on development and humanitarian evaluations for France's DG Treasury has **decreased** from 2016 to 2021.
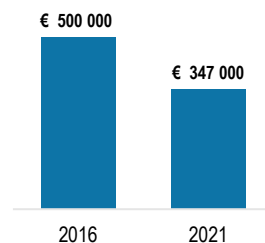
STAFFING

In 2021, France's DG Treasury evaluation unit consisted of **3** full-time equivalent staff

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in France's DG Treasury's evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ❌ Developing evaluation terms of reference
- ⭐ Facilitating/participating in data collection/field visits
- ❌ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Agence Française de Développement (AFD)

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
France's AFD Development **Evaluation Policy**, approved in 2013, guides evaluation efforts.	France's AFD has a **three-year evaluation plan** in place.	Guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment**.	France's AFD **Requires** a **management response** to evaluations. The **evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, France's AFD conducted **74** evaluations, including **62** decentralised evaluations. Between 2017 and 2021, it completed **284** evaluations.
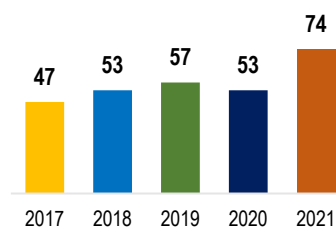
	47	53	57	53	74
	2017	2018	2019	2020	2021

SPENDING

The total amount spent on development and humanitarian evaluations for France's AFD has **increased** from 2016 to 2021.

€ 500 000 — 2016
€ 2 200 000 — 2021
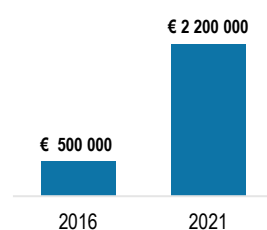
STAFFING

In 2021, France's AFD evaluation unit consisted of **20** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in France's AFD evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ✅ Deciding to undertake specific evaluations
- ✅ Developing evaluation terms of reference
- ⭐ Facilitating/participating in data collection/field visits
- ✅ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Germany

Overall responsibility for development evaluation in Germany falls under the remit of the Federal Ministry for Economic Co-operation and Development (**BMZ**). Under the direction of BMZ, other larger actors in this system are the German Institute for Development Evaluation (**DEval**), the German Corporation for International Co-operation (**GIZ**) and the KfW Development Bank (**KfW**).

Within **BMZ**, the BMZ Evaluation Unit provides direction to the evaluation system as a whole, and has a key role as an interface between the BMZ and DEval. BMZ has largely delegated evaluation work to other actors. In the very few cases where BMZ conducts (commissions) evaluations itself, the Head of the Evaluation and Research Division reports to the State Secretary via the Director of Directorate GS2 "Data and Effectiveness". As a Federal Ministry the BMZ does not have an independent oversight committee. It is controlled by the German Bundestag.

DEval is mandated by BMZ to independently conduct scientifically sound, strategic and policy-relevant evaluations of German development co-operation. Other tasks include evaluation capacity development and conducting research on evaluation methods and standards. The Director of DEval reports to the BMZ State Secretary and, through them, to relevant committees of the German Bundestag. The institute also reports to its own independent oversight committee – the Advisory Board.

GIZ implements German ("technical") development co-operation, under the purview of BMZ. As part of this, its Corporate Evaluation Unit conducts organisation-specific project and strategic evaluations. The Director of the Evaluation Unit reports to the GIZ Management Board, and for project evaluations, the unit also reports to BMZ. GIZ also has a supervisory board.

Similarly, **KfW** is an implementing agency of BMZ (for "financial co-operation") and conducts ex-post and impact evaluations of projects that it implements on behalf of BMZ. The Director of the Financial Co-operation Evaluation Department reports to the member of the independent Board of Supervisory Directors responsible for development cooperation. For project evaluations, KfW's evaluation department also reports to BMZ.

Evaluation work done by civil society organisations (CSOs) complements the government's evaluation system and are also guided by the BMZ Evaluation Policy, mentioned below.

Purpose

BMZ's system-wide evaluation policy describes the evaluation system's overarching goal to be improving development effectiveness and the legitimacy of interventions. In addition, evaluations support knowledge and learning and hence decision making, as well as accountability. Evaluations can also promote an informed, evidence-based dialogue with development professionals and the interested public.

DEval highlights that its research-based evaluation enhances the effectiveness, efficiency, and sustainability of development co-operation. By conducting transparent and independent evaluative work, DEval aims at fulfilling the learning and accountability function of evaluation and contributing to the legitimacy of development policy.

GIZ outlines that evaluations generate evidence-based results and recommendations for decision making, as well as providing credible proof of effectiveness and increasing the transparency of results.

KfW notes that evaluations aim to determine the impact of its co-operation efforts, as well as improve effectiveness, accountability, transparency, and institutional learning.

Principles

Evaluations conducted under BMZ's remit are informed by the following principles: usefulness of the results, and credibility of the findings, independence of assessments, partnership, and ethical standards that incorporate human right principles. These evaluation principles, and the associated standards and application of the criteria, are closely aligned with the principles of the Aid Effectiveness agenda agreed in Paris (2005), Accra (2008) and Busan (2011), as well as the 2030 Agenda for Sustainable Development.

Deval's evaluations are informed by the following principles: utility, evaluability, fairness, independence, accuracy, and comparability.

GIZ's evaluations are informed by usefulness, credibility, independence, partnerships and ethical standards, as well as taking into account human rights principles.

KfW's evaluations are informed by the principles of independence, efficiency, innovation, learning and ethics, as well as the OECD DAC's Principles for Evaluation of Development Assistance.

Germany's Federal Ministry for Economic Co-operation and Development (BMZ)

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Germany's BMZ **Evaluation Policy,** approved in 2021, guides evaluation efforts.	Germany's BMZ follows Deval's **three-year evaluation plan**.	Germany's BMZ guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment**.	Germany's BMZ **requires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

The number of evaluations for Germany's BMZ is not available since overall responsibility for development evaluation in Germany falls under the remit of the BMZ. Please refer to German Institute for Development Evaluation (DEval); the German Corporation for International Cooperation (GIZ); and the KfW Development Bank (KfW).

SPENDING

The total amount spent on development and humanitarian evaluations for Germany's BMZ has **increased** from 2016 to 2021.

€ 2 500 000

€ 1 500 500

2016   2021
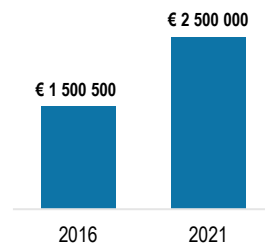
STAFFING

In 2021, Germany's BMZ evaluation unit consisted of **5** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Germany's BMZ evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ❌ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ❌ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- 🟡 OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

German Institute for Development Evaluation (DEval)

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Germany's DEval **Evaluation** efforts are guided by the **institute's strategy, renewed every five years.**	Germany's DEval has a **three-year evaluation plan**.	Germany's DEval guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment**.	Germany's DEval **monitors the implementation of its recommendations** based on the **implementation plan provided by BMZ**. A synthesis report of this **monitoring is published every two years.**

NUMBER OF EVALUATIONS

In 2021, Germany's DEval conducted **9** evaluations, including **5** centralised, **1** joint and **3** other evaluations. Between 2017 and 2021, it completed **36** evaluations.
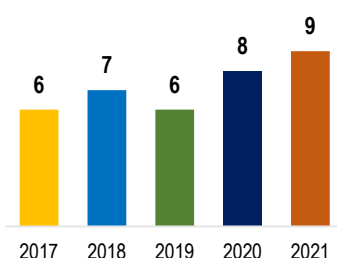
SPENDING

