Norad

Use of Evaluations in the Norwegian Development Cooperation System

Report 8/2012
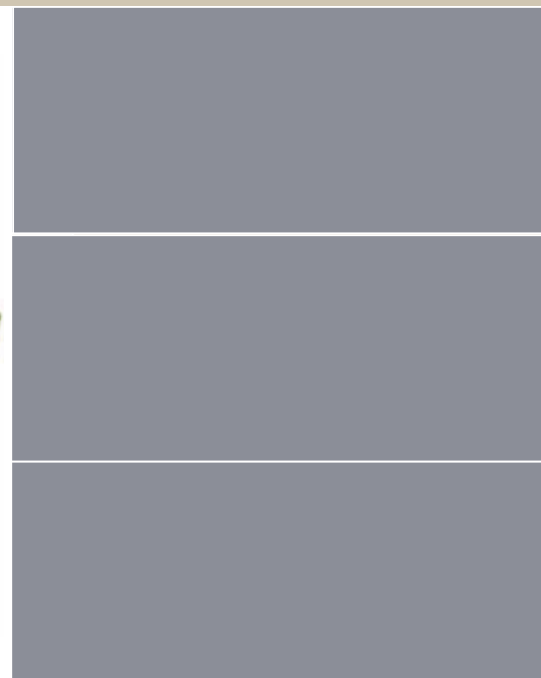
Use of Evaluations in the Norwegian Development Cooperation System

March 2013
Patrick G. Grasso, Linda Morra Imas, Nils Fostvedt
RAND Europe

Note on layout and language

The layout of the document has tried to conform to guidelines for accessibility and ease of reading, which require Arial font and left (not full) justification of the text.

The report has tried to avoid unnecessary use of acronyms and abbreviations.

Preface

We are often asked: Do your evaluations lead to anything? Are they used and useful?

We wanted to know and commissioned this study of the use of evaluations in Norwegian aid. It looks at our department, as well as at the units responsible for the management of aid in the Ministry of Foreign Affairs and in the Norwegian Agency for Development Cooperation (Norad). The study assesses the current evaluation follow-up system, what effects evaluations have had, and constraints in the use of evaluations for decision making and learning. It presents findings based on in-depth interviews, case studies, document reviews and a questionnaire survey with 230 respondents across Norad, the Ministry of Foreign Affairs and the embassies.

The study was carried out by the consultancy company RAND Europe, which is responsible for the content of the report, including the findings, conclusions and recommendations.

Oslo, April 2013

Hans Peter Melby
Acting Head, Evaluation Department

Acknowledgements

received her doctorate, with concentration in evaluation, from the University of Virginia.

Nils Fostvedt is a Senior Consultant and long-time practitioner in development, with particular focus on development evaluation. His consulting assignments since 2007 have included tasks with the Independent Evaluation Group (IEG) of the World Bank, the Independent Evaluation Department (IED) of the Asian Development Bank, and the Quality Assurance Group (QAG) of the World Bank. Prior to retiring from the World Bank, Fostvedt was Senior Adviser to the Director General, Evaluation, where he worked with the Board and senior management of the World Bank Group.

Quality assurance was provided by Lynn Karoly, Director of RAND's Quality Assurance Office and Xavier Legrain, Director of the Development Portfolio Management Group. Tarra Kohli, Project Associate, provided research and report production support to the team.

Table of Contents

Figures

Acronyms and Abbrevations

DAC	Development Assistance Committee
DANIDA	Danish official development assistance
DfID	Department for International Development (United Kingdom)
ICAI	Independent Commission for Aid Impact
IOB	Policy and Operations Evaluation Department (Netherlands)
MFA	Ministry of Foreign Affairs
NGO	Non-Government Organization
Norad	Norwegian Agency for Development Cooperation
OECD	Organisation for Economic Development
RBM	Results-Based Measurement
SADEV	Swedish Agency for Development Evaluation
Sida	Swedish International Development Agency
UN	United Nations
USAID	United States Agency for International Development

Executive Summary

Executive Summary

This evaluation report was commissioned by the Norad Evaluation Department in order to respond to the need for understanding whether and how the evaluations it conducts are used, and how the systems for learning from and using those evaluations can be improved.

Evaluation use is important because, unlike basic research, evaluation is intended to provide accountability for achieving results from the use of funding, and learning from experience in ways that can be put to practical use. The extensive literature on evaluation use suggests that evaluation can affect policy and operational decisions. Much attention focuses on the explicit use of evaluation through the implementation of recommendations. While this is an important kind of evaluation use, this study highlights a number of ways in which evaluations can have more indirect influence on actions, perceptions, and even how specific issues are conceptualized.

In conducting this evaluation the team used a mixed-methods approach. This entailed document reviews, interviews with key informants, case studies of four specific evaluations, detailed review of the content and structure of related evaluation reports, a citation search for Norad evaluations, and a comparative organizational analysis of evaluation units in Norad peer organizations.

