



Partner-led evaluation for policy research programmes

A thought piece on the KNOWFOR programme evaluation

Tiina Pasanen, Stuart Raetz, John Young and Jess Dart

January 2018

Key messages

- In this paper, a 'partner-led evaluation' is where the implementing partners play a leading role in the evaluation design, data collection, analysis and reporting. An evaluation facilitator coordinates and supports activities, and a quality assurer helps to increase rigour, validity and quality.
- This evaluation approach is most appropriate when there is a strong focus on internal learning and capacity-building. It is more likely to increase partners' sense of ownership and, in turn, uptake of results.
- While the approach may be more prone to biases, such as assessing one's own work more favourably, it is possible to introduce measures to mitigate these potential biases and increase impartiality.



Clear Horizon



Overseas Development Institute

203 Blackfriars Road
London SE1 8NJ

Tel: +44 (0) 20 7922 0300
Fax: +44 (0) 20 7922 0399
Email: info@odi.org.uk

www.odi.org
www.odi.org/facebook
www.odi.org/twitter

Readers are encouraged to reproduce material from ODI publications for their own outputs, as long as they are not being sold commercially. As copyright holder, ODI requests due acknowledgement and a copy of the publication. For online use, we ask readers to link to the original resource on the ODI website. The views presented in this paper are those of the author(s) and do not necessarily represent the views of ODI.

© Overseas Development Institute 2018. This work is licensed under a Creative Commons Attribution-NonCommercial Licence (CC BY-NC 4.0).

Acknowledgements

The authors would like to thank the KNOWFOR partners, especially Bethany Davies from CIFOR, Julien Colomer from IUCN and Nalin M. Kishor from PROFOR who all provided invaluable feedback and insights, and showed exceptional enthusiasm to reflect and learn during the KNOWFOR evaluation, and during the write-up of this thought piece. We would also like to thank DFID and especially Gaia Allison for supporting the writing process.

The authors are grateful to the peer reviewers for their invaluable feedback and suggestions on an earlier draft of this paper, Simon Hearn, ODI Research Associate and independent consultant and Tom Aston, Governance Advisor at CARE International UK.

The opinions expressed and the arguments employed in this report do not necessarily reflect those of the funders and peer reviewers, nor should they be attributed to them.

Contents

Acknowledgements	3
List of boxes, figures and tables	5
Acronyms	6
Executive summary	7
1. Introduction	8
1.1. About the KNOWFOR programme	8
1.2. About the KNOWFOR evaluation	9
2. What is a partner-led evaluation and why might you choose it?	11
2.1. How partner-led evaluation is different from a traditional approach	11
2.2. The KNOWFOR case study	11
2.3. Assessing the partner-led evaluation	12
2.4. When and why partner-led evaluation might be useful	13
3. Who leads the process, and who does the work?	16
3.1. How this approach compares to others	16
3.2. Roles and responsibilities in KNOWFOR evaluation	18
4. What about credibility and independence?	20
4.1. Lack of independence	20
4.2. Biases related to case studies and how they are selected	21
4.3. KNOWFOR strategies to increase impartiality and mitigate biases	22
5. Conclusions and recommendations	24
References	25

List of boxes, figures and tables

Boxes

Box 1. DMEL working group in KNOWFOR	9
Box 2. What do we mean by theory-based evaluation and why is it useful for research and knowledge programmes?	10
Box 3. Categorisation of evaluation types	12
Box 4. Partner-led evaluation as a sub-group of joint evaluations	13
Box 5. Examples of how independence is discussed in evaluation manuals	20

Figures

Figure 1. Partner-led evaluation: KNOWFOR process and timeline	15
---	-----------

Tables

Table 1. Matrix of roles in key evaluation stages by evaluation type	17
Table 2. Roles, responsibilities and estimated days in the KNOWFOR evaluation	18
Table 3. Strategies to address biases and increase impartiality applied in different evaluation stages	23

Acronyms

CH	Clear Horizon Consulting
CIFOR	The Center for International Forestry Research
COR	Collaborative Outcomes Reporting
DFID	The UK Government Department for International Development
DMEL	Design, monitoring, evaluation and learning
ICAI	Independent Commission for Aid Impact
IUCN	The International Union for Conservation of Nature
M&E	Monitoring and evaluation
MLLE	Multiple Lines and Levels of Evidence
ODI	Overseas Development Institute
PROFOR	The World Bank Programme of Forests
QA	Quality assurance
ToC	Theory of Change
ToR	Terms of reference

Executive summary

This paper discusses the strengths, challenges and appropriateness of a partner-led evaluation approach, drawing on the experience of the UK Government Department for International Development (DFID)-funded International Forestry Knowledge programme, 'KNOWFOR'.

What is a partner-led evaluation?

In this paper, partner-led evaluation refers to a type of evaluation used by KNOWFOR where partners responsible for implementing the programme had a leading role at each stage of its evaluation: design, data collection, analysis and partner-level reporting. There was an evaluation facilitator to help coordinate activities and provide ongoing support and advice, and an external quality assurer to increase rigour, validity and quality of products.

If we compare this type of evaluation to other typical evaluations, we take the view that it belongs between independently led participatory evaluations and internal self-evaluations, including elements from both types.

When is it appropriate or useful?

While it is impossible to make sweeping generalisations based on one example, combining our experiences with existing literature indicates that partner-led evaluations are most appropriate when there is a strong focus on internal learning and capacity-building. Partner-led evaluations are likely to increase the partners' ownership of the evaluation process and findings, which in turn has potential to increase the use of evaluation findings and recommendations within the organisation. As partner-led evaluations may be process heavy, they are less appropriate when evaluation results need to be produced quickly, and/or when the main or only objective of the evaluation is to provide upward accountability for a commissioner or donor.

In KNOWFOR, partner-led evaluation was chosen for three main reasons:

1. To build on and maintain strong partner ownership of monitoring and evaluation (M&E) activities
2. To further embed organisational learning on knowledge uptake evaluation
3. To enhance partners' abilities to learn from the evaluation.

Roles and responsibilities

In partner-led evaluations, implementing partners are not just evaluation informants, but joint decision-makers and doers. However, unlike in typical internally led self-evaluations, the evaluation facilitator and quality assurer play more substantial roles.

Independence and credibility

Partner-led evaluations are geared towards learning. However, demands for accountability and rigour cannot be ignored either. Independence in evaluation is seen as a key mechanism in reducing biases (such as seeing one's work in a more positive light or only cherry picking the most positive cases or units of analysis) and increasing the credibility of the process and findings. While it is not possible for partner-led evaluations to be institutionally independent, it is possible to increase the impartiality by introducing measures to mitigate potential biases. The first step is to acknowledge the limitations and biases and then to put a clear strategy to address them in place, thereby increasing robustness and credibility.

In this paper, we specifically discuss biases related to case studies and how they are selected. Overall, the KNOWFOR evaluation team made considerable efforts to increase the impartiality and robustness of the approach. As a starting point, the evaluation methodology used multiple sources of evidence. More specific strategies included:

1. incorporating independent elements to the evaluation process such as the external quality assurance (QA) function
2. utilising existing internal QA systems
3. ensuring transparency and consistent standards across evaluation processes and products by introducing a set of tools and frameworks to be used across projects and organisations and carefully documenting the methods and process.

While some of these worked well, some had challenges.

1. Introduction

KNOWFOR



UK supported partnership between CIFOR, IUCN and PROFOR, which ran from 2012–2017.



The mission was to equip developing country policy-makers and practitioners with forestry knowledge.



The evaluation's main aim was to assess the contribution of partner activities to intended programme outcomes.



It used a 'theory-based' evaluation methodology combined with a 'partner-led' evaluation approach.

This paper discusses the strengths, challenges and appropriateness of a partner-led evaluation approach, drawing on the experience of the Department for International Development (DFID)-funded KNOWFOR International Forestry Knowledge programme.

The paper has five short chapters:

1. The initial chapter introduces the KNOWFOR programme and its evaluation.
2. The second chapter presents the partner-led approach, discussing when and why you might choose this non-traditional approach, why it was chosen for the KNOWFOR programme and how it worked out.
3. This looks at one of the most common questions asked by people wondering if a partner-led approach is right for their programme: who is in charge and who does the work?
4. This section considers impartiality – is the partner-led approach credible? Here we look at some of the common challenges and biases especially relating to non-independently led evaluations.
5. We finish by providing some recommendations for future evaluations looking to use a similar approach.

1.1. About the KNOWFOR programme

KNOWFOR was a £38 million forestry research and knowledge programme operating between 2012 and 2017. It was funded by the UK Government, and it involved a partnership between DFID and three implementing partners: the Center for International Forestry Research (CIFOR),¹ the International Union for Conservation of Nature (IUCN)² and The World Bank Programme of Forests (PROFOR),³ hereafter referred to as 'partners'. These KNOWFOR partners represented very different types of organisations working in the forestry sector: CIFOR is a research for development organisation, IUCN is a global membership union composed of both government and civil society organisations and PROFOR is a quasi-independent programme within a Multilateral bank. The underlying idea was to draw on each organisation's experience, strengths and networks to further increase the uptake and reach of forest knowledge and research supported by the KNOWFOR programme.