The total amount spent on development and humanitarian evaluations for Germany's DEval has **increased** from 2016 to 2021.
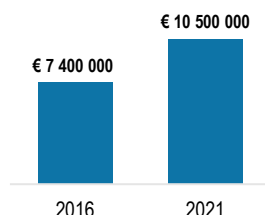
STAFFING

In 2021, Germany's DEval consisted of **55.5** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Germany's Deval evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ✅ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ✅ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

German Corporation for International Co-operation (GIZ)

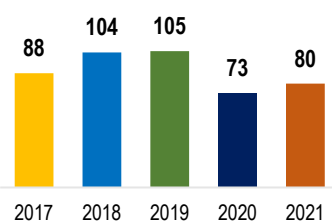| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Germany's GIZ **Evaluation Policy,** approved in 2018, guides evaluation efforts.	Germany's GIZ has an **annual evaluation plan.**	*Data not available*	Germany's BMZ r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Germany's GIZ conducted **85** centralised evaluations. Between 2017 and 2021, it completed **332** evaluations.
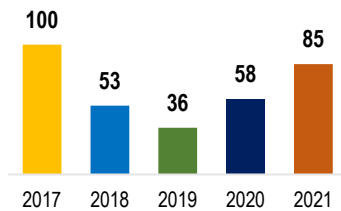
Year	Value
2017	100
2018	53
2019	36
2020	58
2021	85

SPENDING

The total amount spent on development and humanitarian evaluations for Germany's GIZ has **increased** from 2016 to 2021.

Year	Value
2016	€ 2 200 000
2021	€ 6 000 000
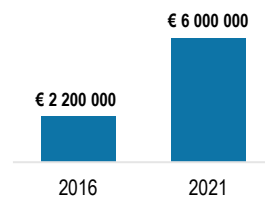
STAFFING

In 2021, Germany's GIZ evaluation unit consisted of **20.7** full-time equivalent staff.
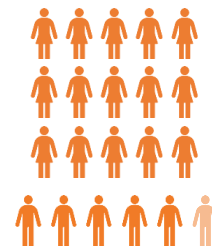
The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Germany's GIZ evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- 🟡 Deciding to undertake specific evaluations
- ✅ Developing evaluation terms of reference
- ⭐ Facilitating/participating in data collection/field visits
- 🟡 Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- 🟡 OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

German KfW Development Bank (KfW)

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Germany's KfW **Evaluation Policy,** approved in 2021, guides evaluation efforts.	Germany's KfW has an **annual evaluation plan.**	Germany's KfW guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Germany's KfW **does not** requires **a management response** to evaluations. However, the **Evaluation unit reports the results** of each evaluation **to management and operational units**.

NUMBER OF EVALUATIONS

In 2021, Germany's KfW conducted **80** centralised evaluations. Between 2017 to 2021, it completed **450** evaluations.

88 — 2017
104 — 2018
105 — 2019
73 — 2020
80 — 2021

SPENDING

The total amount spent on development and humanitarian evaluations for Germany's KfW has **decreased** from 2016 to 2021.

€ 4 800 000 — 2016
€ 2 400 000 — 2021
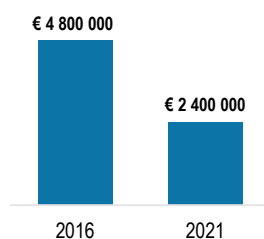
STAFFING

In 2021, Germany's KfW evaluation unit consisted of **15.5** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Germany's KfW evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ✅ Developing evaluation terms of reference
- ⭐ Facilitating/participating in data collection/field visits
- ⭐ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Iceland

Within the Government of Iceland, the Directorate of International Development, Ministry for Foreign Affairs (MFA) is responsible for development co-operation, including allocating and overseeing the implementation of nearly all of the country's official development assistance (ODA) budget. Within the MFA, the Directorate for Internal Affairs is responsible for the evaluation of development co-operation and humanitarian assistance initiatives. It is responsible for managing a wide range of evaluations and reviews, supervising and revising evaluation methodology, and participating in international co-operation on related issues.

The Director of Internal Affairs reports to the Permanent Secretary of State of the Ministry for Foreign Affairs. While no committee formally provides oversight of evaluations, the Development Cooperation Committee carries out some oversight functions.

Purpose

The MFA highlights that evaluations support evidence-based policymaking, effectiveness, and institutional learning, while enhancing accountability to Icelandic taxpayers and other stakeholders. In addition, they support monitoring and results-based management for strategic planning and intervention design.

Principles

The MFA evaluations are informed by the following principles: impartiality, independence, transparency, credibility, and usefulness.

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Iceland's MFA **Evaluation Policy** approved in 2020, guides evaluation efforts.	Iceland's MFA has an **annual evaluation plan.**	Iceland's MFA guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Iceland's MFA **requires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Iceland's MFA conducted **4** decentralised evaluations. Between 2017 and 2021, it completed **21** evaluations.

Year	Value
2017	5
2018	8
2019	2
2020	2
2021	4
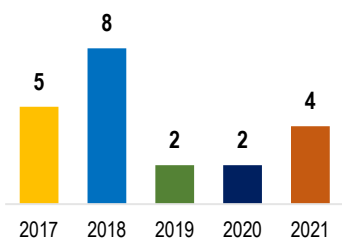
SPENDING

The total amount spent on development and humanitarian evaluations for Iceland's MFA is only available for 2021.

€ 215 617

2016	2021
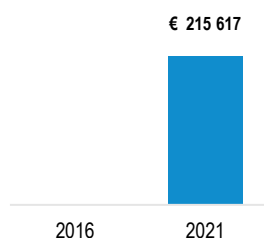
STAFFING

In 2021, Iceland's MFA Independent Evaluation and Results Unit consisted of **1.2** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Iceland's MFA evaluations in the following ways:

- 🟡 Setting evaluation strategy or work programme
- 🟢 Deciding to undertake specific evaluations
- ⭐ Developing evaluation terms of reference
- ⭐ Facilitating/participating in data collection/field visits
- 🟢 Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- 🟡 OCCASIONALLY
- 🟢 REGULARLY
- ⭐ ALWAYS

Ireland

Within the Department of Foreign Affairs (DFA), the Evaluation and Audit Unit is responsible for evaluation across the full range of the DFA's activities, including development and humanitarian assistance. Responsibility for development co-operation and humanitarian assistance is fully integrated within DFA and has its own division, the Development Co-operation and Africa Division.

The Director of Evaluation and Audit reports to the Secretary General of the DFA. The DFA's Audit Committee, which is formally appointed by and advises the Secretary General, provides independent appraisal of the audit and evaluation arrangements, with a view to strengthening governance, risk management and internal controls, as well as helping to enhance the effective operation of the audit and evaluation functions within the department.

Purpose

Evaluation in the DFA serves two main purposes: accountability to show and share the results of investment with key stakeholders; and learning to support decision making and improvements in policy, strategy, programme, and project implementation.
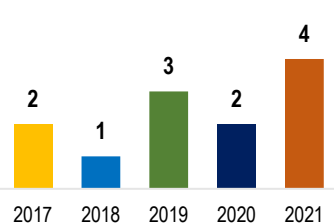
Principles

DFA evaluations are informed by the following principles: utility, credibility, independence, inclusion, and participation.

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Ireland's DFA **Evaluation Policy** approved in 2022, guides evaluation efforts.	Ireland's DFA has a **three-year rolling evaluation plan.**	Ireland's DFA guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Ireland's DFA **requires a management response** recommendations in its evaluation and review reports, setting out agreed actions, timelines and responsible persons. The **Evaluation and Audit Unit** plays a lead role in **tracking implementation** of the agreed actions, and **providing progress updates to the Audit Committee**.