Major conclusions

The evaluation found that the Evaluation Department has done a number of things to enhance the usefulness and actual use of its evaluation work. Among these has been a practice of broadly consulting with key stakeholders in developing its evaluation program to ensure that there is interest in the topics to be covered. This is recommended international good practice for evaluation. The Department also has an established system for disseminating the results of its formal reports through seminars and other mechanisms. The Department has been successful in having its evaluation reports posted to a range of websites that focus on issues of development and/or evaluation and research.

At the same time there have been limitations to its effectiveness in some areas. In some cases evaluations are not targeted on high-priority issues or are not delivered when they could affect decisions. There also has been tension between the need to adhere to the original terms of reference for an evaluation in order to maintain quality and meet procurement requirements, and the need to respond to changes on the ground or opportunities for different analyses that could provide more useful or interesting information.

The reports themselves, though generally of good quality, sometimes are not user-friendly. Excessive use of jargon and acronyms, for example, makes it more difficult for some potential users to access the information readily. Moreover, some officials indicated that, from their perspectives, the reports frequently read more as academic papers than as action-oriented evaluation reports. In part, this may reflect the frequent use of academics for carrying out the evaluation work. Relatedly, some stakeholders raised concerns that a relatively small number of consultants may be carrying out a large number of the evaluations, providing a somewhat restricted view of the issues addressed, though the evidence on this is not strong.

Report recommendations in some cases are not well-targeted or practical for implementation. This limits the usefulness of the evaluation work, since recommendations are a major vehicle for influencing policy-making and management decisions.

Finally, the underlying theory of change for the Evaluation Department needs to take account of the fact that Norad evaluation reports actually are used by a wide array of actors. So that theory needs to include channels for both direct and indirect influence that can amplify the direct use of the findings and recommendations.

Recommendations

Based on the findings and conclusions from this study, we have the following recommendations. First, the Evaluation Department should strengthen the quality of its evaluation reporting by taking the following steps:

- Monitor the frequency of use of individual consultants to ensure that the range of those used is sufficiently broad to avoid the appearance—and perhaps reality—that it relies too heavily on a small group with pre-digested views; at the same time, continue efforts to broaden the range of competitors.

- Develop guidelines to ensure that report recommendations are clear and fact-based, that they can be implemented by the agencies to which they are addressed, and that it is feasible to measure whether and to what extent they have been carried out.

- Provide short, clear executive summaries for all reports.

- Develop short, non-technical briefs for all reports to make the results more widely accessible; the summaries prepared for Annual Reports can be used for this purpose.

- Provide targeted briefings to those units, particularly country offices, involved in the evaluations by arranging to brief them on the findings of relevant evaluations as part of regularly-scheduled unit staff meetings.

- Emphasize dissemination—beyond the dissemination seminar—as a key part of the work of the Evaluation Department by adding time for dissemination activities as a standard evaluation activity, and actively explore additional channels for communication, such as external networks, social media, and podcasts, to better target and disseminate its messages.

- Improve the "visuability" of reports, as discussed in Annex 3.

Second, the Department should improve its processes by:

- Formulating Terms of Reference that are clear, but also flexible enough to allow evaluators to pursue the most promising paths for answering the evaluation questions.

- Making efforts to keep stakeholders more engaged with the evaluation while the work is on-going, perhaps through periodic updates on progress.

Third, the Department should work with the leadership of Norad and MFA to address the structural ambiguities under which it operates currently. One option might be to provide a routine venue for the Director to deal directly with the Ministry's top management in order both to allow for an airing of organizational issues affecting the Department, but also to open a channel of communication that can increase the use of evaluation findings and recommendations at the highest levels. A more far-reaching possibility is to have the Department report directly to Parliament, as is the practice in the Netherlands. (A full review of such options was beyond the scope of this evaluation, however.)

Fourth, on the issue of follow-up of recommendations, we believe that this is the responsibility of MFA as the implementing organization. MFA could take several steps to ensure adherence to the requirements for responding to Evaluation Department reports, including:

- Making timely production of follow-up plans and actions an explicit element in managers' performance metrics.

- Conducting an annual review of compliance with requirements for responding to evaluation recommendations, with reference to specific reports, and disseminating the results to its management team and to Parliament.

- Promoting evaluation use by highlighting "success stories" of good use.

The Evaluation Department also should consider how it might promote accountability. One option, which would not detract from its independence, would be to produce and issue a report every two years on the status of recommendations with a focus on those which MFA has agreed to implement.

1. Introduction

This study was commissioned by the Evaluation Department of the Norwegian Agency for Development Cooperation (Norad) to examine how and to what extent the evaluations it undertakes are used by others in the Norwegian development system.