KNOWFOR aimed to improve the way knowledge of forests can be understood and used internationally, by addressing the disconnect between the supply and uptake of knowledge by practitioners and decision-makers in the forestry sector. KNOWFOR was designed to increase the interaction of policy-makers and forestry practitioners with relevant research and other knowledge products, tools and capacity-building

1 <https://www.cifor.org/>

2 <https://www.iucn.org/>

3 <https://www.profor.info/>

Box 1. DMEL working group in KNOWFOR

The KNOWFOR DMEL working group was formed after the first DFID annual review in 2013. Though partners said there were no major reasons why the design, monitoring, evaluation and learning (DMEL) concept was chosen instead of the more commonly used monitoring and evaluation (M&E) approach, emphasising the design and learning functions were viewed to be important for partners.

Supported by DFID, partners worked together with Clear Horizon Consulting (CH) to redesign KNOWFOR's programme approach and DMEL strategy in a highly participatory manner. The working

group developed a new programme Theory of Change (ToC) and M&E framework, and also co-designed a suite of monitoring tools (including a Strength of Evidence Tool and Uptake Rubric) that focused on the programme's sphere of influence. DMEL focal points also ensured that sub-projects or funded activities aligned with the overall KNOWFOR programme design.

Though not explicitly articulated, a considerable part of KNOWFOR DMEL work was about developing ongoing, outcome-level evaluative thinking. For example, partners led the internal aggregation of monitoring data and assessed and reflected progress against rubrics annually.

activities. It was expected to do this through improved planning for knowledge use and more deliberate learning and reflection. The varied range of activities were undertaken at the global, regional, national and sub-national levels. KNOWFOR took a decentralised approach to funding, where partners determined the priorities for delivery, in consultation with DFID. This took place under the overall theme of improving the connection between research and practice, action and policy in the forestry and land-use sectors.⁴

1.2. About the KNOWFOR evaluation

The KNOWFOR evaluation aimed to assess the contribution of partner activities to intended programme outcomes and equipping decision-makers. The working definition of 'equipped' used in the KNOWFOR programme by partners recognised that policy-makers and practitioners need access to high-quality, evidence-based knowledge and information, but that decision-making and action are also driven by ideology, influence and the institutional context. KNOWFOR partners had primary mandates for knowledge creation and knowledge translation. In this work, they recognised the need to pay deliberate and explicit attention to the context in order to encourage and support the use of knowledge for better environmental and social outcomes. The KNOWFOR evaluation had a smaller emphasis on the programme approach and whether and how KNOWFOR had influenced partner business practices. The evaluation had three key evaluation questions (which were supported by a number of sub-questions):

1. How and to what extent did KNOWFOR contribute to equipping decision-makers and intermediaries? What lessons can be drawn from KNOWFOR's approach to translating knowledge for action?
2. What influence has KNOWFOR had on how partners undertake their core business and how enduring these changes are likely to be?

3. What were KNOWFOR's programme approach and management processes and were these effective, appropriate and relevant?

The KNOWFOR evaluation included several actors. Firstly, there were DMEL focal points from the partner organisations who had been a part of the KNOWFOR DMEL working group/community of practice from 2014 onwards (see Box 1 for more information). Having had key roles in developing a coherent programmatic design for KNOWFOR and ensuring that KNOWFOR DMEL frameworks complimented and built on partner organisations systems, these partner organisations were heavily involved in the KNOWFOR evaluation design and also took a lead role in managing and coordinating data collation, analysis and reporting. This process was supported and facilitated by Clear Horizon Consulting (hereafter referred to as the evaluation facilitator), who was contracted as a DMEL partner for the programme from 2014 onwards. The approach was complemented by an independent QA process, led by the Overseas Development Institute (ODI), to provide additional rigour and transparency (hereafter referred to as the quality assurer). These roles and what they entailed in plans and in practice are discussed in more detail in Chapter 3.

The primary evaluation audiences were KNOWFOR partner organisations, DFID and the UK International Climate Fund (the cross-departmental mechanism through which the UK channels climate change finance). It was also expected that the evaluation process and its findings may also be of interest to an emerging DFID-supported community of practice bringing together DFID-funded knowledge-for-development programmes as well as the wider evaluation community. This, however, has been a secondary audience and might not be relevant for all evaluations.

The KNOWFOR evaluation was a theory-based, partner-led evaluation. While theory-based refers to **ways the data is analysed and interpreted**, partner-led refers to **the way the**

⁴ More information about KNOWFOR can be found at: <https://devtracker.dfid.gov.uk/projects/GB-1-203034>

evaluation is conducted: who makes decisions, who is in charge and who does the work.

While this paper focuses on the partner-led aspect, understanding the data analysis and interpretation approach that was used in the KNOWFOR evaluation is important too. This drew heavily on Collaborative Outcomes Reporting (COR) (Dart and Roberts, 2014), which belongs under the wide umbrella of theory-based evaluations. COR is a participatory approach to outcomes reporting developed by Jess Dart from CH. It is based around a performance story that presents evidence of how a programme has contributed to outcomes and impacts that is then reviewed by both technical experts and programme stakeholders. COR combines contribution analysis (BetterEvaluation, 2017a) and Multiple Lines and Levels of Evidence (MLLE) (BetterEvaluation, 2017b), mapping existing data and additional figures against the ToC to produce a performance story. Performance story reports are essentially narratives about how a programme has contributed to outcomes. Although they may vary in content and format, most are short, mention the programme's context and aims, relate to a plausible results chain and are backed by empirical evidence (Dart and Mayne, 2005: 307-309). The Results Chart is a particular feature of COR summarising the key claims, along with the evidence supporting each one and assessing the strength of this evidence for each claim.

COR usually includes a verification step by way of an expert panel and stakeholder 'summit' workshop, to check for the credibility of the evidence on what impacts have occurred and the extent to which these can be credibly attributed to the intervention. The KNOWFOR evaluation used the summit workshop but not the expert panel,

Box 2. What do we mean by theory-based evaluation and why is it useful for research and knowledge programmes?

The KNOWFOR evaluation drew heavily from the COR approach, which belongs under the wide umbrella of theory-based evaluations. Theory-based approach refers to how data is analysed and interpreted. Theory-based evaluations have a programme theory, ToC or series of hypotheses on how research is expected to lead to outcomes. These theories or hypotheses will then be tested in the evaluations by collecting evidence to either confirm or reject them. The use of theory-based evaluations in assessing the outcomes of policy research programmes seems to be increasing, though how explicit theories are varies a lot.

There are a number of benefits to using a theory-based approach, mainly that this can provide an analytical framework to investigate the impact of policy research programmes beyond the number of outputs produced and the immediate uptake of those outputs (e.g. report downloads). Furthermore, as identified by Belcher et al. (2017), theory-based approaches can also facilitate learning at the project or programme scale, and provide generalisable lessons about how research can contribute to outcomes and how to design research to be more effective.

because this was partly fulfilled by the independent QA role. The expert panel is usually composed of content knowledge experts.

2. What is a partner-led evaluation and why might you choose it?

Key messages

- KNOWFOR used a partner-led evaluation, giving implementing partners a leading role at each stage: evaluation design, data collection, analysis and partner-level reporting.
- Compared to other evaluations, we take the view that partner-led evaluation sits between independently led participatory evaluations and internal self-evaluations, including elements from both types, while also being slightly different from them.
- These types of evaluations are most appropriate when there is a strong focus on internal learning and capacity-building. They are less appropriate when results need to be produced quickly, or when the main or only objective of the evaluation is to provide upward accountability to a donor.

2.1. How partner-led evaluation is different from a traditional approach

Partner-led evaluation refers to the ways that evaluation is carried out. It is about who makes decisions, who is accountable for what and who does the work. By ‘partner-led approach’, we refer to an evaluation where partners contribute to the design and take a lead role in managing and coordinating data collation, analysis and reporting.

But where do partner-led evaluations sit in the field and how do they differ from other types, such as externally led evaluations? If we compare them to other typical evaluation types, we see them sitting between independently led participatory evaluations and internal self-evaluations, including elements from both types, while also being slightly different from them. It has to be noted that these characterisations of different evaluation types are simplified in order to highlight key differences. In practice, there can be much variation within each category. In some ways, partner-led evaluations might not be significantly different from heavily participatory evaluations. However, there are small but important differences. For instance, participatory evaluations are typically led by an external evaluator or consultant who is ultimately responsible for the process (design, data collection, analysis and reporting) and is accountable to the evaluation commissioner. Though in the KNOWFOR evaluation there was an evaluation facilitator, those involved were not responsible for partner-level data collection or analysis. In this sense, partner-led evaluations

are closer to self-evaluations because partners are, to a large extent, analysing and interpreting their own work. This brings some questions about the impartiality and credibility of results (which will be addressed in Chapter 4).