NUMBER OF EVALUATIONS
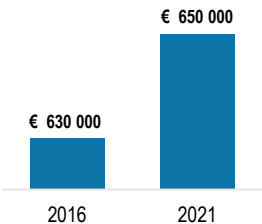
In 2021, Ireland's DFA conducted **4** evaluations, including **3** centralised evaluations and **1** spending review. Between 2017 and 2021, it completed **12** evaluations.

SPENDING

The total amount spent on development and humanitarian evaluations for Ireland's DFA has **increased** from 2016 to 2021.

STAFFING

In 2021, Ireland's DFA Evaluation and Audit Unit consisted of **6.5** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Ireland's DFA evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ❌ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ❌ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Italy

Italian development co-operation is primarily channelled through the Ministry of Foreign Affairs and International Cooperation (ESTERI) and the Italian Agency for Development Cooperation (AICS).

Within ESTERI, the Directorate General for Development Co-operation (DGDC), Office III is responsible for evaluating the impact of development co-operation initiatives. The Head of Office III reports to the Director-General of Development Cooperation. Office III is supported by the Evaluation Advisory Committee – an independent oversight committee – appointed by the Director-General of the Development Cooperation Directorate General.

AICS is responsible for implementing development co-operation activities and its evaluation unit is responsible for ex-ante, in-itinere and ex-post evaluations.

Purpose

In Italy, evaluations serve the purpose of learning and supporting conscious decision making, as well as informing the programming of future initiatives.

Principles

Evaluations are informed by the following principles: utility, credibility, independence, impartiality, transparency, ethics, professionalism, human rights, gender parity, and leave no-one behind.

*Italy Ministry of Foreign Affairs and International Cooperation (ESTERI)*

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Italy's **Evaluation Policy** approved in 2014, guides evaluation efforts within both ESTERI and AICS.	Italy's ESTERI has a **three-year evaluation plan**.	Italy's ESTERI guidance documents include **cross-cutting areas: inclusion, gender and human rights.**	Italy's ESTERI **requires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

The number of evaluations for Italy's ESTERI is not available.

SPENDING

The total amount spent on development and humanitarian evaluations for Italy's ESTERI has **increased** from 2016 to 2021.

€ 495 709

€ 345 445

2016   2021
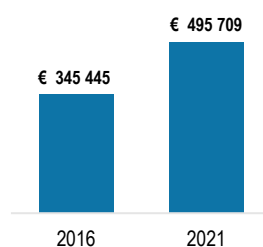
STAFFING

In 2021, there was **no full-time equivalent staff** in Italy's ESTERI's evaluation unit.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Italy's ESTERI evaluations in the following ways:

❌ Setting evaluation strategy or work programme
❌ Deciding to undertake specific evaluations
❌ Developing evaluation terms of reference
✅ Facilitating/participating in data collection/field visits
❌ Tracking/implementing follow up actions

LEGEND:
❌ NO
🟡 OCCASIONALLY
✅ REGULARLY
⭐ ALWAYS

Italy AICS

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	MANAGEMENT RESPONSE
Italy's **Evaluation Policy** approved in 2014, guides evaluation efforts within both ESTERI and AICS.	Italy's AICS has a **three-year evaluation plan.**	Italy's AICS guidance documents include **cross-cutting areas: inclusion, gender and human rights.**	*Data not available*

NUMBER OF EVALUATIONS	SPENDING	STAFFING
The number of evaluations for Italy's AICS is not available.	*The total amount spent on development and humanitarian evaluations for Italy's AICS is not available.*	In 2021, there was **no full-time equivalent staff** in Italy's AICS evaluation unit

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Italy's AICS evaluations in the following ways:

❌ Setting evaluation strategy or work programme

❌ Deciding to undertake specific evaluations

❌ Developing evaluation terms of reference

✅ Facilitating/participating in data collection/field visits

❌ Tracking/implementing follow up actions

LEGEND:
❌ NO
🟡 OCCASIONALLY
✅ REGULARLY
⭐ ALWAYS

Japan

The Ministry of Foreign Affairs (MOFA) is responsible for setting development co-operation policy and priorities, as well as co-ordinating collaboration across ministries. The Japan International Cooperation Agency (JICA) is responsible for delivering the majority of Japan's development co-operation, under the strategic guidance of MOFA.

Within MOFA, the ODA Evaluation Division is responsible for policy-level evaluations such as country/regional assistance evaluations and thematic/aid modality evaluations. The Director of the ODA Evaluation Division reports evaluation results to the ministry's senior officials, including the Deputy Assistant Minister, the Management and Cooperation Division of the Minister's Secretariat, and the Director-General of MOFA's International Cooperation Bureau.

Within JICA, the Evaluation Department is responsible for project evaluations and it reports directly to the JICA President. JICA has an oversight committee, the Advisory Committee on Evaluation.

Purpose

MOFA highlights that evaluations improve ODA management and promote public accountability.

JICA notes that its evaluations support accountability as well as learning and taking actions.

Principles

Both MOFA and JICA evaluations are informed by the following principles: independence, competence and capacity building, transparency and participation, co-ordination with donors and partner countries, and quality assurance.

Japan Ministry of Foreign Affairs (MOFA)

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Japan's **Development Cooperation Charter**, which was approved by the Cabinet in 2015, guides evaluation efforts.	Japan's MOFA has an **annual ODA evaluation plan.**	Japan's MOFA guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Japan's MOFA r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Japan's MOFA conducted **6** evaluations, comprising **5** centralised evaluations and **1** decentralised evaluation. Between 2017 and 2021, it completed **31** evaluations.

Year	Value
2017	8
2018	5
2019	5
2020	7
2021	6
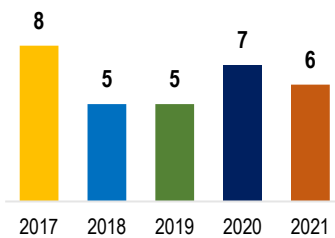
SPENDING

The total amount spent on development and humanitarian evaluations for Japan's MOFA is only available for 2016.

€ 6 571 000 (2016)
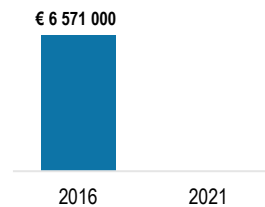
STAFFING

In 2021, Japan's MOFA evaluation unit consisted of **7** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Japan's MOFA evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ❌ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ❌ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Japan International Cooperation Agency (JICA)

| ✓ **EVALUATION POLICY**	✓ **EVALUATION PLAN**	✓ **EVALUATION GUIDANCE**	✓ **MANAGEMENT RESPONSE**
Japan's **Development Cooperation Charter**, which was approved by the Cabinet in 2015, guides evaluation efforts.	Japan's JICA has an **annual ODA evaluation plan.**	Japan's JICA guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Japan's JICA **requires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Japan's JICA conducted **157** evaluations, including **73** centralised evaluations and **84** decentralised evaluations. Between 2017 and 2021, it completed **721** evaluations.

192	181	167	181	157
2017	2018	2019	2020	2021
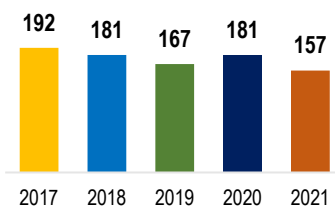
SPENDING

The total amount spent on development and humanitarian evaluations for Japan's JICA is only available for 2016.