A.   Purpose and Objectives of the Evaluation

1.   The purpose of this study is to respond to the needs of Norad for understanding whether and how the evaluations it conducts are used, and how the systems for learning from and using those evaluations can be improved. Evaluation serves the dual functions of providing accountability for the results of Norwegian development cooperation assistance and learning from experience. Evaluation itself involves the use of taxpayer resources, so if they are not used those resources are not being well-deployed. Thus, the purpose of the evaluation is to assess whether and to what extent Norad evaluations are being used to help improve the work of the Norwegian development cooperation system in producing results.

2.   The specific objectives of the evaluation are to: (1) identify to what extent, how and when in the evaluation process, evaluations produced by the Evaluation Department are used; (2) assess which factors have been most decisive in promoting/hampering use of evaluations in the Norwegian development cooperation system; (3) assess whether the existing system for follow-up of evaluations meets the current needs of the Norwegian development cooperation system; (4) assess whether the existing system for dissemination (evaluation product, communication channels and mechanisms) of evaluations meets the current needs of Norad; (5) identify possible mechanisms for cumulative learning; (6) assess the main factors influencing decision-making ; and (7) provide recommendations on how the Norwegian aid administration can improve systems for learning and evaluation use, including explicit recommendations for the Evaluation Department and the evaluation function.

B.   Evaluation in Norad

3.   Norad evaluations are conducted by the Evaluation Department under the terms of the Instruction for Evaluation Activities approved in 2006.[1] Under the terms of these Instructions, the Evaluation Department is charged with

---

[1]   "Instructions for Evaluation Activities in Norwegian Aid Administration," Norwegian Ministry of Foreign Affairs (MFA), May 2006.

evaluating "all activities linked to the administration of ODA-reportable expenditure in the Ministry of Foreign Affairs' budget." This is a broad grant of authority to the Department, providing it with ample space to evaluate development activities carried out through the Ministry and Norad. This authority extends to evaluations of "goal achievement and results relative to adopted plans [i.e., effectiveness];…whether the consumption of resources is commensurate with the results achieved [i.e., efficiency]." It also allows the Department to systematically synthesize learning from experience through its evaluation work, and to provide information from evaluations to both funding authorities and the general public.

4. In carrying out these activities, the Department is required to build on the evaluation guidelines developed by the OECD-DAC.[2] Evaluations are to be staffed by external consultants, who have "transparent and full access" to documents relevant to the evaluation, and to staff of the Ministry, Norad, and the embassies. Consultants are selected using international tendering processes that are open and transparent. The Department is responsible for facilitating the evaluation process and may act as an observer, but is not to interfere with the "neutrality and independence" of that process.

5. The Department also manages the process of formal comments on draft evaluation reports, acting as interlocutor between the evaluation team and those providing comments, including Ministry, Norad, and embassy staff, as well as other reviewers. When the final report has been accepted, the Department prepares a memorandum summarizing the findings and recommendations from the evaluation, and highlighting any comments received. Most important, the memorandum provides suggestions on needed follow up. It is sent from the Evaluation Department Director to the Secretary General of the Ministry through the Director General of Norad, who may append his/her own comments. The Ministry then determines what issues require follow-up, by whom, and within what timeframe. A follow-up plan is to be published within six weeks. No later than one year after the plan is published the responsible unit is required to report to the Secretary General on the actions taken in response to the evaluation. As discussed in Chapter 4, this formal process is not always carried out as described.

C. Evaluation Use

6. Use is one of the perennial issues for evaluation. Unlike basic research, evaluation generally is intended to be used (although even this is questioned by some[3]). But this simple assertion raises a set of issues that can be summarized into three broad questions: (1) What are the uses of evaluation? (2) What factors facilitate or impede evaluation use? And (3) how can evaluation use be measured?

---

2   OECD, *Evaluation in Development Agencies*, Better Aid, OECD Publishing, 2010.
3   Gary T. Henry and Melvin Mark, "Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions," American Journal of Evaluation 24,(2003), pp. 293-314.

7. The evaluation literature includes a number of typologies of evaluation use. The earliest work focused on three types of use: instrumental or decision-making, conceptual or understanding, and strategic or advocacy.[4] More recent work has expanded on this list by focusing on how the evaluation process itself may provide other uses.[5]

8. Fleischer and Christie synthesized thinking on evaluation use based on an extensive review of the history and development of the concept, deriving these categories[6]:

   a. Instrumental—Decision-makers use evaluation to modify the program, project, activity evaluated.

   b. Conceptual—The evaluation provides a new way to think about an issue.

   c. Enlightenment—The evaluation adds knowledge to the field that is available to others, not just the decision-makers.

   d. Process use—Participation in the evaluation process leads to evaluative thinking, and thus to organizational, program, cognitive, and behavioural change.

   e. Persuasive or symbolic value—The very fact of the evaluation helps to persuade stakeholders that there is accountability for the activity and can persuade them of its value.