Partner-led evaluations can also be seen as a sub-group of ‘joint’ evaluations, to which different donor agencies and/or partners participate, according to Breier (2005). The degree of ‘jointness’ varies ‘depending on the extent to which individual partners cooperate in the evaluation process, merge their evaluation resources and combine their evaluation reporting’ (Breier, 2005: 26). This is discussed in more detail in Box 4.

2.2. The KNOWFOR case study

The partner-led approach was chosen for KNOWFOR for three main reasons:

1. It aimed to build on and maintain strong partner ownership of the programme’s M&E.
2. It was geared towards further embedding of organisational learning with regard to knowledge uptake evaluation.
3. It was intended to enhance partners’ abilities to learn from the evaluation via ownership of evaluation findings (KNOWFOR, 2016).

While DMEL was not heavily featured in the initial KNOWFOR programme design, it became a more prominent feature after the first annual review in 2013. From 2014 onwards, KNOWFOR put considerable

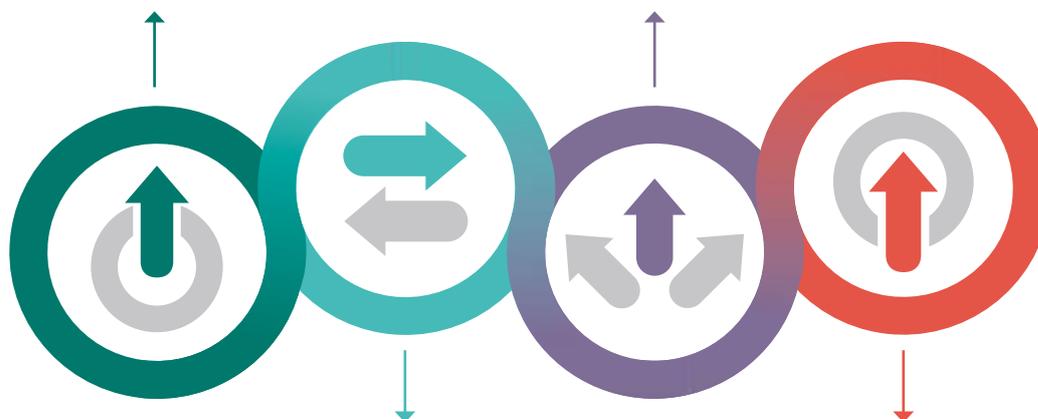
Box 3. Categorisation of evaluation types

EXTERNAL, INDEPENDENT EVALUATION

Conducted by organisations or people who are not part of the design and implementation of the project or directly accountable to those responsible for it.

PARTNER-LED EVALUATION

By partner-led we mean evaluations where the implementing partners are part of the design and take a lead role in managing and coordinating data collation, analysis and reporting.



EXTERNAL BUT PARTICIPATORY EVALUATION

Typically led by an external evaluator, though representatives of implementing organisations and other stakeholders (often including beneficiaries) are involved in design, data collection and analysing the results. The degree of participation can vary.

INTERNAL SELF-EVALUATION

An evaluation carried out by those who are also responsible for the design and delivery of the project.

emphasis on DMEL work which, in fact, was the area where partners collaborated most. This DMEL work included:

- redeveloping the KNOWFOR ToC on how partners' research, knowledge sharing and engagement with stakeholders was expected to influence changes in policy and practice
- introducing and/or supporting ToC thinking into a project design
- supporting the development of partners' monitoring systems.

In the final phase of the project, the idea to try something with a more learning orientation for the required evaluation was supported by everyone, including the DFID Senior Responsible Owner. Partners used several months on discussing the approach and agreed that a partner-led evaluation was well aligned with the DMEL work, with its strong emphasis on learning. Partners also wanted to try something different from a typical external evaluation. KNOWFOR is a complicated and technical programme and partners noted that, based on their experiences in previous programmes, it often took external evaluators a long time to understand the programme and their findings often seemed to add little to what was already known. The rationale

was that a partner-led evaluation could be more efficient and in-depth because it would harness partners' technical knowledge and capability. On the practical side, having an embedded DMEL support from evaluation facilitator who had worked alongside the KNOWFOR partners since 2014 providing ongoing M&E support helped to make this approach possible.

2.3. Assessing the partner-led evaluation

From the beginning of the evaluation, it was recognised that this process would be likely to provide good learning, not just for the partners, but also for the wider audience. Thus, it was decided that an assessment of the partner-led evaluation (referred to hereafter as 'evaluation of the evaluation') would be undertaken. This would capture learning in a systematic manner. During this procedure, people involved in the KNOWFOR evaluation reflected on the evaluation process, along with its strengths, challenges and usefulness. The process was led by ODI, who also provided the QA function for the KNOWFOR evaluation. This included several components to capture a wide set of perspectives and experiences. The process had five parts in total.⁵

5 (1) It started in the reflective session during the summit workshop on March 2017, facilitated by CH.
(2) Based on the above discussion, partners presented their reflections at the DFID seminar on the following day.
(3) After the key evaluation pieces were complete, ODI organised an 'After Action' Review with the core M&E people (which took place in July).
(4) During August and September, ODI interviewed eight senior members of staff in three partner organisations.
(5) Finally, an online survey was conducted in September to capture views from a wider range of actors within partner organisations. The findings from this were shared with partners in an internal 'evaluation of the evaluation report'.

Box 4. Partner-led evaluation as a sub-group of joint evaluations

While the term ‘partner-led evaluation’ is fairly new, many of the ideas behind it are not. For example, ‘joint’ evaluation has been on the development agenda since the early 1990s (Breier, 2005). In 1991, the DAC Principles for Evaluation of Development Assistance stated that ‘joint donor evaluations should be promoted in order to improve understanding of each other’s procedures and approaches to reduce the administrative burden on participants’.

Definitions of ‘partner-led’ and ‘joint’ evaluation have varied. In 2000, the OECD stated that ‘joint’ evaluation involves multiple donors. However, this definition, which originally excluded implementing partners, was subsequently revised by the OECD DAC in 2002 when they described it as involving ‘donor agencies and/or partners’.

Interest in partner-led evaluation peaked in the early 2000s, when there was strong interest and support for the approach in the OECD DAC Evaluation Network. This was led by two consecutive chairs of the network, Niels Dabelstein and Rob van den Berg, who were both strongly supportive of the approach. This convergence

of interest in the network resulted in publications on the practical implementation of partner-led approaches (Binnendijk, 2000) and a review of 53 joint evaluations conducted between 1990 and 2005 (Breier, 2005).

In KNOWFOR, partners shared equal responsibility for the planning and execution of the evaluation. This makes it an example of what the OECD defines as a ‘classic multi-partner evaluation’ (Breier, 2005). It is different from a ‘qualified multi-partner’ evaluation, where participation is not equally shared, because different partners bring different levels of commitment, resources and capacity.

Based purely on the OECD typology, the KNOWFOR evaluation is unique for two reasons. Firstly, a synthesis of 53 evaluations between 1990 and 2005 indicated that only a very small minority took a ‘classic multi-partner’ approach in which resourcing and commitment to the evaluation was shared more or less equally by partners. In KNOWFOR, partners all shared equal responsibility for evaluation design, planning and implementation. Secondly, the OECD 2005 review found that it is also very rare for multi-donor trust-funds to be involved in joint evaluation activities because of the governance of such funding arrangements. In the KNOWFOR evaluation one partner, PROFOR, was governed by a multi-donor trust fund.

The ‘evaluation of the evaluation’ concluded that, overall, the majority of people involved in the evaluation were broadly positive about its approach. Though it was seen as ambitious and time-consuming, people felt it useful due to providing a good learning opportunity for individuals and organisations. Programme staff in partner organisations felt they had built capacity and learnt from each other and the evaluation facilitator. Depending on previous experience and engagement with evaluations, participants learnt different skills and capacities. These ranged from using evaluation tools and practices to understanding the importance of thinking strategically about how research can influence policy and practice. The usefulness of developing and using the ToC was mentioned on several occasions during the process, but participants also considered the importance of (joint) reflection and analysis and whether this was in the form of sense-making workshops or simply a case of ‘being forced to think about impact of our work’.⁶

Furthermore, ownership was perceived to be higher than in a conventional evaluation. This helped to ensure buy-in of the process and findings from management and other stakeholders. Validity and quality of information was perceived to be enhanced because the people collecting it knew the programme and/or project leads.

It was generally agreed that the diversity of KNOWFOR in terms of organisational types and activities taken across the different partners made it difficult to tell a convincing overall story about what worked and why, or to provide

a convincing overall performance rating. The level of analysis in headline findings was a little disappointing to some people involved in the evaluation. We will discuss this tension between focusing on the programme and partner levels in the next chapter.

