Organisation for Economic Co-operation and Development

DCD/DAC/EV/M(2023)2

Unclassified

English - Or. English

2 October 2024

DEVELOPMENT CO-OPERATION DIRECTORATE
DEVELOPMENT ASSISTANCE COMMITTEE

Cancels & replaces the same document of 4 June 2024

DAC Network on Development Evaluation

Summary Record of the 30th Meeting of the DAC Network on Development Evaluation

15-16 November 2023, OECD Conference Centre

Members approved this draft summary record at EvalNet's 31st Meeting and is now issued as FINAL.

The 30th EvalNet Meeting took place at the OECD Conference Centre in Paris, France (with an option for virtual participation) on 15-16 November 2023. Sessions were open to all registered staff from participating EvalNet units/departments, and to guests by invitation.

A special 40th Anniversary session on 16 November was open to the broader DAC community and guests from partner countries. The meeting, including side events, was attended by more than 150 participants spanning more than 62 organisations. A list of participants is available on EMS or upon request from the Secretariat.

Megan Grace Kennedy-Chouane, megangrace.kennedy-chouane@oecd.org + (33-1) 85 55 47 49

JT03550744

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.

Summary Record: 30th Meeting of the DAC Network on Development Evaluation

This meeting was held at the OECD Conference Centre, Paris, France with a virtual option. Interpretation French/English was provided for all formal sessions.

15 November 2023

Item 1. Opening Session

1. The Chair of the DAC Network on Development Evaluation (EvalNet), Jörg Faust (Germany, DEval), opened the meeting by welcoming all participants. Members approved the agenda [DCD/DAC/EV/A(2023)2] for the 30th Meeting and the Secretariat confirmed that the summary record for the 29th EvalNet Meeting [DCD/DAC/EV/M(2023)1] had been approved via written procedure, issued as final and declassified.

2. The Secretariat led the process of confirming the EvalNet Bureau composition for 2024. Lena Johanssen de Chateau (Head of Evaluation Unit at the Swedish International Development Co-operation Agency, Sweden) was elected as the new vice chair. Chair, Jörg Faust (Executive Director at the German Institute for Development Evaluation, Germany), and Vice-Chair, Daniel Jacobo Orea (Advisor at the Division of Development Policy Evaluation and Knowledge Management, Spanish Ministry of Foreign Affairs and Cooperation, Spain) were confirmed in their roles. Members expressed support and appreciation to the Bureau, and thanked outgoing Vice-Chair Isabelle Mercier for her contributions. The Secretariat reminded members that the network strives to achieve gender and geographic balance in the Bureau composition over time and encouraged non-European members to consider a role in the future.

3. The Chair welcomed Lithuania and Estonia as the newest DAC members, as well as the Asian Infrastructure Investment Bank (AIIB) and the Islamic Development Bank (IsDB), whose status as observers of the network was approved by the DAC as part of the mandate renewal process.

Item 2. Member updates, work proposals and other business

4. The Network welcomed the following new participants:

- Rachel Ingwersen, Acting Director, Development Evaluation and Assurance Section, Department of Foreign Affairs and Trade, Australia
- Jean-Claude Pires, Directeur du Département Évaluation et Apprentissage, Direction Innovation Recherche et Savoirs, French Agency for Development (AFD), France
- Loukianos-Eleftherios Klint, Expert Counsellor, Directorate for Development Co-operation Policy, Greek Ministry of Foreign Affairs, Greece
- Toshiya Abe, Director, Evaluation Unit, Japan International Cooperation Agency, Japan
- Mindaugas Cesiulis, Third Secretary, Development Co-operation Department, Ministry of Foreign Affairs, Lithuania
- Silvana Hogg, Director, Evaluation and Control Unit, Swiss Agency for Development Cooperation, Switzerland
- Moa Chenon, Senior Evaluation Advisor and Deputy Head of Unit, Swedish International Development Agency, Sweden
- Sarah Honeywell, Deputy Head, Evaluation Unit, Foreign and Commonwealth Development Organisation, United Kingdom

- Adam Trowbridge, Senior Evaluation Advisor, Office of Foreign Assistance, United States Department of State, United States
- Marvin Taylor-Dormund, Managing Director, Complaints Resolution Evaluation and Integrity, Asian Investment, Infrastructure, and Investment Bank (AIIB)
- Vivek Arora, Deputy Director, Independent Evaluation Office, International Monetary Fund (IMF)
- Sabine Bernabe, Director General, Independent Evaluation Group, World Bank

5. Megan Grace Kennedy-Chouane (EvalNet Secretariat) updated members that the Evaluation Systems Study, was launched as a pre-publication version in November, and shared the timeline for the final publication. She noted the need for additional work on comparing resourcing for evaluations across institutions, and proposed that this work will go forward in the new year depending on capacity. She announced recruitment of an Evaluation Policy Analyst to fill the current vacancy in the Secretariat and reminded members of the option to provide staff on loan or in-kind staff support, and encouraged participants to consider this modality, which can be of mutual benefit to the staff member, the network, and the sending institution.

6. Isabelle Mercier (UNDP) briefed members on the SDG Synthesis Coalition. The Coalition was established to provide evidence on SDG achievement. The group has established five groups to synthesise evidence on each of the SDG pillars. While there have been successes, several challenges remain including an imbalance in evidence and the inclusion of different types of evaluation. Ms. Mercier proposed further consultations to discuss the future direction of the coalition, deepening partnerships with organisations outside the evaluation space and the establishment of new funding partners. She also proposed a more focused approach, prioritising studies and sharing living evidence syntheses within pillars. Members were invited to participate in these consultations and provide feedback.