€ 942 000

2016	2021
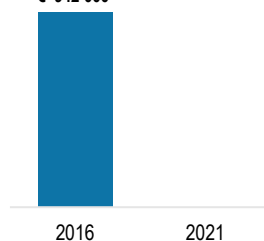
STAFFING

In 2021, Japan's JICA evaluation unit consisted of **23** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Japan's JICA evaluations in the following ways:

- 🟢 Setting evaluation strategy or work programme
- 🟡 Deciding to undertake specific evaluations
- 🟢 Developing evaluation terms of reference
- 🟡 Facilitating/participating in data collection/field visits
- 🟢 Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- 🟡 OCCASIONALLY
- 🟢 REGULARLY
- ⭐ ALWAYS

Korea

Development co-operation in Korea is provided by the Korean International Cooperation Agency (KOICA) and the Economic Development Cooperation Fund (EDCF). KOICA is responsible for bilateral grants and EDCF is responsible for bilateral loans.

Within KOICA, the Evaluation Department undertakes evaluations of development co-operation projects. The head of the Evaluation Department reports to the President of KOICA. KOICA has an independent oversight committee – the Government Evaluation Committee.

Within EDCF, the Evaluation and Operation Review Department is responsible for evaluating projects. The department is comprised of multiple evaluation teams that each report to the Executive Director. EDCF has an independent oversight committee – the Subcommittee for Integrated Evaluation, under the Committee for International Development Cooperation in the Prime Minister's Office.

Purpose

KOICA highlights that evaluations support accountability, as well as learning, which in turn guides evidence-based decision making and contributes to continuous policy and programme improvement.

EDCF highlights that evaluations aim to improve future policies, programmes and projects by providing insight into lessons learned. Evaluations also provide a basis for enhanced accountability, including by making development co-operation information publicly available.

Principles

KOICA and EDCF evaluations are informed by the following principles: impartiality, independence, credibility, usefulness, and partnership.

Korean International Cooperation Agency (KOICA)

| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Korea's KOICA **Evaluation Policy** approved in 2018, guides evaluation efforts.	Korea's KOICA has an **annual evaluation plan.**	Korea's KOICA guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Korea's KOICA r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Korea's KOICA conducted **105** centralised evaluations. Between 2017 and 2021, KOICA completed **255** evaluations:

10 — 2017
12 — 2018
80 — 2019
48 — 2020
105 — 2021
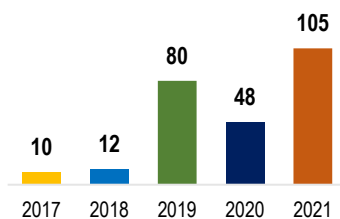
SPENDING

The total amount spent on development and humanitarian evaluations for Korea's KOICA has **increased** from 2016 to 2021.

€ 1 100 000 — 2016
€ 1 599 905 — 2021
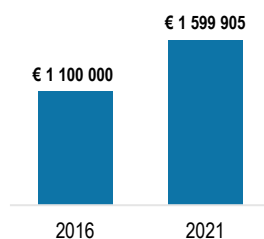
STAFFING

In 2021, Korea's KOICA evaluation unit consisted of **7** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Korea's KOICA evaluations in the following ways:

- ✓ Setting evaluation strategy or work programme
- ✓ Deciding to undertake specific evaluations
- ✓ Developing evaluation terms of reference
- ✓ Facilitating/participating in data collection/field visits
- ✗ Tracking/implementing follow up actions

LEGEND:
- ✗ NO
- ✓ OCCASIONALLY
- ✓ REGULARLY
- ★ ALWAYS

Korean Economic Development Cooperation Fund (EDCF)

| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Korea's EDCF **Evaluation Policy** approved in 2020, guides evaluation efforts.	Korea's EDCF has an **annual evaluation plan.**	Korea's EDCF guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Korea's EDCF r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Korea's EDCF conducted **3** centralised evaluations. Between 2017 and 2021, it completed **34** evaluations.

Year	2017	2018	2019	2020	2021
	8	10	7	6	3
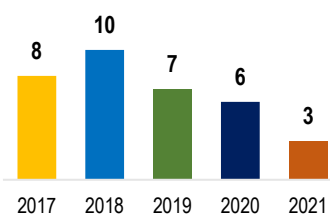
SPENDING

The total amount spent on development and humanitarian evaluations for Korea's EDCF has **decreased** from 2016 to 2021.

€ 622 000 (2016)
€ 107 330 (2021)
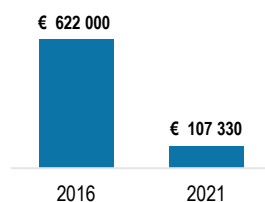
STAFFING

In 2021, Korea's EDCF evaluation unit consisted of **6** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Korea's EDCF evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ✅ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ✅ Tracking/implementing follow up actions

LEGEND:
❌ NO
✅ OCCASIONALLY
✅ REGULARLY
⭐ ALWAYS

Luxembourg

The Ministry of Foreign and European Affairs (MFEA) is responsible for setting out the high-level principles and policy that guide Luxembourg's development co-operation efforts. Lux-Development (LuxDev) is the country's lead development co-operation implementation agency.

Within MFEA, the Evaluation Office is responsible for setting evaluation priorities and conducting a small number of evaluations directly. The head of the Evaluation Office reports directly to the Office of the Minister. MFEA does not have an independent evaluation oversight committee.

Within LuxDev, the Evaluation, Quality and Knowledge Management Department is responsible for evaluating its projects. The head of the department reports to the General Directorate of LuxDev. LuxDev does not have an independent oversight committee.

Purpose

MFEA and LuxDev highlight that evaluation aims to improve the quality of programming, provide evidence on effectiveness and enhance transparency and accountability.

Principles

MFEA evaluations are informed by the OECD DAC Principles for Evaluation of Development Assistance: impartiality, independence, credibility, usefulness, and participation.

LuxDev evaluations are informed by the following principles: systematisation, impartiality and independence, partnership, feedback, transparency, and capitalisation.

Luxembourg Ministry of Foreign and European Affairs (MFEA)

| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Luxembourg's MFEA **Evaluation Policy** approved in 2015, guides evaluation efforts.	Luxembourg's MFEA has a **biennial evaluation plan.**	Luxembourg's MFEA guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Luxembourg's MFEA **requires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Luxembourg's MFEA conducted **4** centralised evaluations. Between 2017 and 2021, MFEA completed **27** evaluations:
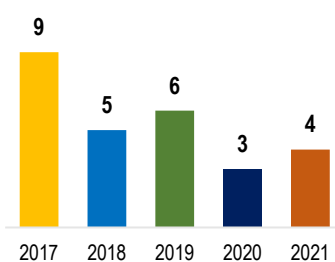
9 — 2017
5 — 2018
6 — 2019
3 — 2020
4 — 2021

SPENDING

The total amount spent on development and humanitarian evaluations for Luxembourg's MFEA has **decreased** from 2016 to 2021.