The majority of people involved in the evaluation would recommend the partner-led approach to colleagues and other organisations, but were less sure about its cost-effectiveness. Some of the organisations were already using elements or tools used in KNOWFOR evaluation in other evaluations.

2.4. When and why partner-led evaluation might be useful

The rationale for and strength of evaluations where partners take a lead role are, in many ways, quite apparent. There has been a lot of discussion and concern about the extent to which the evaluation results are being used or are useful, especially in relation to implementing organisations (see Torres and Preskill, 2001; Patton, 2012). Involving programme staff in the evaluation process and supporting evaluative thinking is likely to increase the ownership of the process and the findings, and expected to reach a wider range of programme staff members. This, in turn, has the potential to increase the uptake and use of evaluation processes and findings within the organisations (Taut, 2007). As Patton (2012: 6) states on utilisation-focused evaluations:

6 Online survey respondent ID24.

Intended users are more likely to use evaluations if they understand and feel ownership of the evaluation process and findings; they are more likely to understand and feel ownership if they've been actively involved.

Furthermore, as recommendations in non-externally led evaluations are based on in-depth knowledge of the programme and context, they can be seen by programme staff and senior management as more reliable and relevant than recommendations coming from external consultants, who often have limited time to spend with the programme and organisation. These strengths also featured heavily in the KNOWFOR evaluation.

While it is impossible to make sweeping generalisations based on one evaluation, combining KNOWFOR experiences with existing evaluation knowledge and literature offers some strong indications on when and why the partner-led approach might be appropriate and feasible.

2.4.1. Evaluation purpose

- Partner-led evaluation is most appropriate when there is a strong focus on internal learning and capacity-building for evaluation. According to OECD DAC (2005), jointly undertaking evaluations is among the most effective tools for evaluation capacity-building
- It is less appropriate when the **main** or **only** objective of the evaluation is to provide upward accountability to a donor or an evaluation commissioner. If the internal learning is not among the key evaluation purposes, partner-led evaluation may come across as a process-heavy approach and the results may not always be seen to be as legitimate and useful for external audiences as they may be for internal audiences.

2.4.2. Trust

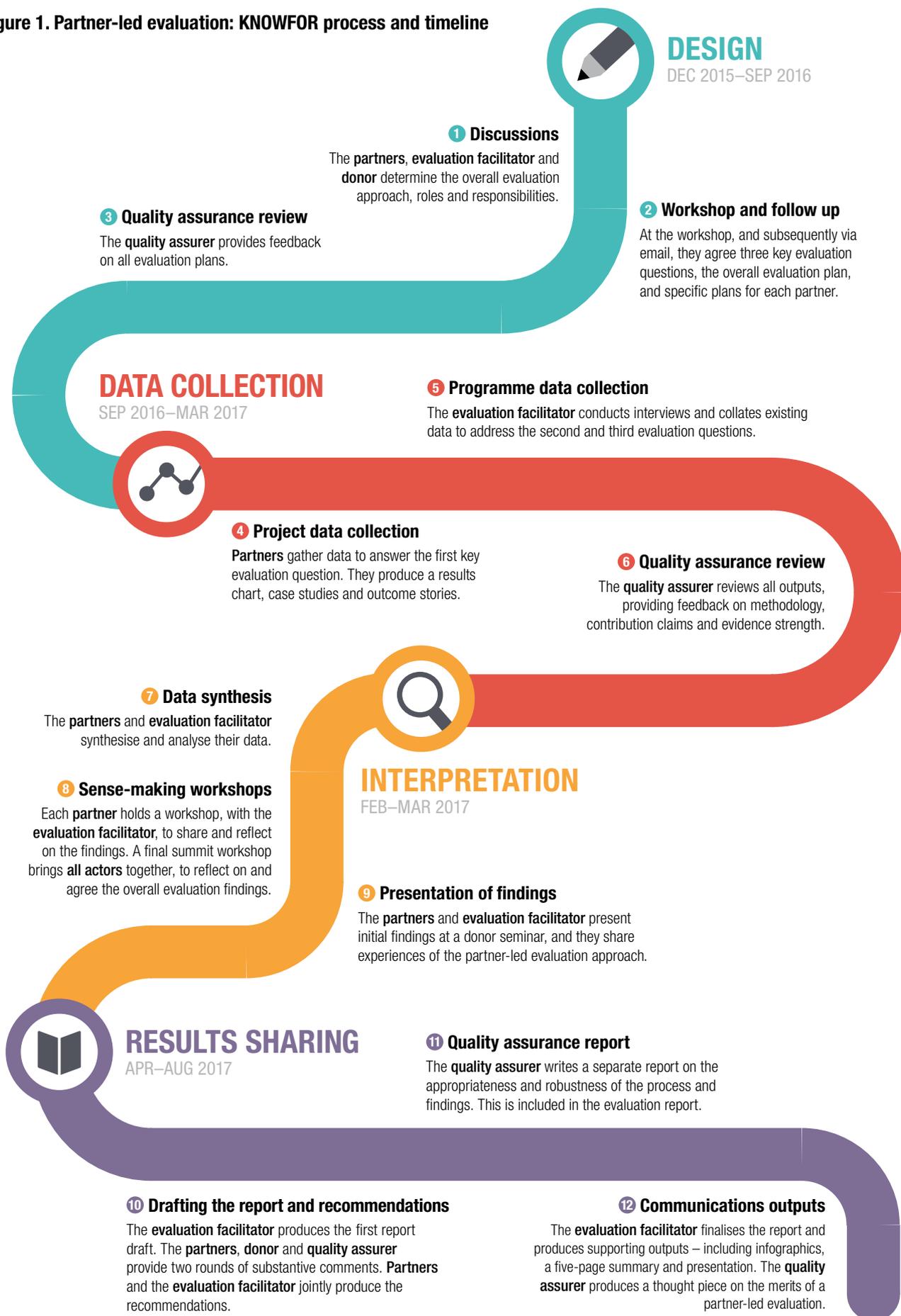
- This approach is more appropriate and feasible to implement in multi-partner programmes when there is trust between partners and they do not see themselves **mainly** or **only** as competitors, which may easily happen if there are not many joint activities or if the partnership is designed or driven by a donor. Competition between partners, though often not unsurprising, can easily distort the evaluation process.

2.4.3. Capacity, resources and timing

- For an organisation to be able to focus on learning and building evaluation capacity, there needs to be a dedicated and preferably flexible form of funding for monitoring, evaluation and learning functions and processes. For example, in KNOWFOR, DFID significantly invested in partners' DMEL work and capacities from 2014 onwards.
- Partner-led evaluation is more appropriate when there is a certain level of **existing M&E capacity and skills** within each organisation. In KNOWFOR, most key people involved already had either solid M&E or research experience. Targeted evaluation skills and methods training can strengthen M&E capacity, but if the existing M&E resources and skills are limited at the start of the evaluation this approach might not be appropriate or present value-for-money for the evaluation commissioner, given partners' responsibilities for data collection and analysis. In this case, an externally led participatory approach might be more appropriate.
- This approach is less appropriate when evaluation results need to be produced quickly. Any type of evaluation done well can take a long time, and it is often difficult to assess and compare the exact time and effort spent on different types of evaluations. However, the partner-led approach places emphasis on the process of evaluation. Thus, it may take longer than independent, non-participatory evaluation (for example) because partners have more say in the process, and coordination and agreement between partners is likely to bring additional time demands.
- While the KNOWFOR evaluation was conducted towards the end of the programme, partner-led evaluation could also be implemented mid-term. In this case, organisations have the opportunity to make use of lessons in ongoing projects. However, with relatively short programmes (e.g. three years or less), the investment may not be proportional to the benefits. However, developing ongoing, outcome-level evaluative thinking, which was part of the embedded DMEL work in KNOWFOR, would also be appropriate for shorter programmes.

Figure 1 provides a simplified visual outline of the partner-led evaluation process and timeline.

Figure 1. Partner-led evaluation: KNOWFOR process and timeline



3. Who leads the process, and who does the work?

Key messages

- In KNOWFOR partner-led evaluation, implementing partners were joint decision-makers and doers. These partners developed the evaluation terms of reference (ToRs) and key questions in collaboration with the evaluation facilitator and donor. Data collection and analysis was shared between the partners (who were responsible for their projects) and the facilitator (who was responsible for programme-wide data).
- The evaluation facilitator was responsible for bringing everyone together and keeping them on track. An external quality assurer provided feedback at each stage.
- There was some tension between project- and programme-level focus, especially in the final report. However, this is not uncommon for any type of evaluation with multiple projects or partners.

3.1. How this approach compares to others

In Table 1 below, we present some broad differences between **who makes the decisions and who does the work** at different stages in different types of evaluations. The degree of participation of the various stakeholders is to some extent an independent dimension, but is frequently related to others in that external evaluations tend to be much less participatory than self-evaluations. It is important to note that these characterisations are simplified and there can be a lot of variation within each evaluation type in practice. In the partner-led column, we specifically refer to the partner-led evaluation carried out in KNOWFOR.