7. Marisa Berbegal Ibanez (OECD) Working Party on Statistics (WP-STAT) discussed revisions to how evaluation costs are reported in the Creditor Reporting System (CRS) and requested member feedback on whether the typology was appropriate. Three types of evaluations were identified: those linked to projects or programmes, strategic and cross-cutting evaluations, and investments in evidence generation. Members were invited to provide feedback on these classifications. Astrid Ganterer (Austria) proposed adding "institutional evaluation" to the "strategic evaluation" category for further clarity. In Austria, strategic evaluations involve policymaking and operational engagement. Jan Pettersson (Sweden, EBA) shared that current reporting directives are detailed and burdensome to report on and proposed combing the categories and to avoid inflation of administrative costs. Members had no further discussion.

Item 3. What is quality evaluation? Evaluation quality standards in practice

8. Vera Hundt (Secretariat) presented a paper reflecting on 40 years of EvalNet's discussions on evaluation quality. There has been a shift in focus, from rigorous impact evaluations as the sole marker of quality, to broader acceptance of a range of designs and methodologies, with an increased emphasis on considerations of relevance and use. The interplay between methodological quality, process quality, timing and utilisation is evident, with observed synergies and trade-offs between different dimensions. The OECD/DAC quality standards and principles for evaluating development assistance are widely recognised and embedded within EvalNet participant policies and practices. In some cases, these are supplemented by other standards or guidance to address specific needs. However, the use of these standards varies across central and decentralised evaluations, depending on the management of evaluation processes and quality assurance systems of each organisation.

9. As described in last years' Evaluation Systems in Development Co-operation (OECD, 2023) report and confirmed in interviews of members, centrally managed, strategic evaluations (the type generally managed by EvalNet participants) tend to be found to be of higher quality, while decentralised evaluations – usually managed by project or operational units – tend to have lower quality with weaker methodological approaches, less robust conclusions and gaps in application of ethical standards. Although these

evaluations are considered useful for the projects they evaluate, their wider utility - at the sectoral, regional or organisational level – is low, due to limited access to reports (no repositories or system for finding), limited support from evaluation units/experts, and limited know how among project staff.

10. During the discussion, participants confirmed the continued relevance and usefulness of the OECD/DAC quality standards. Some members discussed areas where the quality standards and principles could be expanded. Gottfried von Gemmingen (BMZ, Germany) described BMZ's new evaluation policy which explicitly mentions the quality in terms of methods, process, and utilisation of evaluation. He also suggested the importance of institutionalising the quality principles. Helge Østtveiten (Norway) underlined independence as a requirement for quality assurance and encouraged teams to clearly explain how methods and processes are decided.

11. Members agreed on the importance of investing in use, that learning and accountability are interlinked in a dynamic process, and that the term accountability also includes the accountability to learn from and utilise evaluation findings. Julie Chen (United States) described work on creating an organisational culture for an environment conducive to learning to improve evaluation quality. Ivory Yong-Protzel (IADB) challenged whether ensuring learning was the role of evaluators, suggesting that other stakeholders should take the lead. She also suggested that accountability activities be more geared towards development results and achieving better outcomes, to avoid an over emphasis on mechanistic reporting and due diligence (which may not add value). Karen Rot-Munstermann (AfDB) shared the Bank's efforts to strengthen the utilisation of evaluations which include increased consultation with senior management and using expert staff to promote dissemination.

12. Members raised other recurring challenges around ensuring quality evaluations. These included data quality, staffing and budgetary constraints, and communicating evaluation findings in an accessible manner. Members discussed the costs attached to ensuring quality in evaluations and trade-offs between 'perfect' evaluations and 'good enough' evaluations that can be delivered on time. Members that rely on external consultants discussed how they support consultants in producing high quality evaluations which included open lines of communication and dialogue and mutual adherence to methodological standards.

13. Capacity strengthening and the development of tools, templates and guidance were two areas where members are investing to improve evaluation quality. Jean-Claude Pires (AFD) shared how AFD developed an action plan for evaluations in 2018 after agency-wide concern about the poor quality of project evaluations. To enhance evaluation quality, central learning and evaluation processes were implemented and earmarked resources were dedicated to each evaluation project. In-house tools, staff trainings and pooled references on external consultants across agencies strengthened the evaluation function.

14. Members discussed the implications of recent shifts towards more locally-led development. Australia and the United States mentioned that they are instating programmes to strengthen local partners' capacities and improve their access to respond to evaluation bid.

Item 4. Evaluation in, on and for crisis

INCAF-EvalNet Joint Task Team on Afghanistan

15. In 2021, a temporary joint EvalNet–International Network on Conflict and Fragility (INCAF) task team was convened to respond to the DAC's request for both networks to unpack questions around learning (and not learning) from evaluations in conflict and fragile settings. Sarah Spencer-Bernard (OECD/INCAF) updated members on the joint task team's progress exploring links between learning and action in Afghanistan (underway), a working paper on development co-operation in politically constrained contexts and a meta-review of ten years of development co-operation in different fragile contexts (Afghanistan, Mali and South Sudan). Members welcomed the progress and further options for reflecting on these topics will be explored.

Climate Crisis Working Group (CCWG)

16. Sven Harten (DEval, Germany) presented a proposed management structure and priorities for the CCWG, underlined the usefulness of leveraging EvalNet for joint work and peer support, and invited members to identify areas of focus. Members expressed interest in the work of the CCWG, citing shared challenges around measurement, data quality, reporting guidelines, and benchmarking. Lori Bell (UNHCR) informed members that heads of evaluation across the UN system have committed to looking systematically at environmental impact in evaluations and welcomed linking UNEG's Working Group on Climate and EvalNet's CCWG.