€ 600 000 — 2016
€ 396 480 — 2021
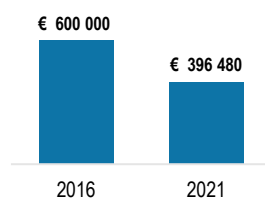
STAFFING

In 2021, Luxembourg's MFEA evaluation unit consisted of **1.5** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Luxembourg's MFEA evaluations in the following ways:

- ⊘ Setting evaluation strategy or work programme
- ✗ Deciding to undertake specific evaluations
- ⊘ Developing evaluation terms of reference
- ★ Facilitating/participating in data collection/field visits
- ★ Tracking/implementing follow up actions

LEGEND:
✗ NO
⊘ OCCASIONALLY
✓ REGULARLY
★ ALWAYS

Luxembourg Lux-Development (LuxDev)

| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Luxembourg's LuxDev **Internal Evaluation Policy** approved in 2019, guides evaluation efforts.	Luxembourg's LuxDev has a **biennial evaluation plan.**	Luxembourg's LuxDev guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Luxembourg's LuxDev r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Luxembourg's LuxDev conducted **15** evaluations, including **13** centralised and **2** decentralised evaluations. Between 2017 and 2021, it complete **53** evaluations.
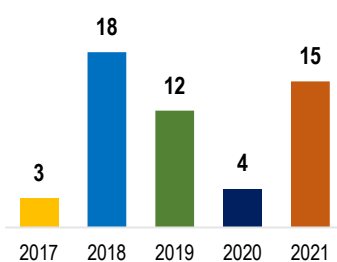
SPENDING

The total amount spent on development and humanitarian evaluations for Luxembourg's LuxDev is only available for 2021.
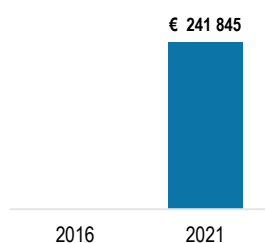
STAFFING

In 2021, Luxembourg's LuxDev evaluation unit consisted of **1.5** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Luxembourg's LuxDev evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ✅ Deciding to undertake specific evaluations
- ✅ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ⭐ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Netherlands

Within the Ministry of Foreign Affairs of the Netherlands, the Policy and Operations Evaluation Department (IOB) is responsible for evaluating the ministry's entire field of operations: foreign policy, trade, and development co-operation. The IOB is structurally part of the Ministry, but its independence is save-guarded by its formal protocol and its independent director.

The IOB reports are made public and its policy evaluations are actively shared with the relevant Parliamentary commissions and the general public. The ministry has no independent oversight committee but an independent external reference group is composed for each study.

Purpose

The Netherlands highlights that evaluations support both accountability and policy-oriented learning. They aim to use the findings for future policymaking and strategy development.
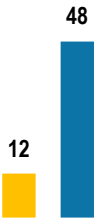
Principles

Evaluations are informed by the following principles: independence, objectivity, transparency, validity, reliability, partnership, feasibility, correctness, cost effectiveness, usability, gender, anonymity and confidentiality, responsibility, and respect and understanding for other cultures and universal values.

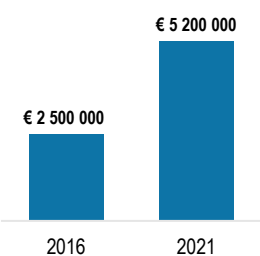| ✓ **EVALUATION POLICY**	✓ **EVALUATION PLAN**	✓ **EVALUATION GUIDANCE**	✓ **MANAGEMENT RESPONSE**
The Netherlands' **Evaluation Policy** follows the national directives of the Ministry of Finance, responsible for the national co-ordination and quality control of policy evaluations. National directives for evaluation were approved in 2003, 2009, and updated in 2021.	Netherlands' IOB has a **five to seven-year evaluation cycle** that is **updated annually** and shared with Parliament and the general public.	Netherlands' IOB guidance documents includes **gender as cross-cutting area.**	Netherland's MFA r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Netherlands' IOB conducted **60** evaluations, including **12** centralised evaluations and **48** decentralised evaluations.

48

12

SPENDING

The total amount spent on development and humanitarian evaluations for Netherlands' IOB has **increased** from 2016 to 2021.

€ 5 200 000

€ 2 500 000

2016   2021

STAFFING

In 2021, Netherlands' IOB evaluation unit consisted of **30** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Netherlands' IOB evaluations in the following ways:

- ✓ Setting evaluation strategy or work programme
- ✓ Deciding to undertake specific evaluations
- ✓ Developing evaluation terms of reference
- ✓ Facilitating/participating in data collection/field visits
- ✓ Tracking/implementing follow up actions

LEGEND:

❌ NO
✓ OCCASIONALLY
✓ REGULARLY
⭐ ALWAYS

New Zealand

In New Zealand, the Monitoring, Evaluation, Research and Learning (MERL) unit based in the Pacific and Development Group (PDG) of the Ministry of Foreign Affairs (MFAT) is responsible for evaluating development co-operation. It gives internal technical advice on the commissioning and design of evaluations, builds evaluation capacity, and provides technical quality assurance for evaluations.

The MERL unit manager reports to the Divisional Manager of the Development, Capability and Insights Division who in turn reports to the Deputy Secretary of the Pacific and Development Group. PDG has an independent oversight committee – the Pacific and Development Strategic Governance Group.

Purpose

MFAT highlights that evaluations support evidence-informed design and delivery of development assistance, continuous improvement and adaptation, accountability and transparency.

Principles

MFAT's evaluations are informed by the following principles: clarity of intent, credibility and rigour, utility, culturally and contextually responsive, partnership and participation, inclusion, resilience, donor co-operation and collaboration, safeguarding and ethical practice as well as a strengths-based approach to capacity building.

| ✓ **EVALUATION POLICY**	✓ **EVALUATION PLAN**	✓ **EVALUATION GUIDANCE**	✓ **MANAGEMENT RESPONSE**
New Zealand's MFAT **Evaluation Policy** approved in 2020, guides evaluation efforts.	New Zealand's MFAT has a **three-year evaluation plan.**	New Zealand's MFAT guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	New Zealand's MFAT r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, New Zealand's conducted **10** decentralised evaluations. Between 2017 and 2021, it completed **54** evaluations.

9	11	18	6	10
2017	2018	2019	2020	2021
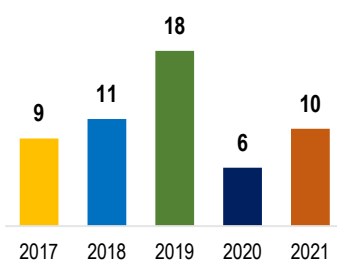
SPENDING

The total amount spent on development and humanitarian evaluations for New Zealand's has **decreased** from 2016 to 2021.

€ 727 709 € 509 263

2016	2021
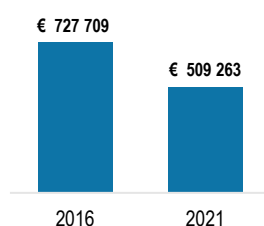
STAFFING

In 2021, New Zealand's evaluation unit consisted of **4** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in New Zealand's evaluations in the following ways:

- ✓ Setting evaluation strategy or work programme
- ✓ Deciding to undertake specific evaluations
- ✓ Developing evaluation terms of reference
- ★ Facilitating/participating in data collection/field visits
- ✓ Tracking/implementing follow up actions

LEGEND:
- ✗ NO
- ✓ OCCASIONALLY
- ✓ REGULARLY
- ★ ALWAYS

Norway

The Department for Evaluation located in the Norwegian Agency for Development Cooperation (Norad) performs independent evaluations of Norwegian development co-operation, as well as of Norway's International Climate and Forest Initiative. The department's work is governed under a separate mandate and associated evaluation strategy and guidelines issued by the Ministry of Foreign Affairs and the Ministry of Climate and Environment.

The head of the Department for Evaluation reports directly to the Secretary Generals of the two ministries. The evaluation department in the Norwegian development aid administration does not have an independent oversight committee.

Purpose

The purpose of Norway's evaluations is in part to promote a knowledge-based approach to Norwegian development co-operation, and in part to hold Norwegian development policy actors accountable for the management of funds.

Principles

Evaluations are informed by the following principles: credibility, independence, transparency, use, ethical aspects, and participation.

| ✓ **EVALUATION POLICY**	✓ **EVALUATION PLAN**	✓ **EVALUATION GUIDANCE**	✓ **MANAGEMENT RESPONSE**
Norway's Evaluation Department's work is governed under a **dedicated mandate and associated evaluation strategy and guidelines**, as **issued** jointly by the **Ministry of Foreign Affairs and the Ministry of Climate and Environment.**	Norway's Norad has a **three-year evaluation plan, updated annually.**	Norway's Norad guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Norway's completed evaluation reports, together with a memo in Norwegian summing up the main findings from the evaluation and the Evaluation Department's recommendations for follow up, are sent to the concerned ministry i.e., the Ministry of Foreign Affairs and/or the Ministry of Climate and Environment. The ministries assess the need to take any decisions and/or draw up follow-up plans.