As briefly discussed in Chapter 2, we take the view that partner-led evaluations fit between participatory but externally led evaluations and self-evaluations, combining elements from both of them.⁷ The roles and responsibilities in partner-led evaluations differ mostly from those that are externally led and independent, where an external evaluator or consultant is in charge of data collection, analysis and producing evaluation products.

3.1.1. Partner-led evaluations compared to externally led and participatory evaluations

As Table 1 shows, while partner-led evaluation is in many ways close to typical participatory evaluations, the key difference in the KNOWFOR evaluation is that partners had a much more significant role in all key evaluation stages. These were not just evaluation informants or involved

in some of the evaluation stages, they jointly made key decisions about the approach and were in charge of key evaluation stages to a large extent.

While the KNOWFOR evaluation had an **evaluation facilitator**, their role was to **coordinate and support the process and produce the key written products that reflected agreements between partners** (e.g. evaluation plans, data quality ratings and quality rubrics). They also undertook **synthesis work, bringing all the evidence together, coordinating the programme-level analysis**. In typical participatory evaluations, though partners and other stakeholders (often including beneficiaries) are involved in data collection and analysis, it is the external evaluator who is usually ultimately responsible for these stages and accountable to the donor or evaluation commissioner.

3.1.2. Partner-led evaluations compared to self-evaluations

The KNOWFOR partner-led evaluation is, in many evaluation stages, close to self-evaluation. There are, however some small but significant differences. Though self-evaluations can also involve a facilitator to help during part of the process, or even contract an external peer reviewer to comment on some of the outputs, the role of the facilitator and quality assurer during KNOWFOR was much more substantial. For example, the facilitator was involved throughout to provide support and advise, synthesise evidence and draw up programme-level conclusions. The quality assurer did not only cover the main evaluation report, but also all key outputs and meetings.

⁷ More information on participatory approaches and methods can be found at the Participatory Methods website (<http://www.participatorymethods.org/>) or Guijt (2014) (http://www.participatorymethods.org/sites/participatorymethods.org/files/Participatory_Approaches_ENG%20Irene%20Guijt.pdf).

Table 1. Matrix of simplified roles in key evaluation stages by evaluation type

STAGE OF EVALUATION	TYPE OF EVALUATION			
	External (non-participatory evaluation)	External but participatory	KNOWFOR partner-led evaluation	Self-evaluation
Development of the ToR	<p>Usually <i>set and developed by the commissioner</i>.</p> <p>May be refined by external evaluators after or as part of procurement.</p>	<p>Usually <i>set and developed by commissioner</i>, with consultation with the project team and other stakeholders.</p> <p>May be refined by external evaluators after or as part of procurement.</p>	<p>Set and <i>developed by partners</i> in collaboration with the evaluation facilitator and DFID. The DFID evaluation advisor comments on the evaluation plan and ToR for QA.</p>	<p>Set and <i>developed by the implementing organisation</i> (e.g. the project team, central evaluation unit or senior management).</p>
Agreement of key evaluation questions/ approach	<p>Key evaluation questions usually set by the commissioner in the ToR, but can be <i>refined by a contracted external evaluator</i>. Approach usually developed by external evaluator.</p>	<p>Key evaluation questions and overall approach usually set by the commissioner in the ToR, but often <i>developed by the external evaluator with inputs from stakeholders</i>.</p>	<p>Set and <i>developed by partners</i> in collaboration with the evaluation facilitator and DFID. Reviewed by the quality assurer.</p>	<p>Set and <i>developed by the implementing organisation</i>.</p>
Data collection and analysis	<p>Required data determined by – and <i>most of the work done by – the external evaluator</i>, though this may require data from the project team.</p> <p>Analysis completed by external evaluator.</p>	<p>Data required usually determined by the evaluator, with <i>data collection and analysis shared between external evaluator and stakeholders</i>.</p>	<p>Required data determined <i>by partners and supported by the evaluation facilitator</i>.</p> <p>Partner-level data collection and analysis completed or managed by partners, with the synthesis undertaken and programme-level analysis coordinated by the evaluation facilitator. Reviewed by the quality assurer.</p>	<p>Required data determined – along with <i>all of the data collection and analysis – by the implementing organisation</i> (though an external evaluation advisor and the quality assurer may be contracted).</p>
Production of recommendations	<p>Recommendations produced by external evaluator.</p>	<p>Recommendations produced by external <i>evaluator with involvement of stakeholders, often in workshops</i>.</p>	<p>Approach defined and <i>coordinated by the evaluation facilitator, with inputs from partners in an online workshop</i>.</p>	<p>Recommendations produced by the implementing organisation.</p>
Communication of findings	<p>Key report usually <i>completed by an external evaluator for commissioner</i>. Other outputs and dissemination usually undertaken or managed by the commissioner.</p>	<p>Main report usually for the commissioner, though additional products may be defined by projects for specific audiences. Key <i>products produced by the evaluator, with involvement of stakeholders</i>.</p>	<p>Main report for partners, though additional products are developed for other audiences. Main report <i>written by evaluation facilitator, based on inputs from partners</i>. Partners produce their own partner-level evaluation reports.</p> <p>Thought piece led by the quality assurer.</p>	<p>Main report and other products for the implementing organisation. <i>Most products produced by the project team</i>.</p>

3.2. Roles and responsibilities in KNOWFOR evaluation

The KNOWFOR evaluation plan described a clear set of responsibilities for each of the key actors. Table 2 includes roles, responsibilities, estimated days for different stages and examples of specific activities taken to further illustrate what the roles included in practice.

There is no final estimate of the number of days, though it became evident during the ‘evaluation of the evaluation’ process that, in most cases, the number of days required had been considerably underestimated, especially for management and coordination.

According to the ‘evaluation of the evaluation’, the roles were well understood to a large extent, and the evaluation broadly proceeded according to the plan, though there

Table 2. Roles, responsibilities and estimated days in the KNOWFOR evaluation

ACTOR	PLANNED DAYS	EXAMPLES OF SPECIFIC ACTIVITIES TAKEN
<p>PARTNERS: CIFOR, IUCN AND PROFOR</p> <p>Main responsibilities: To lead, manage and deliver partner-level evaluation activities and reports, co-lead partner-level workshops, contribute to programme-wide evaluation activities and offer collaborative analysis.</p>	<p>82 days each (a total of 246 days) containing:</p> <ul style="list-style-type: none"> • planning: c. 11 days • data collection: 48 days • analysis and preparation of reports: eight days • partner and programme-level analytical workshops: five days • finalising reports and communicating the results: 10 days. <p><i>Note: these days do not include those used for discussing and deciding which evaluation approach will be used.</i></p>	<p>Partners either wrote or subcontracted case studies (three studies per partner).</p> <p>Partners wrote or coordinated outcomes stories (up to six per partner), lessons learnt (up to three per partner), Results Charts (one per partner) and value-for-money analysis (one per partner).</p> <p>One partner facilitated their own sense-making workshop.</p> <p>Partners commented on the final evaluation report.</p>
<p>EVALUATION FACILITATOR/CH</p> <p>Main responsibilities: Manage the overall evaluation process, undertake specific programme-wide evaluation activities, support the partner evaluation teams, facilitate workshops and produce the final evaluation reports.</p>	<p>76 days including:</p> <ul style="list-style-type: none"> • management: six days • planning: c. 12 days • support to partners: 12 days • programme-wide data collection and analysis: 15 days • facilitating sense-making and summit workshops: 15 days • report writing: 10 days • communication of the results: six days. 	<p>Provided ongoing evaluation support.</p> <p>Facilitated two of the partner-level workshops.</p> <p>Facilitated KNOWFOR-level summit workshop.</p> <p>Produced KNOWFOR-level Results Charts.</p> <p>Produced KNOWFOR final report.</p>
<p>DONOR/DFID</p> <p>Main responsibilities: High-level evaluation oversight to ensure the evaluation is delivered to the required quality standards.</p>	<p>A total of 20 days by the DFID Senior Responsible Owner and Programme Manager.</p>	<p>Initiated the partner-led approach.</p> <p>Participated in summit workshop</p> <p>Commented on the final evaluation report.</p>
<p>Quality assurer/ODI</p> <p>Main responsibilities: To provide improvement-focused reviews of evaluation planning processes and products, ensuring rigour, validity and quality of products.</p>	<p>72 days including:</p> <ul style="list-style-type: none"> • commenting on plans and initial reports and meeting partners: 20 days • participating in and commenting on the outputs from the sense-making and summit workshops: 24 days • commenting on partners’ final reports: five days • commenting on the final overall evaluation report, conducting an ‘evaluation of the evaluation’, writing this thought piece and contributing to final communication products: 23 days. 	<p>Reviewed case studies.</p> <p>Provided strength of evidence ratings for case studies.</p> <p>Participated in partner-level and summit workshops.</p> <p>Conducted ‘evaluation of the evaluation’.</p> <p>Reviewed the final evaluation report.</p> <p>Led the writing of this thought piece.</p>

were several delays. Developing the overall plan and partner-level plans took much longer than expected, partly because of the volume of work involved and the fact that the approach was new for many participants, but also because it was difficult to find a time when all participants were available for the coordination meetings. Also, some new elements emerged during the evaluation. For example, a light value-for-money analysis was mentioned in the evaluation plan, but these particular approaches were decided only in the summit workshop, where other key findings were agreed on.⁸ However, this type of evolution of plans is not unusual in evaluations, and especially so in developmental evaluations.⁹

Though roles were clear on paper, some participants raised questions whether the role of QA was clear enough:

[I] think their [QA] role could have benefited from some greater clarity. It was unclear if they had “teeth” in partner-led evaluation with three equal and independent partners, and it is not surprising that there were some tensions between the partner- and programme-levels’ focuses and findings.