Lessons from Evaluation for Ukraine and other fragility and conflict-related work

17. The Secretariat provided an update on work to leverage evidence to support international efforts towards the recovery and reconstruction of Ukraine. Members discussed rising concern regarding the development co-operation challenges faced at the local, regional and national level in the context of the ongoing war and welcomed the Secretariat's paper for the High-Level Meeting (HLM) of the DAC. Participants agreed to reconvene to explore evaluation needs related to Ukraine, including potential joint work, and to further develop the Lessons from Evaluation for Ukraine note (building on the HLM paper).

18. Vivek Arora (IMF) discussed evaluations of IMF operations in Fragile and Conflict-affected States (FCS). The recent escalation of Russia's unprovoked war of aggression in Ukraine has prompted a doubling of the IMF's engagement with FCS. Lessons that have emerged from evaluations include the importance of long-term engagement in fragile states, the need for greater field presence, and flexibility in programme design. A significant mismatch was identified between the needs of FCS and the available financial instruments. For instance, the current structure relies on successive short-term programmes, whereas long-term loans are preferred by partners and are likely to be a better fit. There is tension between maintaining high standards and tailoring conditions to country-specific circumstances. Challenges in conducting evaluations in FCS include security, capacity constraints, and political fragmentation. Collaboration with development partners is important but difficult in such contexts, and attracting skilled staff to work in FCS has been challenging.

19. Toshiya Abe (Japan) shared JICA's experience in crisis contexts, which is challenging and requires a specific lens for timeliness, relevance, coherence and impact. Interventions in such contexts must be coherent with the political situation, reach socially vulnerable populations, and avoid conflict relapse. Keeping accurate project records is crucial, given the dynamic nature of these situations. JICA developed new guidelines for evaluations in Sudan, South Sudan, and the Philippines and found that flexibility in delivering development assistance in FCS is key.

20. Véronique Salze-Lozac'h (EBRD) discussed the role of independent evaluation units during crises and how to remain relevant in situations that demand rapid learning. EBRD uses three approaches to do so: leveraging existing evaluative knowledge, providing feedback on programme results measurement, and establishing rapid feedback and learning loops. The rapid assessment of EBRD's response to the COVID-19 pandemic was timely and useful in shaping the Bank's response in Ukraine. The implications of these lessons – including the need for flexibility, focus on long-term objectives, incorporation of climate change considerations, and knowledge exchange with other institutions – also extend to how EBRD conducts its evaluations.

21. Eleni Kalampoka (EU) affirmed the EU's support for Ukraine. The EU has proposed the Ukraine facility instrument with an allocation of 50 billion euros for loans and guarantees, for which a M&E system with strong Ukrainian ownership has been set-up. The EU's Strategic Evaluation Plan for Ukraine includes a focus on recovery, reconstruction, and modernisation aligned with previous evaluations' findings. The plan acknowledges past issues, such as delayed responses to early warnings. Efforts to improve M&E systems are being implemented, with a shift towards performance-oriented interventions, and integrating

results frameworks into every intervention, which has supported tracking of progress of the interventions delivered in Ukraine.

22. Lori Bell (UNHCR) discussed the challenges and considerations involved in evaluating humanitarian operations. The criteria for humanitarian evaluations are different from those that govern development evaluations and issues such as access and security for evaluation teams, and duty of care for staff and partners demand high attention. Limited access to affected populations and staff turnover pose additional challenges in gathering reliable data. In this context, considering the substantial resources allocated to humanitarian action, establishing a learning focus early in the evaluation processes is key.

23. The Secretariat invited members to provide feedback on and express interest in the proposal for a series of Systematic Reviews on Aid Effectiveness at Country Level in a range of fragile contexts. Members acknowledged the work's potential benefits and raised methodological and conceptual questions. Numerous participants expressed interest.

COVID-19 Global Evaluation Coalition

24. Frank Kirwan (Ireland) updated members on the COVID-19 Global Evaluation Coalition. The coalition is a strong example of effective collaboration and demonstrates the value of collaboration, especially among small to medium-sized evaluation systems where joining forces has allowed for greater achievement. The Coalition has provided timely and credible evaluative evidence during the pandemic, facilitated several events for dialogue and learning, and is now in the critical stage of harvesting lessons from the collective response to the pandemic through the Strategic Joint Evaluation.

25. Megan Kennedy-Chouane (EvalNet Secretariat) updated members on the *Strategic Joint Evaluation of the Collective International Development and Humanitarian Assistance Response to the COVID-19 Pandemic*. The evaluation aims to answer questions around the relevance, coherence, efficiency, and effectiveness of the pandemic response overall. Module 1 of the evaluation, a synthesis of completed evaluations is available, and Modules 2-4 are underway. These include case studies of development co-operation providers and partner country cases, as well as a study of the philanthropic response. The Secretariat is fundraising for this evaluation, as well as to support country-led evaluations in Burkina Faso and Ghana, and members are encouraged to contribute. The Secretariat is also seeking a consultant to lead the case study for Mozambique and the SIDS case studies.

26. Daniel Jacobo Orea (Spain) updated members on the evaluation of Spain's Joint Strategy For COVID-19 Response and Vaccine Donation, which is also a case study for the Strategic Joint Evaluation having been inspired by the work of the COVID-19 Global Evaluation Coalition. The evaluation explores the strategy's effectiveness in addressing partner country needs and examines the impact of a 20% increase in ODA in 2021. Case studies in Peru, Bolivia, and Ethiopia provide further insights. The evaluation is scheduled to be presented to the Spanish Cooperation Council and Monitoring Committee.

27. The Chair closed the session by highlighting that joint efforts across the evaluation community are vital in answering questions around collective failure in conflict settings and how to prevent it.

Outcomes and next steps

- The Secretariat will circulate a draft Lessons from Evaluation for Ukraine piece, and members will be invited to validate it and add examples.
- The Climate Working Group will propose a work plan.
- The Secretariat will invite those members that expressed interest in the proposal for fragility meta-syntheses for further discussion at a virtual meeting.