NUMBER OF EVALUATIONS

In 2021, Norway's Norad Evaluation Department conducted **5** evaluations, including 2 centralised, 2 joint evaluations and 1 mapping. Between 2017 and 2021, it completed **44** evaluations.

2017	2018	2019	2020	2021
11	13	2	13	5
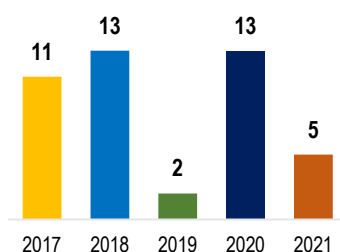
SPENDING

The total amount spent on development and humanitarian evaluations for Norway's Norad has **decreased** from 2016 to 2021.

2016	2021
€ 3 300 000	€ 1 588 820
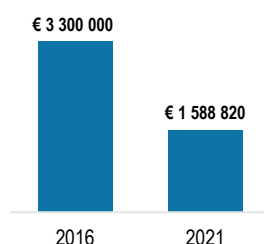
STAFFING

In 2021, Norway's Norad evaluation department consisted of **7.5** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Norway's Norad evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ❌ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ❌ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Portugal

The Institute for Cooperation and Language (Camões, I.P.) is responsible for monitoring and evaluating Portuguese development co-operation efforts and language and cultural promotion activities. The Office of Evaluation and Audit (*Gabinete de Avaliação e Auditoria* - GAA) within Camões, I.P. is in charge of evaluating the development co-operation delivered directly and through line ministries, as well as internal audits.

The Head of GAA reports directly to the Board of Camões, I.P. Camões, I.P. further reports to the Secretaries of State within the Ministry of Foreign Affairs. The evaluation function in the Camões, I.P. does not have an independent oversight committee.

Purpose

Camões, I.P. highlights evaluation as an important mechanism for accountability, responsibility and learning, and supporting Portuguese development co-operation's planning and management.

Principles

Evaluations are informed by the following principles: ethics, transparency, credibility and quality, independence and impartiality, competence and capacity, timeliness and usefulness.

| EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Data not available	Portugal's Camões, I.P. has a **three-year rolling evaluation plan**, which is reviewed annually.	Portugal's Camões, I.P. guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Portugal's Camões, I.P. **requires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Portugal's Camões, I.P. conducted **3** centralised evaluations. Between 2017 and 2020, it completed **9** evaluations.

6
2
1
0
3

2017 2018 2019 2020 2021
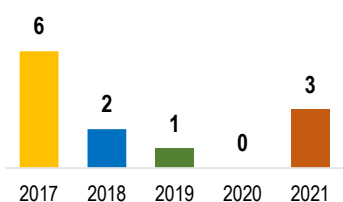
SPENDING

The total amount spent on development and humanitarian evaluations for Portugal's Camões, I.P. has **increased** from 2016 to 2021.

€ 105 993
€ 100 000

2016 2021
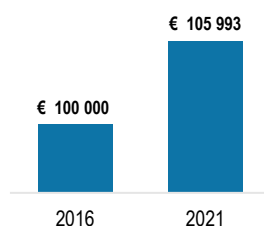
STAFFING

In 2021, Portugal's Camões, I.P. evaluation unit consisted of **3** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Portugal's Camões, I.P. evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ✅ Deciding to undertake specific evaluations
- ✅ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ✅ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Spain

Spain's Development Policy Evaluation and Knowledge Management Division (DEGCO) within the Ministry of Foreign Affairs, European Union and Cooperation (MAEC) is responsible for evaluating international development co-operation policy.

The DEGCO has a dual reporting line. On the one hand it is integrated into the Directorate-General for Sustainable Development Policies (DGPOLDES) (organic level), while on the other it reports directly to the titular head of the Secretary of State for International Cooperation (SECI). MAEC does not have an independent oversight committee.

Purpose

MAEC highlights that evaluations generate learning which improves the effectiveness of co-operation systems. Furthermore, they provide elements for transparency and accountability.

Principles

MAEC's evaluations are informed by the following principles: independence, impartiality, transparency, credibility, usefulness, efficiency, participation and capacity building, co-ordination and complementarity.

✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Spain's MAEC **Evaluation Policy** approved in 2013, guides evaluation efforts.	Spain's MAEC has a **nine-year evaluation plan.**	Spain's MAEC guidance documents include **cross-cutting areas: inclusion, gender and human rights.**	Spain's MAEC r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Spain's MAEC conducted **1** evaluation. Between 2017 and 2021, it completed **111** evaluations.

Year	Value
2017	10
2018	12
2019	8
2020	80
2021	1
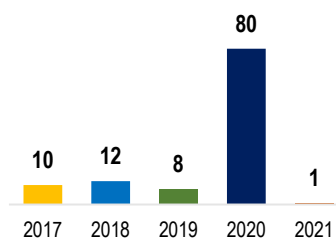
SPENDING

The total amount spent on development and humanitarian evaluations for Spain's MAEC is not available.

STAFFING

In 2021, Spain's MAEC evaluation unit consisted of **4** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Spain's MAEC evaluations in the following ways:

- ★ Setting evaluation strategy or work programme
- ✓ Deciding to undertake specific evaluations
- ✓ Developing evaluation terms of reference
- ✓ Facilitating/participating in data collection/field visits
- ✗ Tracking/implementing follow up actions

LEGEND:
- ✗ NO
- ✓ OCCASIONALLY
- ✓ REGULARLY
- ★ ALWAYS

Sweden

Two government institutions are responsible for the evaluation of Swedish international development co-operation: the Swedish International Development Cooperation Agency (Sida) and the Expert Group for Aid Studies (EBA).

Sida conducts central evaluations on strategic issues, as well as decentralised evaluations, mostly on Sida contributions. Partner-led evaluations are conducted by Sida's partners. Sida's evaluation unit is responsible for planning, commissioning, and co-ordinating central evaluations. For decentralised evaluations, the evaluation unit's role is primarily advisory and supportive. The head of the evaluation unit reports directly to the Sida Director General with regard to central evaluations. The central evaluation plan is suggested by the head of the evaluation unit to the Director General, who makes the final decision on the central evaluation plan.

The EBA is a government committee mandated to evaluate and analyse the direction, governance and implementation of Sweden's official development assistance with a specific focus on results and effectiveness. EBA focuses primarily on overarching issues within Swedish development assistance, not on individual projects.

The committee reports to the Ministry of Foreign Affairs at the Government Offices of Sweden. EBA does not have an independent oversight committee.

Purpose

Sida sees evaluations as providing an understanding of how and why certain results were (or were not) achieved; if a project or programme led to any unintended effects; and if it was implemented in a cost-efficient manner. Hence, evaluations contribute to well-informed decision making in projects, programmes, and co-operation strategies.

EBA highlights that evaluations contribute to the government's (and other actors') development and governance of aid.

Principles

Sida evaluations are informed by the following principles: utility, relevance, fitness for purpose and alignment with the OECD/DAC principles.

EBA evaluations are informed by the following principles: advancement of knowledge (does the project offer something new for actors in development assistance?); policy relevance (will the expected results be of use to policymakers, can they contribute to the development of aid?); current relevance (link to current developments in the aid sector?); share of the development assistance budget or significance of the expected impacts of the phenomenon to be studied; and important questions of principle relating to development assistance (for example, ethical issues in this area).