This was a KNOWFOR programme evaluation, and the three key evaluation questions were explicitly pitched at programme level (see Chapter 1). While this was needed to capture programme-level results, the ‘evaluation of the evaluation’ revealed some dissatisfaction in its value within each organisation. As one interviewee said:

The individual case studies were useful and provided an opportunity to look at how activities led to outcomes [but] the questions focused on approaches that were applied across all three organisations, rather than exploring specific issues in each organisation, so the final report was rather lacklustre and not very insightful.¹⁰

While comments like this are valid, it has to be noted that this type of programme- versus project-level tension is not applicable only to partner-led evaluation and would have been present even if the evaluation had been independently led by an external evaluator.

The question of who should write the final report is interesting too. In KNOWFOR’s case, it was decided that partners would write their own partner-level reports while the evaluation facilitator should write the main report, on the basis of inputs from partners. But, given the

decentralisation of ownership throughout the evaluation, would it have been feasible and perhaps more appropriate for partners to write the joint report by themselves? By not being deeply involved in the aggregation of results at the programme level, partners likely lost an opportunity to learn from each other. However, having a third party engage deeply with the evidence base across partners added an additional check and balance for data quality. The evaluation facilitator had to dig deeply enough into partner-level information to be confident that claims were robust and evidence existed to support programme-wide claims.

Also, from the practical point of view, DMEL focal points would have probably struggled to dedicate enough time to engage with this level of information in an in-depth manner, given other internal time and project commitments. It is also very likely that this would have considerably prolonged an already-long process.

Perhaps the underlying issue was that the evaluation tried to cover everything in the three key evaluation questions. However, this resulted in missing interesting lessons that emerged at the partner or project level. This is common in many evaluations. Also, in the case of KNOWFOR, the partner-led reports (which were meant to capture the distinctness or richness of what is happening at a partner level) were not prepared before the main KNOWFOR report. In practice, it is probably impossible to completely remove this tension in any type of evaluation where partners represent different types of organisations, have independent activities and varied influencing strategies and limited collaboration. However, it is important to first recognise this tension, discuss it openly and then find strategies to mitigate it. If possible, it is good to use partners’ existing M&E systems as much as possible, which is what took place in KNOWFOR – as far as was feasible, given the very different nature of the organisations.

Another strategy is to produce different evaluation outputs for different audiences. For instance, in KNOWFOR, partners found different parts of the evaluation most useful for them. For example, one partner felt that one of the most valued learning points was the value-for-money analysis, while another had significant concerns about this chosen approach. While the final KNOWFOR evaluation report included a massive amount of synthesised information, partners were able to highlight their individual strengths, strategies and findings in their individual partner-level reports and other communication products.

8 The chosen value-for-money approaches were the Independent Commission for Aid Impact (ICAI) framework for the programme level (<https://icai.independent.gov.uk/wp-content/uploads/ICAIs-Approach-to-Effectiveness-and-VFM2.pdf>), and the Redstone Strategy-influenced approach for partner-level (<https://www.redstonestrategy.com/wp-content/uploads/2013/09/2013-09-30-IDRC-Helping-think-tanks-measure-impact.pdf>)

9 More information on developmental evaluations is available at http://www.betterevaluation.org/en/plan/approach/developmental_evaluation

10 Skype interview with a senior manager from one of the partner organisations.

4. What about credibility and independence?

Key messages

- Partner-led evaluations are geared towards learning. However, demands for accountability and rigour cannot be ignored. Robust results are important for both learning and accountability purposes.
- Partner-led evaluations may be more prone to some of the challenges and biases related to self-evaluations, such as seeing one's own work in a more favourable light.
- Though it is not possible for partner-led evaluations to be institutionally independent, it is possible to increase impartiality by introducing measures to mitigate potential biases.
- The first step is to acknowledge these limitations and biases and then put a clear strategy in place to address them to increase both robustness and credibility.

Questions about credibility and independence come easily to mind when talking about partner-led and self-evaluations and, sometimes, heavily participatory evaluations. Partner-led and self-evaluations are not, by their nature, independent. And, while their main purpose is usually learning and capacity-building (as discussed in Chapter 2), they also aim to produce results that are seen as credible and robust by both internal and external audiences. Having robust analysis and findings is not only important for accountability purposes, learning also needs to be based on credible and reliable evidence.

But how can we increase impartiality? And how can biases, especially those related to self-evaluations, be addressed? This chapter will discuss some of these challenges, along with what strategies KNOWFOR used to mitigate these and what else could be done. However, it has to be noted that some of these challenges and potential biases can be present in any type of evaluation, though partner-led (and self-) evaluations can amplify the likelihood of them.

4.1. Lack of independence

One of the main challenges in partner-led evaluations is lack of independence. Independence is considered to be one of the key evaluation principles, and is closely linked with the avoidance of biases and external influences. An example of biases, especially in non-independent evaluations, is that people may tend to see their own work in more positive light than an external evaluator would (positive bias) or might only want to cherry pick the most successful case studies

Box 5. Examples of how independence is discussed in evaluation manuals

‘Impartiality contributes to the credibility of evaluation and the avoidance of bias in findings, analyses and conclusions. Independence provides legitimacy to evaluation and reduces the potential for conflict of interest which could arise if policy-makers and managers were solely responsible for evaluating their own activities.’ —OECD DAC, 1991

‘Independent evaluation: An evaluation carried out by entities and persons free of the control of those responsible for the design and implementation of the development intervention. Note: The credibility of an evaluation depends in part on how independently it has been carried out. Independence implies freedom from political influence and organizational pressure. It is characterized by full access to information and by full autonomy in carrying out investigations and reporting findings.’ —OECD DAC, 2002

‘High quality evaluation depends on evidence that is objective, trustworthy and credible. Evaluations should be carried out from oversight of specialist who are independent from those responsible for the design and implementation of the development intervention being evaluated.’ —DFID, International Development Evaluation Policy, 2013

or units of analysis (sampling bias). They may also only see or put more emphasis on evidence that confirms what they already ‘know’ or believe (confirmation bias). Independence is seen as a way to avoid these types of biases and, thus, guarantee or at least increase the credibility of the evaluation process and findings. In line with this, Box 5 includes examples of how independence is defined and associated with credibility in evaluation manuals and guidance documents. However, it is important to recognise that independence usually refers to ‘institutional independence’ and might not mean the same as ‘analytical independence’. By this we mean that institutional independence (i.e. freedom from control or avoidance of conflict of interest) does not guarantee that external evaluators are not influenced by the programme, donor or their own values, preferences or even evaluation capacities and skills.

Being external can increase objectivity and reduce the likelihood of some of the biases mentioned above. External evaluators are accountable to an evaluation commissioner or donor, rather than the programmes themselves, whereas institutional M&E people are accountable for their organisation. However, being external does not magically take away all biases, as the evaluators might form close relationships with people involved in the programme or want to present their findings in a more positive light to obtain the next assignment.

We discuss some of these challenges in more detail before presenting the strategies that KNOWFOR used to mitigate them and to increase impartiality.

4.2. Biases related to case studies and how they are selected

As mentioned, partner-led evaluations (similarly to self-evaluations) can be more prone to biases such as assessing one’s work in more positive light or only seeing evidence that confirms what people already know or believe in. These and other types of challenges can be especially present in case studies, which often form the main evidence base for research or knowledge programmes such as KNOWFOR. In particular, large multi-project or portfolio research programmes may rely heavily on case studies, as it is usually impossible to investigate the entire programme.

4.2.1. Biased case study sampling and over-emphasising the role of research

While case studies are useful for illustrating in detail how research and other knowledge mobilisation activities can contribute to changes in policy and practice, they can also be associated with several caveats, such as difficulties in sampling and an overemphasis on the importance of research. This is a significant risk, as the case studies might come across as biased and promotional (Boaz et al., 2008).