Item 5. Development evaluation: The central role of partner countries

28. The Chair kicked off the session by welcoming in-person and online guests from Indonesia's Ministry of National Development Planning, Cambodia's Council for Development, Colombia's National Planning Department, Indian Development Monitoring and Evaluation Office, Malawi's Ministry of Finance, Economic Planning and Development, Thailand's International Cooperation Agency, the Commission for Monitoring and Evaluation Office of the Prime Minister of Uganda, and CLEAR Anglophone Africa (AA).

29. Development co-operation evaluation systems are well established compared to other policy fields, with well established, common norms, and standards. The discourse on partner engagement and inclusion in evaluation processes has been a feature at EvalNet since its founding forty years ago. The Paris Agenda on Aid Effectiveness in 2005 called for strengthening ownership of partner countries and mutual accountability, including in evaluation, but progress has been slow and fragmented. There is increasing demand from partner countries for a more prominent role in evaluation and more partner-orientated approaches.

30. The Chair introduced the panel of experts: Candice Morkel (CLEAR Anglophone Africa (AA)), Peter van der Knaap (IOB, the Netherlands), Isabelle Mercier (Independent Evaluation Office, UNDP) and Timothy Lubanga (Monitoring and Evaluation Office of the Prime Minister, Uganda).

31. Candice Morkel (CLEAR-AA) expressed concern over the slow progress in achieving country ownership of evaluation systems despite commitments on alignment and mutual accountability made under the Paris Declaration, the Accra Agenda, and the Busan Agreement. The involvement of partner country governments and the broader ecosystem in which they operate, including citizens and parliaments, is critical to strengthening country ownership. She challenged the notion that donors hold greater accountability rights than partner countries. Responding to calls to decolonise development systems, partner involvement can help rebalance asymmetrical power relations between donors and recipients and build trust-based collaborative partnerships. Partner governments and donors should engage in joint planning to assess institutional capacities and strengthen national evaluation systems. There are examples of effective partnerships, such as the one between CLEAR AA and UNICEF, that have successfully increased national evaluation capacities. Ms. Morkel closed by urging DAC members to adhere to the key effectiveness principles of inclusive partnerships, transparency, and mutual accountability, as they operationalise partner involvement in evaluation processes.

32. Peter van der Knaap (the Netherlands) noted the importance of partner involvement for the legitimacy and effectiveness of development co-operation. A partner-oriented approach increases the relevance of policy interventions by ensuring that they are responsive to needs and interests of partner countries and their citizens. Trust between actors and collaboration with partners has been proven to drive effectiveness. When considering which potential partners to involve, donors should engage in trust-based dialogue to clarify and understand expectations around the interventions and evaluations. He cautioned against engaging with actors with interests that did not align with the interests of beneficiaries and reminded members of the importance of balancing demands related to accountability mechanisms to maintain public trust in development spending. Investments in building evaluation culture and capacity can contribute to strengthening good governance. Though evaluation capacity strengthening carries additional costs, other innovations such as new tools for data collection could offset some of them. Mr. van der Knaap closed by suggesting strengthening national policies should be the ultimate development objective, which would lead to more effective and sustainable outcomes in the long run.

33. Isabelle Mercier (UNDP) explained that UNDP's country programme evaluations are conducted in partnership with partner countries and funded by both donors and partner countries. Partner governments are involved throughout the evaluation process, including reviewing ToRs and providing feedback on draft reports. However, UNDP's decentralised and thematic evaluations have limited engagement with government officials and beneficiaries. The lack of entry points for the views and voices of those not

already part of the system is a key area of concern. Such evaluations could be strengthened by giving greater consideration to the viewpoints of beneficiaries, orienting more to the formats and approaches most useful to partner governments, and reflecting on the ethical importance of including partners in evaluations. The evaluation community at large should systematically consider the ethical implications of what is measured and prioritised in evaluation work and for whom evaluation results are the most useful. Partner-orientation also entails strengthening systems in partner countries that can absorb and apply evaluation findings. Targeting evaluation products and results to different stakeholders serves an accountability and a learning purpose that can ultimately bolster national evaluation systems and lead to more impactful development outcomes.

34. Timothy Lubanga (Uganda) also noted the importance of involving partner countries at all stages of the evaluation process based on - their national evaluation systems and capacities. Involving partner countries in identifying evaluation topics ensures that evaluations address questions relevant to achieving national development outcomes, and maximise alignment and harmonisation with national systems , while also strengthening national evaluative capacity. Procuring and managing evaluation teams makes greater use of local expertise, and can develop a pool of skilled professionals capable of evaluating policies and programmes effectively. Mr. Lubanga acknowledged the additional costs that accompany partner-oriented approaches but stressed the crucial role of partner countries in disseminating and utilising evaluation results to inform national policies and promote learning. Donors could strengthen national evaluation systems by increasing demand for evidence and enhancing national capacity to supply evaluations. Donors can also foster a culture of evidence use in partner countries by promoting evidence-based practices across different policy areas. Finally, national evaluation systems need strong evaluators that are equipped with applying innovative evaluative methodologies, including artificial intelligence (AI) and GIS technology.

35. During the discussion, participants strongly agreed regarding the importance of local ownership, country-led evaluations, and engaging partners in planning, designing, and promoting the use of evaluation findings for evidence-based policymaking. There was consensus on the need to foster meaningful partnerships and address frequent challenges such as the safety and independence of local evaluators in fragile contexts, and resource constraints. Members also discussed ways to promote inclusive and culturally sensitive evaluation practices.