Swedish International Development Cooperation Agency (Sida)

| × EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Sweden's Sida **does not have a specific evaluation policy**. Its **evaluation efforts are guided** by the **operational strategy and quality policy**.	Sweden's Sida has a **three-year evaluation plan for central evaluations.**	Sweden's Sida guidance documents include **cross-cutting areas: inclusion, gender and human rights.**	Sweden's Sida r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Sweden's Sida conducted **39** evaluations, including **3** centralised, **35** decentralised and **1** other evaluation. Between 2017 and 2021, it completed **166** evaluations.
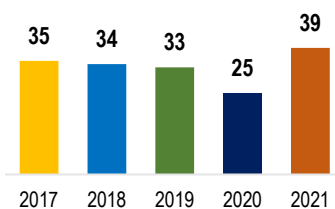
Year	Value
2017	35
2018	34
2019	33
2020	25
2021	39

SPENDING

The total amount spent on development and humanitarian evaluations for Sweden's Sida has **increased** from 2016 to 2021.

Year	Value
2016	€ 600 000
2021	€ 6 371 148
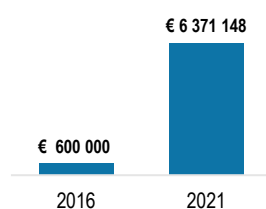
STAFFING

In 2021, Sweden's Sida evaluation unit consisted of **6** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Sweden's SIDA evaluations in the following ways:

- 🟡 Setting evaluation strategy or work programme
- 🟢 Deciding to undertake specific evaluations
- 🟢 Developing evaluation terms of reference
- ⭐ Facilitating/participating in data collection/field visits
- 🟢 Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- 🟡 OCCASIONALLY
- 🟢 REGULARLY
- ⭐ ALWAYS

Swedish Expert Group for Aid Studies (EBA)

× EVALUATION POLICY	× EVALUATION PLAN	✓ EVALUATION GUIDANCE	× MANAGEMENT RESPONSE
Sweden's EBA **does not have a specific evaluation policy**. Its evaluation efforts are **guided by its committee terms of reference, operational strategy and quality policy**.	Sweden's EBA **does not have a formal annual plan.**	Sweden's EBA guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	Sweden's EBA **does not** requires **a management response** to evaluations.

NUMBER OF EVALUATIONS

In 2021, Sweden's EBA conducted **8** centralised evaluations. Between 2017 and 2021, it completed **255** evaluations.
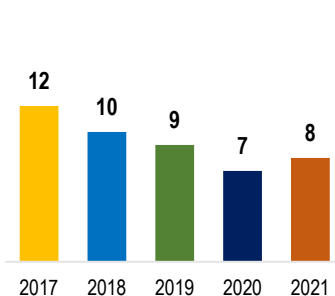
2017	2018	2019	2020	2021
12	10	9	7	8

SPENDING

The total amount spent on development and humanitarian evaluations for Sweden's EBA is only available for 2016.

€ 16 000 000

2016	2021

STAFFING

In 2021, Sweden's EBA evaluation unit consisted of **8.5** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Sweden's EBA evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ❌ Deciding to undertake specific evaluations
- ❌ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ❌ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- ✅ OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Switzerland

Two agencies in the Swiss Government are responsible for development co-operation: the Swiss Agency for Development and Co-operation (SDC) and the Economic Co-operation and Development Division of the State Secretariat for Economic Affairs (SECO).

SDC, within the Federal Department of Foreign Affairs, is responsible for evaluating technical, humanitarian, and financial co-operation programmes, targeting mostly least developed and fragile countries. The head of evaluation reports directly to the Board of Directors.

SECO, within the Federal Department of Economic Affairs, Education and Research, is responsible for evaluating the economic and trade programme targeting middle-income countries. The evaluation unit is structurally independent. The head of the evaluation reports directly to the head of policy and quality. SECO has an external evaluation committee which provides independent oversight.

Purpose

SDC highlights that evaluations advance learning with a view to improving the quality and results of development co-operation, and allow for evidence-based decision making and steering, as well as accountability through reporting and communicating the results. In addition, they complement results-based management.

SECO identifies learning, accountability, and steering of future decisions as the purposes of its evaluations.

Principles

SDC's evaluations are informed by the following principles: usefulness, feasibility, correctness, quality and reliability, participation, impartiality and independence, transparency and partnership.

SECO's evaluations are informed by the following principles: clear governance, a sound system of independent evaluations, an integration of evaluation results in knowledge management processes, and transparency about evaluation results.

Swiss Agency for Development and Co-operation (SDC)

| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
Switzerland's SDC **Evaluation Policy** approved in 2018, guides evaluation efforts.	Switzerland's SDC has an **annual evaluation plan.**	Switzerland's SDC guidance documents include **cross-cutting areas: inclusion, gender and human rights.**	Switzerland's SDC r**equires a management response** to evaluations. The **Evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Switzerland's SDC conducted **88** evaluations, comprising **4** centralised, **83** decentralised and **1** joint evaluation. Between 2017 and 2021, it completed **414** evaluations.
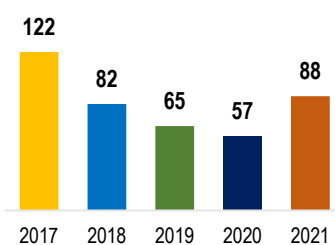
122
82
65
57
88

2017 2018 2019 2020 2021

SPENDING

The total amount spent on development and humanitarian evaluations for Switzerland's SDC has **decreased** from 2016 to 2021.

€ 1 000 000
€ 400 856

2016 2021
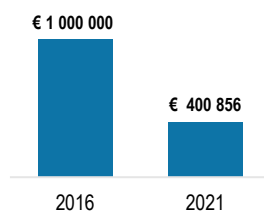
STAFFING

In 2021, Switzerland's SDC evaluation unit consisted of **4.4** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved Switzerland's SDC evaluations in the following ways:

- ✓ Setting evaluation strategy or work programme
- ✓ Deciding to undertake specific evaluations
- ✓ Developing evaluation terms of reference
- ✓ Facilitating/participating in data collection/field visits
- ✓ Tracking/implementing follow up actions

LEGEND:
⊗ NO
✓ OCCASIONALLY
✓ REGULARLY
★ ALWAYS

Swiss Economic Co-operation and Development Division of the State Secretariat for Economic Affairs (SECO)

| ✓ **EVALUATION POLICY**	✓ **EVALUATION PLAN**	✓ **EVALUATION GUIDANCE**	✓ **MANAGEMENT RESPONSE**
Switzerland's SECO **Evaluation Policy** approved in 2021, guides evaluation efforts.	Switzerland's SECO has a **four-year evaluation plan.**	Switzerland's SECO guidance documents include **inclusion, gender and human rights.**	Switzerland's SECO **requires a management response** to evaluations. The **evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, Switzerland's SECO conducted **32** evaluations, comprising **1** centralised and **31** decentralised evaluations, as well as **29** internal reviews. Between 2017 and 2021, it completed **157** evaluations.
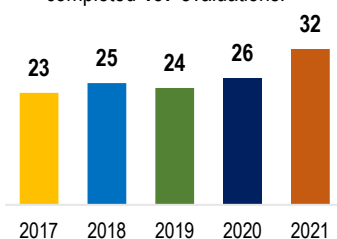
23	25	24	26	32
2017	2018	2019	2020	2021

SPENDING

The total amount spent on development and humanitarian evaluations for Switzerland's SECO is only available for 2021.

€ 165 968

2016	2021
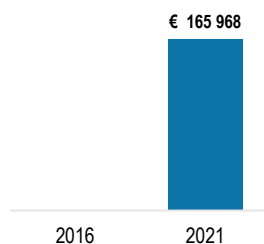
STAFFING

In 2021, Switzerland's SECO's evaluation unit consisted of **1.2** full-time equivalent staff.