A review completed by DFID in 2014 concluded that research uptake case studies need to be analysed with caution for the following reasons:

1. Case studies are often written to prove positive impact in order to justify continued investment in research. Thus, the selection of case studies is likely to be biased and lead to an overestimation of the policy impact of research.
2. Symbolic use of research is likely to be underrepresented and instrumental impacts are prone to being overrepresented.
3. Quantifying and proving the extent to which research has contributed to change is difficult and claims of contribution need to be viewed with caution.

Case studies on how research has contributed to changes in policy or practice typically focus only or mainly on the role of a particular research project. This means they may easily neglect the role of other contributing factors and context. To be fair, evaluations rarely have enough resources to properly investigate the roles of other influencing actors, power dynamics or politics taking place in that particular place and time. However, this may lead to a situation where the role of the research is overrated and other factors are neglected.

As self- and partner-led evaluations are not by their nature independent, their case studies may be more easily viewed as promotional pieces than as a way of providing robust and credible evidence of the research influence.

This means that in partner-led evaluation one has to pay additional attention to how case studies are selected (such as which type of purposive sampling is used), how data is collected and how the role of research is interpreted.

4.2.2. Case study challenges in KNOWFOR

Many of these case study and sampling challenges were also present in the KNOWFOR evaluation, which used purposive success case sampling to learn more about enabling factors that led to positive outcomes. Partners chose case studies themselves, meaning they were inclined to pick up ones they knew or anticipated had good or interesting outcomes. This means that KNOWFOR case studies may not represent ‘typical’ cases and the results therefore cannot be generalised across projects. However, this is very common and individual case studies are not normally used for generalisation, but for providing specific examples of influence or contribution. In KNOWFOR, the analysis of other contributing factors and context varied significantly. Some included more in-depth analysis, while others had shorter descriptions of context and other potential factors contributing to observed changes. However, the key challenge in some of the KNOWFOR case studies was that the conclusions were mainly drawn from staff interviews with limited triangulation. While this was commented on by the quality assurer, some of the case studies were not updated, given time and resource pressures, and thus were given lower strength of evidence ratings.

4.3. KNOWFOR strategies to increase impartiality and mitigate biases

The team undertaking the KNOWFOR evaluation made considerable efforts to increase the impartiality and robustness of the evaluation approach and process. As a starting point, the evaluation methodology used multiple sources of evidence. While case studies looked at particular projects in an in-depth manner, they were supplemented with shorter outcome stories as well as analysis across projects. While case studies looked at particular projects in an in-depth manner, they were supplemented with shorter outcome stories and analysis that looked across all projects (such as the Results Chart or the Uptake Rubric).¹¹

These strategies consisted of:

1. including independent elements in the evaluation process, such as the external QA function
2. utilising existing internal QA systems
3. ensuring transparency and consistent standards across evaluation processes and products by introducing a set of tools and frameworks to be used across organisations and carefully documenting methods and the process.

While some of these worked well, others had challenges. Below, we describe the strategies in more detail and Table 3 illustrates how different strategies have been used in different evaluation stages.

4.3.1. Including independent elements in the evaluation process

Contracting an independent QA body to review processes and products. An external QA body was contracted towards the end of the design phase shortly after partners had prepared their evaluation plans. This QA body not only reviewed the partner and joint evaluation plans, case studies (giving them strength of evidence ratings, which are included in the final report) and final evaluation report, but also attended partner- and KNOWFOR-level sense-making workshops.

The challenges for the QA body partly stemmed from the fact they came on board slightly late. While the quality assurer commented on the evaluation plans, a majority of partners had already moved on with their plans to proceed to the data collection stage, and some of the recommendations came when it was too late for them to be acted upon. One of the things the quality assurer commented on was ‘success case sampling’ and how focusing only on the most positive examples (though understandable) might limit the scope of internal learning. To balance this, it was decided that including some ‘lessons learnt briefs’ would capture situations where activities did not always work exactly as planned or expected. While some of these briefs did indeed capture these elements, others were less effective. This may have reflected the reality that being open about ‘failures’ or actions not working out as planned is still a major issue for most

organisations. While it is understandable that organisations competing for funding will want to focus on successful cases, this may limit the scope of (internal) learning.

Including a broader set of experts or external informants in the process (besides the quality assurer). Originally, the KNOWFOR evaluation plan included an external evaluation informant exercise to discuss and validate findings. However, this was kept quite limited and, ultimately, only three people were interviewed. However, external people attended some of the sense-making workshops to provide outsiders’ perspectives. The evaluation process and initial findings were also introduced and discussed with DFID evaluation advisors in an internal seminar. At least one value-for-money analysis was reviewed by an external expert. It has to be noted that having a broader set of external experts can be challenging, as it is not always easy to find people with a sufficient level of understanding of the whole programme and what it aims to achieve. More widely, this means there may be a trade-off between independence and prior knowledge of a programme.

4.3.2. Utilising existing internal QA systems

Utilising DFID’s internal QA system. An internal donor QA system called ‘Equals’ was used to review Evaluation ToR and the first draft of the final evaluation report. A DFID evaluation advisor not working on the programme completed a review, which included assessments of the evaluation methodology and design (including how inherent imbalances or biases are addressed), the robustness of the analysis and the findings. While the ‘Equals’ review template was more geared towards independent evaluations, one of its conclusions was that the KNOWFOR evaluation methodology was strong and included a good mix of methods. It also concluded that analysis is strong, taking into account varying data sources and using many good examples. On the other hand, one example of an aspect considered to be in need of improvement was the extent of clarity regarding ‘specific findings and their implications’, stating that ‘These are in the report, but often hard to find.’ (DFID, 2017: 5). The majority of the feedback and recommendations were incorporated into the final evaluation report.

4.3.3. Ensuring consistent standards and balance

Designing a set of tools and guidelines to ensure consistent standards across evaluation products and organisations. These items included a Strength of Evidence Tool, gender responsiveness rubric and overall KNOWFOR evaluation plan. There was some confusion about how some of the tools should be used across case studies, with some partners questioning the strength of evidence ratings given by the quality assurer.

Carefully documenting methods and process to increase transparency. The methods and process, together with

11 The KNOWFOR Uptake Rubric assessed the level of which targeted audiences (intermediaries and/or end users) have been reached or equipped as articulated in project plans in each KNOWFOR project.

detailed Results Charts were included in the evaluation report (which the quality assurer report was part of).

4.3.4. What else could have been done?

Given the exploratory nature of the partner-led evaluation, some of outlined strategies did not always work out as planned. For example, the question of how to address QA comments and to what extent was not clear for everyone and, as mentioned, there was some confusion regarding how some of the tools – such as the one related to strength of evidence – should be used across case studies.

What could have been done to further increase the credibility and robustness of the results? In KNOWFOR’s case, the participants’ views differed significantly. While some would have preferred more support and guidance, others felt there was already plenty of support available, but that this was perhaps not properly used by partners. This may reflect partners’ different capacities, skills and expectations.

Based on our experiences, we recommend that future partner-led evaluations use the strategies mentioned above, but also pay attention to:

- Bring in quality assurers early on in the process. In KNOWFOR, the quality assurer came on board at the very end of the design phase and partners had already moved to the data collection phase when feedback on the evaluations was given.
- Facilitate face-to-face, kick-off evaluation workshops with each partner, to work through evaluation approaches, challenges, potential biases, methods and tools.
- Pay additional attention to how case studies are selected across partners. If possible, the ToC should be used to guide the case study selection so that case studies can provide evidence to either support or refuse the identified hypothesis.
- Explore context and other contributing factors in case studies to ensure that the role of research is not overrated.
- Agree from the start how partners are expected to address quality assurer comments.

Many of these recommendations reflect the **need for investing independent elements, QA and guidance already at the planning stage**, not only in reviewing the findings as at that point it may be too late to make significant changes if needed.

Table 3. Strategies to address biases and increase impartiality applied in different evaluation stages

	 PLANNING	 DATA COLLECTION/ CASE STUDIES	 ANALYSIS, SENSE- MAKING AND SYNTHESIS	 REPORTING
Including independent elements in the evaluation process	Quality assurer brought in at the end of the planning stage to review evaluation plans	Quality assurer reviews case studies	Quality assurer attends sense-making workshops Quality assurer reviews Results Charts External evaluation information exercise Introduction of initial findings in a seminar for DFID evaluation advisors	Quality assurer reviews the final draft report Quality assurer report included in the final report
Using existing internal QA systems	DFID ‘Equals’ review of evaluation ToR	–	–	DFID ‘Equals’ review of the draft report; feedback incorporated into the final report
Ensuring transparency, consistent standards and balance	Support available from CH Tools and frameworks designed	Support available from CH Tools used in case studies ‘Lessons learnt’ briefs	Support available from CH Tools used in case studies, Results Charts and synthesis level	Transparent documentation of methods and process

5. Conclusions and recommendations

This paper has discussed the strengths, challenges and appropriateness of a partner-led evaluation approach, drawing on the experience of the DFID-funded KNOWFOR International Forestry Knowledge programme. While partner-led evaluation shares similarities with heavily participatory evaluations and self-evaluations, we also see it as having particular features that differentiate it.