36. Karen Rot-Munstermann (AfDB) offered practical suggestions to enhance partner contribution, including inviting representatives to join evaluation reference groups and consulting with central evaluation departments during the scoping phase of evaluations. Astrid Ganterer (Austria) raised the need for budgets to enable equitable language and fostering culturally sensitive communication to facilitate meaningful engagement and contributions from partners. Finally, Julie Chen (USAID) suggested including specific questions of interest to beneficiary countries could enhance the utility and inclusiveness of the evaluation process and results.

37. The session then moved into smaller group discussions on five practical considerations on increasing partner orientation in development evaluation. Highlights reported by each group included:

Differentiating between different potential partners

38. The group acknowledged the diverse range of potential partners for evaluation efforts and honed in on specific concerns for local civil society organisations and women’s networks. These include challenges with creating safe spaces, ensuring safety and navigating alignment issues with government priorities. The structure of such partnerships requires careful considerations, a thorough with stakeholder mapping and ongoing communication throughout the evaluation process. Language considerations and realistic expectations were also recognised as crucial elements of the partnership approach, along with the importance of triangulating between different perspectives for an inclusive evaluation.

Differentiating between different “intensity” of partner orientation in development evaluation

39. The group discussed strategies to intensify partner orientation in evaluation efforts beyond information sharing and coordination. One approach is to promptly respond to partner requests for evaluations, leveraging their engagement and interest. Another tactic involves packaging partner participation into the evaluation process, allowing them to lead or manage certain aspects, potentially formalising this arrangement in contracts. On this point, the group debated how to balance institutional and national evaluation frameworks with flexibility to engage partners more deeply. Finally, for specific types of evaluations like policy evaluations, it is important to engage with a diverse range of stakeholders. This engagement could potentially be facilitated by platforms like EvalNet.

Differentiating partner involvement during different stages of the evaluation cycle

40. The group mapped out the stages of a generic evaluation process, from intervention design to data collection, analysis, and dissemination. Currently, partner country engagement is primarily seen during the data collection phase, and it is not realistic to expect partner countries to engage on every detail. However, increased involvement of partner countries at the design and inception stages to collectively shape the evaluation direction ensures clarity, consensus on evaluation approaches and instils ownership. Partner participation during the analysis phase to workshop emerging findings and recommendations, can enhance evaluation quality. Involvement during the dissemination phase would ensure that evaluation results are effectively and widely communicated.

The potential role of Evaluation Capacity Development in strengthening partner orientation

41. The group discussed the importance of two-way capacity building (where all partners grow and learn together), raising the need for equitable practices within development organisations and the recognition of local evaluators as crucial actors in the local evaluation ecosystems. Clarity on the mandates of evaluation units to strengthen national evaluation capacity is essential. Finally, the group suggested that building networks and fostering connections between individuals and organisations for knowledge exchange, play a significant role in capacity development.

Academic institutions as potential partners in development evaluations.

42. The group discussed the effectiveness and compatibility of academic institutions as potential partners in collaborative evaluation efforts. They highlighted the pilot approach of the European Commission of adopting an intercultural perspective to involve partner countries during the evaluation process, which helps in eliminating biases and blind spots. A significant challenge faced when involving partner organisations in evaluation is around procurement of services. The group suggested refining ToRs at the donor agencies to simplify the processes around partner orientation.

43. During the discussion, panellists and participants deliberated on how politics, both at global and national levels, influence the conduct of evaluations. A political economy analysis is key to ascertaining the feasibility of an evaluation and ensuring the safety of local evaluators and other partners involved in the process, especially in repressive contexts. Evaluations have an important role to play in democratic strengthening but without a thorough understanding of a national context, risks becoming instrumental. To overcome instrumentalisation, discussants emphasised the importance of adaptation, involving partners and beneficiaries for a stronger understanding of context, and improving the utility of evaluations.

44. Members agreed on the need for peer learning on this topic, noting the strong position of EvalNet to facilitate such joint reflections. Participants recognised the importance of having partner country governments at the table and thanked those who participated in the session. Peter van der Knaap (the Netherlands) suggested convening interested members to draft a sourcebook of existing practices to strengthen partner involvement, bringing cases together with conceptual and institutional frameworks around partner engagement in evaluations.

Outcomes and next steps

- The Chair will publish a policy brief sharing key reflections on the role of partner countries in evaluations from this session.
- Members will consider producing a "source book" or similar compilation of lessons and experience with effective engagement.

16 November 2023

Item 6. Anniversary Celebration

45. The Chair opened the second day by welcoming members and special observers to the special session celebrating the 40th anniversary of the DAC Network on Development Evaluation.

46. Carsten Staur (Chair of the DAC) provided opening remarks. The DAC chair noted the instrumental role of EvalNet in anchoring the evaluation function as an integral part of development co-operation and humanitarian assistance. In the current geopolitical context where political bodies such as the DAC and the EU consider enlargement and crises are plenty, independent evaluation functions within government agencies play a crucial twin role of maintaining accountability and advancing learning. The DAC Chair reflected on ODA to Ukraine and emphasised the importance of embedding evidence in policy responses, drawing on evaluations of previous reconstruction efforts. There are recurring challenges around maintaining a focus on evidence-based decision-making during crises, but with increasing pressure on ODA budgets, evaluations help ensure and demonstrate the effectiveness of ODA. The COVID-19 Global Evaluation Coalition's ongoing Strategic Joint Evaluation is another example of where lessons from a global crisis can inform preparedness and maximise the value of ongoing efforts. ODA decisions are political though a shift towards evidence-based practices is evident and EvalNet is well positioned to leverage this momentum. The DAC Chair expressed gratitude for the work of EvalNet and encouraged the network to continue effectively disseminating its work to inform decision-makers and drive impact.