The head of the evaluation unit identifies as male.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in Switzerland's SECO evaluations in the following ways:

- 🟡 Setting evaluation strategy or work programme
- 🟡 Deciding to undertake specific evaluations
- 🟢 Developing evaluation terms of reference
- 🟢 Facilitating/participating in data collection/field visits
- 🟡 Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- 🟡 OCCASIONALLY
- 🟢 REGULARLY
- ⭐ ALWAYS

United Kingdom

The Foreign, Commonwealth & Development Office (FCDO) operates a blended centralised/decentralised evaluation system, in which the majority of evaluation activity is devolved to spending units across the global network and delivered by a network of 65 embedded Evaluation Advisers. This is complemented and supported by the central Evaluation Unit and the Head of Profession for Evaluation and Government Social Research, which work to support evaluation across the FCDO by implementing the FCDO's Evaluation Strategy and ensuring, through the FCDO's Evaluation Policy, that there is a common understanding of evaluation principles and the minimum standards that support those principles. The Head of the Evaluation Unit and Head of Profession for Evaluation and Government Social Research report to the Chief Economist.

The FCDO has specific scrutiny and accountability obligations as an official development assistance (ODA) spending department. The 2015 International Development Act includes the following clause: "the Secretary of State must make arrangements for the independent evaluation of the extent to which ODA provided by the United Kingdom represents value."

In addition, the International Commission for Aid Impact (ICAI) scrutinises UK aid spending, including approaches to monitoring, evaluation, and learning. ICAI's formal remit is to provide independent evaluation and scrutiny of the impact and value for money of all UK development co-operation. ICAI operates independently of government and reports to Parliament through the House of Commons' International Development Committee (IDC) or their ICAI Sub-Committee.

Purpose

FCDO highlights that evaluations serve both a learning and accountability function, and help to enhance resource efficiency.

The purpose of ICAI's evaluations is to ensure both value for money and scrutiny of impact.

Principles

Both the FCDO's and ICAI's evaluations are informed by the following principles: usefulness, credibility, robustness, proportionate budget to expected outcomes, as well as safe and ethical practice.

| ✓ **EVALUATION POLICY**	✓ **EVALUATION PLAN**	✓ **EVALUATION GUIDANCE**	✓ **MANAGEMENT RESPONSE**
United Kingdom's FCDO **Evaluation Policy** approved in 2022, guides evaluation efforts.	United Kingdom's FCDO has an **evaluation strategy** in place, covering the period from **2022 to 2025**.	United Kingdom's FCDO guidance documents include **cross-cutting areas: inclusion, gender and human rights.**	United Kingdom's FCDO **requires a management response** to evaluations. Programme teams **play a role in monitoring** the i**mplementation of management response** recommendations.

NUMBER OF EVALUATIONS

The number of evaluations for United Kingdom's FCDO is not available.

SPENDING

The total amount spent on development and humanitarian evaluations in United Kingdom's FCDO is only available for 2016.

€ 15 100 000

2016　　　2021
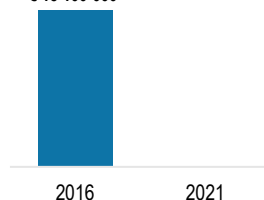
STAFFING

In 2021, United Kingdom's FCDO evaluation unit consisted of **8** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

The information for partner country engagement for United Kingdom's FCDO is not available.

United States of America

Four US government agencies – the US Department of State, USAID, the Development Finance Corporation (DFC) and the Millennium Challenge Corporation (MCC) – manage and evaluate a significant share of US foreign assistance, each with a dedicated service overseeing and supporting implementation.

Within the Department of State (State), the Office of Foreign Assistance and Resources is tasked with providing technical assistance and overviewing the implementation of the Evaluation Policy. The head of evaluation reports to the Assistant to the Administrator.

Within USAID, the Bureau for Policy, Planning and Learning's Office of Learning, Evaluation and Research is responsible for evaluation. The head of evaluations reports to the Assistant to the Administrator. The US General Accounting Office performs an oversight function.

The Development Finance Corporation (DFC) conducts portfolio performance evaluations, which are deep dives into select portfolios of projects with a sectoral, thematic, or country-level focus. The learnings gleaned from performance evaluations help inform changes to DFC policies, procedures, project structuring, and/or strategy. They also contribute to the learnings of the wider impact investing community.

The MCC outsources evaluations to third-party independent experts. Evaluations are reported to the Vice President of the Department of Policy and Evaluation. MCC has an oversight committee – the MCC Evaluation Management Committee.

Purpose

USAID highlights that evaluations support accountability to stakeholders and learning to improve development outcomes/

Principles

USAID evaluations are informed by the following principles: integration into design of strategies, projects, and activities; unbiasedness; relevance; methodological excellence; transparency; local ownership; and utilisation of findings.

| ✓ EVALUATION POLICY	✓ EVALUATION PLAN	✓ EVALUATION GUIDANCE	✓ MANAGEMENT RESPONSE
United States' USAID **Evaluation Policy** approved in 2020, guides evaluation efforts.	United States' USAID has an **annual evaluation plan.**	United States' USAID guidance documents include **cross-cutting areas: inclusion, gender, human rights, and climate/environment.**	United States' USAID r**equires a management response** to evaluations. The **evaluation unit plays a role** in **monitoring** the implementation of **management response recommendations.**

NUMBER OF EVALUATIONS

In 2021, United States' USAID conducted **94** evaluations, including **9** centralised and **85** decentralised evaluations. Between 2017 and 2021, it completed **776** evaluations.

162   189   177   154   94

2017   2018   2019   2020   2021
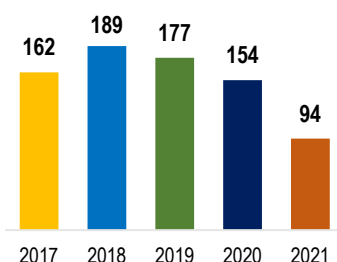
SPENDING

The total amount spent on development and humanitarian evaluations for United States' USAID has **increased** from 2016 to 2021.

€ 52 700 000

€ 14 000 000

2016   2021
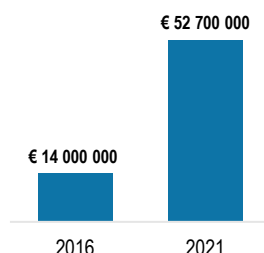
STAFFING

In 2021, United States' USAID evaluation unit consisted of **44** full-time equivalent staff.

The head of the evaluation unit identifies as female.

PARTNER COUNTRY ENGAGEMENT

Partner countries are involved in United States' USAID evaluations in the following ways:

- ❌ Setting evaluation strategy or work programme
- ✅ Deciding to undertake specific evaluations
- ✅ Developing evaluation terms of reference
- ✅ Facilitating/participating in data collection/field visits
- ✅ Tracking/implementing follow up actions

LEGEND:
- ❌ NO
- 🟡 OCCASIONALLY
- ✅ REGULARLY
- ⭐ ALWAYS

Evaluation Systems in Development Co-operation 2023

Drawing on the experiences of the members and observer organisations of the Development Assistance Committee Network on Development Evaluation (EVALNET), this study provides a snapshot of the core elements and ways of working within development evaluation systems. It offers insight to development co-operation organisations as they seek to establish or strengthen credible evaluation systems of their own, to support learning and accountability. The report also explores the role of evaluation in development co-operation and humanitarian assistance, including the various policy and institutional arrangements used with evaluation systems. It then reviews the evaluation process, from deciding on evaluations to disseminating findings and finally, focuses on how evaluation findings are used to improve development co-operation efforts.