The KNOWFOR evaluation process was perceived to be ambitious, time-consuming and sometimes complicated by the people who were involved in it. These challenges were balanced by higher than normal ownership (especially when compared to a typical external evaluation), which led to an increased buy-in, and a high potential for an increased uptake of the process, findings and recommendations. There are already examples of new or improved organisational and evaluation practices, such as the use of the ToC to plan how a particular project is expected to influence policy or practice, based on the work done before and during the KNOWFOR evaluation. It has also provided a rich learning experience for a wide range of people in partner organisations beyond the core M&E staff.

Based on our experiences, partner-led evaluations are most appropriate when there is a strong focus on internal learning and capacity-building. While producing robust results for donor or for external audience should not be partner-led evaluations' only aim (and if it is, another approach should be chosen), demands for accountability and rigour cannot be ignored either.

Partner-led evaluations may be more prone to biases, such as people seeing their own work in a more favourable light and only noticing the evidence that confirms what they already 'know'. This means that one has to pay additional attention to how case studies are selected, how data is collected and how the role of research is interpreted. There is a real risk that people in partner-led evaluation will give an impression of promoting their own work instead of providing robust evidence of how research and engagement and influencing activities have led to change. While it is simply not possible to make self-evaluations or partner-led evaluations institutionally independent, it is possible to include independent elements in the evaluation process as well as carefully and openly mitigate biases to increase the robustness and credibility of the evaluation process and findings. The first step is to be transparent about the existence of these limitations and biases from the start, and then develop a clear strategy on how to address them at each evaluation stage.

Recommendations

- Ensure that there are sufficient resources and time available.
- If partners' existing M&E skills or resources are limited, apply the approach to relatively simple programmes or projects where there is a clear ToC involving a clearly defined set of stakeholders with broadly similar objectives. While it is possible to use partner-led evaluation to assess more complicated or complex programmes, this requires a high level of M&E skills and capacities, as well as more dedicated time for each evaluation stage. Also, one would likely need additional analysis tools on top of those used in KNOWFOR for more complex programmes.
- It is impossible to say anything definite about the 'ideal' number of partners or how many may be too many or too few for partner-led evaluation on the basis of one example. What is crucial, however, is how the decision-making power is shared between partners. In KNOWFOR, partners had equal decision-making power but this may not always be the case. For example, one partner may have a leading role in the programme while others have minor or supporting roles. In this case, questions of proportional investment and decision-making would need to be addressed during the evaluation design stage.
- While it is important to spend time to plan and agree on the approach, we suggest that partners deliberately leave time and resources to explore unexpected opportunities as they emerge. However, it is important to manage emerging opportunities carefully, to avoid extending the time and resources for data collection at the expense of data analysis.
- Ensure there is a common understanding of the approaches among all organisations at the start of the evaluation and provide practical training (e.g. methods workshops) at the start, with continual support as required.
- Case studies that are contracted to independent consultants may bring additional challenges. If some of the approaches are new to partner staff and are not well internalised, they are less likely to convey them to external consultants in the right way. This further underlines the importance of targeted evaluation methods workshops to gain a clear understanding of what was needed at the start of field work.
- Use approaches that draw on the partners' existing M&E systems as much as possible.
- Recognise that the methods expert has final decision-making power regarding quality and acceptability of products.

References

- Belcher, B., Suryadarma, D. and Halimanjaya, D. (2017) 'Evaluating policy-relevant research: lessons from a series of theory-based outcomes assessments'. Palgrave Communications 3, Article number: 17017.
- BetterEvaluation (2017a) *Contribution analysis*. Victoria: Better Evaluation (http://www.betterevaluation.org/en/plan/approach/contribution_analysis).
- BetterEvaluation (2017b) *Multiple lines of evidence*. Victoria: Better Evaluation (<http://www.betterevaluation.org/en/plan/approach/mle>).
- BetterEvaluation (2017c) *Developmental evaluations*. Victoria: Better Evaluation. (http://www.betterevaluation.org/en/plan/approach/developmental_evaluation).
- Binnendijk, A. (2000) *Effective Practices in Conducting a Joint Multi-Donor Evaluation*, commissioned by the OECD DAC Evaluation Network. Paris: Organisation for Economic Cooperation and Development (OECD).
- Boaz, A., Fitzpatrick, S. and Shaw, B. (2008) *Assessing the impact of research on policy: a review of the literature for a project on bridging research and policy through outcome evaluation. Final report*. London: King's College London, Policy Studies Institute (http://www.psi.org.uk/pdf/2008/bridgingproject_report.pdf).
- Breier, H. (2005) 'Joint Evaluations: Recent Experiences, Lessons Learned and Options for the Future'. Working Paper for the DAC Evaluation Network. Paris: Organisation for Economic Cooperation and Development (<https://www.oecd.org/dac/evaluation/dcdndep/35353699.pdf>).
- Dart, J. and Mayne, J. (2005) 'Performance Story', in S. Mathison (ed.), *Encyclopedia of Evaluation* (pp. 307–309). Thousand Oaks, CA: Sage Publications, Inc.
- Dart, J. and Roberts, M. (2014) 'Collaborative Outcomes Reporting'. Carlton: BetterEvaluation, Australia and New Zealand School of Government (ANZSOG) (<http://www.betterevaluation.org/plan/approach/cort>).
- Department for International Development (DFID) (2013) *Evaluation Policy*. London: Department for International Development (<https://www.gov.uk/government/publications/dfid-evaluation-policy-2013>).
- DFID (2014). *What is the evidence on the impact of research on international development?* London: Department for International Development (<https://assets.publishing.service.gov.uk/media/57a089aced915d622c000343/impact-of-research-on-international-development.pdf>).
- DFID Equals Helpdesk (2017) *Quality Assurance, Evaluation Report* (internal document). London: Overseas Development Institute.
- Guijt, I. (2014) *Participatory Approaches, Methodological Briefs: Impact Evaluation 5*. Florence: UNICEF Office of Research (http://www.participatorymethods.org/sites/participatorymethods.org/files/Participatory_Approaches_ENG%20Irene%20Guijt.pdf).
- Independent Commission for Aid Impact (ICAI) (2011) *ICAI's Approach to Effectiveness and Value for Money. Report 1*. London: Independent Commission for Aid Impact (<https://icai.independent.gov.uk/wp-content/uploads/ICAIs-Approach-to-Effectiveness-and-VFM2.pdf>).
- KNOWFOR (2016) *Evaluation plan* (internal document). London: Overseas Development Institute.
- KNOWFOR (2016) *Quality Assurance plan* (internal document). London: Overseas Development Institute.
- Participatory Methods website (2017) Produced by the Participation Research Cluster, Institute of Development Studies (IDS). Brighton: Institute of Development Studies (<http://www.participatorymethods.org>).
- Patton, M.Q. (2012) *Essentials of Utilization-Focused Evaluation*. Thousand Oaks, CA: Sage Publishing.
- Organisation for Economic Cooperation and Development (OECD) (1991) *DAC Principles for Evaluation of Development Assistance*. Paris: Organisation for Economic Cooperation and Development (<http://www.oecd.org/dac/evaluation/50584880.pdf>).
- Organisation for Economic Cooperation and Development, Development Assistant Committee (OECD DAC) (2002) *DAC Glossary of Key Terms in Evaluation and Results Based Management*. Paris: Organisation for Economic Cooperation and Development (<http://www.oecd.org/dac/2754804.pdf>).
- Redstone Strategy Group (2013) *Helping Think Tanks Measure Impact, prepared for the International Development Research Centre's Think Tank Initiative*. Boulder/Redwood: Redstone Strategy (<https://www.redstonestrategy.com/wp-content/uploads/2013/09/2013-09-30-IDRC-Helping-think-tanks-measure-impact.pdf>).
- Taut, S. (2007) 'Studying Self-Evaluation Capacity Building in a Large International Development Organization', *American Journal of Evaluation* 28(1): 45–59.
- Torres, R. and Preskill, H. (2001) 'Evaluation and Organizational Learning: Past, Present, and Future', *American Journal of Evaluation* 22(3): 387–395.



ODI is the UK's leading independent think tank on international development and humanitarian issues.

Readers are encouraged to reproduce material for their own publications, as long as they are not being sold commercially. As copyright holder, ODI requests due acknowledgement and a copy of the publication. For online use, we ask readers to link to the original resource on the ODI website. The views presented in this paper are those of the author(s) and do not necessarily represent the views of ODI or our partners.

© Overseas Development Institute 2018.
This work is licensed under a Creative Commons Attribution-NonCommercial Licence (CC BY-NC 4.0).

All ODI publications are available from www.odi.org

Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ
Tel +44 (0) 20 7922 0300
Fax +44 (0) 20 7922 0399