47. Pilar Garrido (OECD), Director of the Development Co-operation Directorate moderated the panel: Sabine Bernabe (World Bank IEG), Candice Morkel (CLEAR AA), Jörg Faust (DEval, Germany), Rachel Ingwersen (Australia), Ivory Yong-Protzel (IADB). Ms. Garrido set the scene by describing the key functions of EvalNet. These include developing norms and standards that help improve the quality of evaluations, acting as a conduit between policy makers and the evaluation community to provide evidence-based advice, supporting partner countries in undertaking more and better evaluations, and support the DAC to further progress along the SDGs. Panellists were invited to respond to the question on if and how organisations learn and the factors that drive learning.

48. Jörg Faust (DEval, Germany) noticed that an increasing number of quantitative studies on aid effectiveness over the past 20 years reveal positive effects of development co-operation. Evaluations have played a role in this trend by supporting organisations learn better and thereby improve their practices. When it comes to the uptake of recommendations, evaluators face constant challenges around feasibility and political prioritisation by decision-makers. Evaluations can foster learning when evaluation agencies are independent and can communicate to varied audiences. Evaluators have access to large evidence repositories and synthesising this evidence is key. However, decision-makers must also invest in time and resources for evaluative work and improve evidence absorption capacity.

49. Sabine Bernabe (World Bank IEG) noted the increasing influence of evaluations on learning. This influence is driven by three key factors. Firstly, credibility increases uptake and is signalled by the independence, quality, and rigor of evaluations. Secondly, engagement with actors along the evaluation process fosters trust and openness to learning, aiding in a better understanding of evaluands. Lastly, diversification of evaluation products and reporting, including user-friendly formats, is crucial to meet

stakeholder needs, such as faster evaluations and real-time reporting. Technological advancements today around AI, text mining and machine learning provide opportunities for evaluators to enhance evaluation processes. However, the influence of evaluations should extend beyond the immediate internal stakeholders. Evaluations play a key role in learning and evaluation capacity strengthening in partner countries too. Finally, focusing on outcomes and impacts rather than outputs is key to learning and meaningful development.

50. Carsten Staur (DAC Chair) reflected on whether DAC members learn from evidence and the degree to which development co-operation is evidence-based. ODA decisions are political, but there has been a noticeable shift towards evidence-based practices, especially in the methods of ODA delivery. A major obstacle to the uptake of evaluative evidence is the user-friendliness of dissemination methods. Another obstacle is the tension between evaluation and public relation branches of aid agencies and an imbalance between the need to be transparent and the discomfort around publishing 'uncomfortable truths.' Overcoming these tensions require maintaining integrity and staying true to evaluation values. Preliminary discussions about potential evaluation outcomes at higher levels can cement commitment to transparency and accountability early in the process.

51. Rachel Ingwersen (Australia) discussed Australia's accountability structures, which focuses on being accountable to its taxpayers and Parliament. She highlighted gaps around accountability to beneficiaries and partner countries. Efforts to include partners in the evaluation process tend to be focussed on the data collection phase, which is sometimes viewed as extractive and not a legitimate form of engagement. Ms. Ingwersen shared Australia's forward-looking approach to embed accountability in a new international development policy focused on deepening consultation with partner countries. This includes strengthened approaches to country and regional planning and a focus on equal dialogue, dissemination of findings, and learning events. The aim is to ultimately strengthen accountability to partners and beneficiaries.

52. Candice Morkel (CLEAR AA) reflected on the evolving discourse around decolonising evaluation and transforming development practices, particularly in Africa. There is an observable trend towards questioning traditional approaches and advocating for indigenising evaluation systems for true partner ownership. However, there lies a real challenge in translating theoretical concepts into practical action within development institutions. These institutions require transformation along many areas such as the accountability and governance structures, procurement processes and legal frameworks. Navigating post-colonial governance structures is complex and calls for a systemic shift in thinking and framing over short-term changes in practice. As broader conversations begin, examining the value-based aspects of evaluation processes is a first step to fostering a more inclusive and culturally sensitive approach. The Chair welcomed these suggestions and noted EvalNet's role in furthering these conversations.

53. Ivory Yong-Protzel (IADB) discussed the launch of a capacity-building initiative in Latin America and the Caribbean, aimed at enhancing evaluation skills among government officials in the region. Delivering these trainings effectively required extensive tailoring to the regional context and recognition of the linguistic diversity in the region. The initiative also fostered networking opportunities, enabling participants to connect with peers facing similar challenges. EvalNet facilitated collaboration and knowledge-sharing, which ultimately supported IADB in strengthening their evaluation efforts in the region. Despite challenges, including the COVID-19 pandemic, the initiative has shown promise in promoting a common evaluation culture and facilitating the exchange of experiences and best practices among stakeholders in the region.

54. During the discussion participants expressed their profound appreciation for EvalNet. They noted the network's instrumental role in setting quality standards and accountability criteria for practitioners globally. EvalNet acts as a convener and provides a platform for peer exchange, learning and inspiration. It has facilitated critical joint reflection on policy priorities such as the COVID-19 pandemic and continues to do so within the Climate Crisis Working Group. In the context of increased calls for systematic

coherence and co-ordination between countries, multilateral and bilateral agencies, EvalNet can act as a co-ordination platform and foster an evidence-driven culture. Co-ordination, knowledge sharing and resource pooling in the evaluation field can yield great benefits, especially during crises.

55. Participants agreed on the important role of actively involving partner countries in evaluation processes, highlighting the necessity for a normative shift in evaluation practices to accommodate this perspective. Wiwien Apriliani (Indonesia) shared Indonesia's experience in establishing evaluation systems, leveraging resources developed by EvalNet and adapting them to their national context. Central to this endeavour was ownership over the evaluation system, facilitated by the ability to contextualise principles to suit local needs. They advocated for the importance of fostering dialogue with partner countries and enhancing their access to knowledge and learning systems.

56. Participants expressed a desire to move beyond the traditional long timelines for evaluations and explore real-time evaluation methods for great policy influence. There was broad agreement that increased synthesis work could effectively utilise existing evaluation data and that advancements in technology could help accelerate the feedback of evidence to policy making. Exploring how AI could streamline the evaluation process was a common area of interest. However, the discussion also explored how political considerations that underpin decision-making require the evaluation community to critically reflect on navigating the complex reality of influence, including barriers that timing and quality alone cannot address. The discussion touched on the importance of anchoring evaluation in integrated policy frameworks to avoid siloed approaches and nuance evaluations to match political realities.

57. Panellists suggested that EvalNet focus future work on building partner country capacities, explore ways to mine and use decades of evaluative evidence, and continue to leverage its unique position to expand partnerships and foster coherence. The discussion closed with consensus among participants to increase demand for evaluations in decision-making and to strive to increase influence. The EvalNet Chair closed the session by encouraging evaluators to decrease bureaucratic burden on decision-makers during crises and expressed gratitude to the panel, the DAC Chair and the DCD Director for their support.

Item 7. Artificial Intelligence (AI) and evaluation: Implications and next steps

58. Mathieu Audet (Canada) highlighted the growing significance of artificial intelligence (AI) in shaping the future of evaluation and its potential to revolutionise data analysis. The member survey on AI revealed that 60% of members do not currently use AI but that 70% are curious and open to trying. Members also expressed valid concerns such as transparency, ethics, and biases in data. Evaluators play an important role in addressing these challenges, given their contextual understanding and research skills. Members already using AI in their work, now focus on how, where, and when it should be integrated into evaluation processes and dissemination efforts.

59. Estelle Raimondo (World Bank IEG) presented on how the World Bank IEG has integrated AI in their evaluation practice including their most recent Experiments with Generative AI for evaluation processes. IEG used AI to work on complex portfolio reviews which otherwise involve labour-intensive manual text and image data classifications; text mining for identifying success and failure factors; sentiment analysis of media dialogue; classifying components of theory of change in thematic evaluations; and creating systematic models that can be (re-)used systematically for country programme evaluations. Though AI helped analyse large datasets and significantly expand the scope of data covered in evaluations, challenges included finite specialised expertise, code reproducibility, and scaling. The Bank's recent experiments with generative AI (using primarily ChatGPT) demonstrated that it could deliver on the promise to increase speed and quality when writing code and describing functions, simple classification, replicating econometric analyses, and summarisation (with human supervision). However, it could not conduct literature reviews or syntheses due to the high risk of false information. IEG will continue using ChatGPT for simpler tasks for now, and invest in training for prompt engineering and output verification.

60. Isabelle Mercier (UNDP) gave a live demonstration of UNDP's AI-powered evaluation database AIDA. The tool is designed to search and find insights from UNDP's evaluation database, which contains over 6,300 evaluation reports produced by UNDP. Users can search for specific topics, such as governance or effectiveness, and the tool generates summaries based on clusters of insights on the chosen topic. AIDA involves a sequence of five steps: text extraction, AI-driven labelling, human feedback, clustering and summarisation, and sentiment analysis. The tool extracts text from the evaluation reports and then AI algorithms label the extracted text, enabling the tool to categorise and analyse the information. Human evaluators provide feedback to train the AI system to recognise codes and language patterns accurately. AIDA clusters insights and generates summaries of the findings. Finally, it analyses the sentiment of the insights, indicating whether the information is positive or negative. AIDA utilises five distinct models, developed in partnership with Amazon, to process and analyse the data. It incorporates Chat GPT for text generation but modifies it to minimise errors. It operates as a closed system, focusing exclusively on UNDP's evaluation reports (for now). AIDA is a valuable starting point for research and inquiry but not an endpoint. The effectiveness and utility of AIDA are still being explored, and further experimentation and training are needed to determine its full potential and integration into evaluative processes.

61. Javier Mata Fabra (Norway) presented Norad's experiences with using AI to enhance their evaluation processes and improve efficiency. Starting in 2017, Norad focused on using AI to analyse big data and run Twitter analytics. Over the years, Norad has progressed to more advanced AI applications, particularly in natural language processing. In 2021, they undertook AI-driven projects, including analysing textual data and structured data for project classification facilitated richer insights with fewer resources compared to their earlier efforts. Current work includes retrieving archival documents through automated classification and storage and a rapid literature review of corruption measures in development aid to Ukraine. Norad's approach involves combining manual processes with automation, particularly in text mining and synthesis. Unsupervised classification methods, topic modelling, and human feedback loops ensure accuracy. Though AI saves costs and time, human oversight is essential nonetheless. can help increase external accessibility by transforming reports into audio and video formats and improving translation capabilities. There is potential in tapping into existing audio and video sources for data analysis and synthesis, paving the way for cost-effective and efficient knowledge extraction. Norad is considering institutionalising the use of AI beyond specific evaluations. AI

62. Jochen Kluve (KfW) demonstrated how KfW's Interactive Database Evaluation and Learning (IDEaL) tool functions. The tool aims to provide easy access to knowledge and evidence generated from evaluations, allowing users to make informed decisions based on reliable data. The tool aggregates data from over 1,100 project evaluations conducted since 2007, systematically coded according to OECD evaluation criteria. This ensures standardised evaluation results across sectors, regions, and timeframes. The tool provides detailed information on project indicators, allowing users to assess the achievement of project goals and outcomes. Users can start with a brief overview and delve deeper into specific reports as needed, transitioning from summary information to full report access. The tool is designed to work with conventional AI models, leveraging the structured data to perform analysis and generate insights. It combines data from multiple sources, enabling users to analyse factors influencing project success or failure and test AI models for accuracy. From the initial investment in coding and IT capacity building, the tool provides significant added value by centralising evaluation data and streamlining access.

63. Maya Vijayraghavan (ADB) shared that ADB used strong AI models to gather, identify, summarise and rank lessons from their evaluation database for operations. An ADB-wide version of the tool is expected to be ready in about a year.

64. During the discussion participants shared other experiences with using AI in their work, noting the role of AI as a starting point for evaluations and agreeing with the speakers about the requirements for regular human oversight and verification. The Secretariat reminded members of EvalNet's evidence repository DEReC, and ongoing discussions on how to modernise it.

65. The group agreed on the potential efficiency benefits of AI application while also flagging issues of ethics and data protection, transparency, accountability, sustainability, and user-friendliness. Members discussed the value of thinking collectively on this topic to overcome upfront investments (and avoid duplication). Joint work could include reflection on AI guidelines and consolidating evaluation databases for global knowledge mobilisation.

Item 8: Capturing unintended effects

66. The Chair welcomed Dirk-Jan Koch, Chief Science Officer at the Netherlands Ministry of Foreign Affairs who presented key messages from his book on Foreign Aid and its Unintended Consequences.

67. Focusing on unintended effects is crucial for preventing harm and addressing dilemmas, such as the unintended consequences of tourism on local communities. In line with EvalNet guidance (including the 2021 Applying Criteria Thoughtfully), the speaker suggested that evaluators should look beyond target groups and consider longer timeframes to see impacts. Identifying unintended effects can make discussions about aid effectiveness more honest and inclusive, but there are barriers to learning, and research – including using innovative methods to analyse coverage of unintended effects in evaluations – shows that they are often not well covered. A meta-evaluation conducted in 2014 found that only about 15% of USAID evaluations included unintended effects, despite guidelines stipulating their importance. Recent research shows similar trends in Dutch evaluations. Mr Koch introduced a typology of ten unintended effects: backlash effects, conflict effects, migration effects, price effects, marginalisation effects, behavioural effects, government effects, negative spillover effects, environmental safety and side effects, and ripple effects. He provided examples for each and discussed common challenges in evaluating them, including ideological biases, institutional barriers, and technological limitations in measurement.

68. Rahul Malhotra (OECD) linked the politics of evidence use in development co-operation and the discussions around unintended consequences of aid. Well-constructed theories of change are a first step in capturing the complexity of (un)intended effects and assumptions underpinning these models require regular evidence-based monitoring. However, such thorough investigations can be resource intensive.

69. During the discussion participants spoke about the importance of local and indigenous knowledge in identifying unintended consequences. Investing in local research capacity has several benefits, an important one being their ability to create safe spaces for respondents to have more honest conversations. Evaluators should combine hard and soft skills, and leave room for exploratory interviews (that can surface unexpected effects when conducted with tact and integrity). Lori Bell (UNCHR) shared the importance of asking the right questions, to the right people and in the right ways. She stressed the value of open-ended qualitative inquiry in revealing unexpected findings. The Chair highlighted the importance of examining the landscape of development co-operation to uncover macro level unintended effects that may not be readily evident in individual project evaluations. Participants expressed interest in a two-pager summarising the key advice from the book for evaluators.

Item 9: Closing session and AOB [continued: Member updates, proposals, and other business]

70. The Chair commenced the closing session by highlighting the ongoing joint efforts within the Climate Change Working Group and the Global Evaluation Initiative.

71. In other business, members shared the following and invited others to join if interested:

- Astrid Ganterer (Austria) informed members that Austria is conducting a review of existing guidance on country programme evaluations and is planning to draft new guidance next year. Other planned work includes a synthesis of country programme evaluations, an inter-ministerial evaluability assessment on the contribution to climate financing and strategic evaluations on food security.
- Tom Hennessy (Ireland) indicated that Ireland is also interested in reviewing their country programme evaluation approach and seeks collaboration within the Network.

- Ivory Yong-Protzel (IADB) introduced the EVALAC evaluation capacity development initiative in the Latin America and the Caribbean region and invited EvalNet members to participate.
- Julie Chen (USAID) informed members of USAID's focus on promoting locally led development and supporting local researchers and evaluation professionals. Building an evidence base for locally led initiatives is a key priority. USAID seeks collaboration with other development partners to strengthen connections, increase access to evaluation services, and promote the use of locally generated evidence by policymakers and evaluators.
- Adam Trowbridge (United States) informed members of plans to develop sector-specific syntheses of cost-effectiveness data, inspired by similar efforts in the UK and invited members to engage.
- Daniel Jacobo Orea (Spain) and Frank Kirwan (Ireland) shared Spain's and Ireland's engagement in the SDG Synthesis Coalition highlighting the need for increased dialogue and horizontal relationships between partners. They expressed continued support for the next phase of the initiative and encouraged other members to consider participating as well, while also looking for potential adjustments to the scope and approach.
- Megan Kennedy-Chouane (EvalNet Secretariat) shared two new resources: Second Edition of the Glossary of Key Terms in Evaluation and Results-based Management for Sustainable Development, and a Development Co-operation Tools and Practices (TIPs) page on Evaluating Development Co-operation. The Secretariat will circulate a job description for the Policy Analyst position in the team and welcomes potential staff on loan from members.
- Felix Zimmermann (OECD) informed members that the OECD Development Communications Network are interested in an informal exchange with EvalNet on how to communicate about development evaluations and how to evaluate communications work. More information will follow.

Outcomes and next steps

- Several members have raised questions about the organisation of country programme evaluations, the Secretariat will convene interested members to discuss.
- The Evaluation Systems Study is complete, and profiles will be finalised ahead of full publication.
- The Secretariat will undertake additional research to advance the questions on resourcing evaluation (discussed at the previous EvalNet meeting).

72. The Chair reminded members and observers that the next EvalNet meeting will take place on 18 and 19 June 2024. He thanked participants for their active engagement and the Secretariat for organising and closed the meeting.