9. There is no consensus on the categories of evaluation use. In fact, there is an extensive literature on evaluation use going back more than three decades that includes a range of different categorizations or frameworks. Since the 1980s, three broad categories of use widely cited have been instrumental, conceptual, and symbolic use. In the 1990s the notion of process use was added to this inventory, and subsequently the enlightenment or educative function of evaluation was incorporated into use frameworks.[7] The Fleischer-Christie approach used here captures these developments. In addition, it allows this study to be compatible with a separate study of evaluation in Norwegian civil society organizations.[8]

10. Much of the literature on evaluation use is theoretical. This study seeks to contribute to our understanding of evaluation use by presenting empirical information from a combination of case study and institutional analyses.

---

4 William R. Shadish, Thomas D. Cook, and Laura C. Leviton, *Foundations of Program Evaluation: Theories of Practice*, Thousand Oaks, CA: Sage, 1991.
5 Michael Q. Patton, *Utilization-Focused Evaluation*, 3rd ed., Thousand Oaks, CA: 1998.
6 Dreolin N. Fleischer and Christina A. Christie, Evaluation Use, "Results from a Survey of American Evaluation Association Members," American Journal of Evaluation 30:2, June 2009, pp. 158-75. This study builds on a previous survey reported in Hallie Preskill and Valerie Caracelli, "Current and Developing Conceptions of Use: Evaluation," American Journal of Evaluation, 18:3, September 1997, pp. 209-25. A third survey currently is underway.
7 A useful summary of this history can be found in Robert Wood Johnson Foundation, An Ecological Understanding of Evaluation Use: *A Case Study of the Active for Life Evaluation,* Princeton, NJ, 2010.
8 *A Study of Monitoring and Evaluation in Six Norwegian Civil Society Organisations. Report 7.12.* Norad, 2012.

In particular, the study helps to identify both the opportunities for and the impediments to increased use of evaluation findings, lessons, and recommendations in a complex, multi-organizational environment, such as Norway's development cooperation assistance system.

D. Theory of Change

11. Over a period of decades evaluators have come to recognize that to do justice to a program, project, or other activity being evaluated it is important to understand just what it was intended to do. Typically, such activities are designed to affect some form of change; certainly this is the case in international development work. In her pioneering work, Carol Weiss promoted the idea of laying out the chain of events from activities to final goals or objectives they are expected to achieve, and clarifying the underlying assumptions that specify just how this is supposed to happen.[9] This has proven to be not only an effective way to design evaluations, particularly of complex activities, but also to plan and design the activities themselves so as to achieve better results.

12. At the start of this evaluation, the Evaluation Department provided a preliminary theory of change setting out a set of propositions about how evaluations promote change in the Norwegian international development system. See figure 1. This acted as a roadmap as we began our work. This is a straightforward theory that posits a relationship between the Department's independence, its work program and products, and their use by actors in the Norwegian development system.

13. As formulated, the theory argues that the Department's Instruction[10], giving it a broad mandate to carry out evaluations, allows it to adopt a relevant evaluation program, one informed by stakeholder consultations but decided independently. The evaluations carried out through this program and the resulting reports provide an opportunity for learning from experience as they are shared with key stakeholders through various forms of dissemination. The theory assumes that through this sharing the evaluations affect the views of development-related actors, as expressed in follow-up memoranda and action plans, and that these subsequently are translated into actions, such as policy changes or new/revised project and program designs. Of course, an underlying assumption is that the findings and recommendations from the evaluations are themselves fact-based and clear, properly targeted to the right actors, implementable, and possible to monitor as they are put into practice. We return to the theory of change in Chapter 5.

14. Chapter 2 sets out our approach and methods used in this study. It describes in detail the kinds of data that were collected and how they were analysed, including any limitations. In Chapter 3 we present the findings from four case

---

9 Carol H. Weiss, " Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families, " in J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New Approaches to Community Initiatives: Concepts, Methods, and Contexts* (Vol. 1), pp. 65-92. Washington, DC: Aspen Institute.
10 Op.cit, MFA, 2006.

studies. There were selected, in consultation with Norad, so as to allow for contrasts between instances where use was high and those where it was low. This selection strategy is useful for helping to identify the factors that may lead to more or less use. Chapter 4 contains findings from broader institutional analyses we undertook. Much of the data for this chapter comes from an internal survey conducted by the Evaluation Department among Norad, Ministry of Foreign Affairs, and embassy staff. In addition, it includes comparative information on evaluation in other development organizations and an analysis of how the Evaluation Department fits into the overall development architecture. Finally, Chapter 5 presents our conclusions and recommendations.

Figure 1: Formal theory of change of the evaluation function in the Norwegian Development Cooperation System.
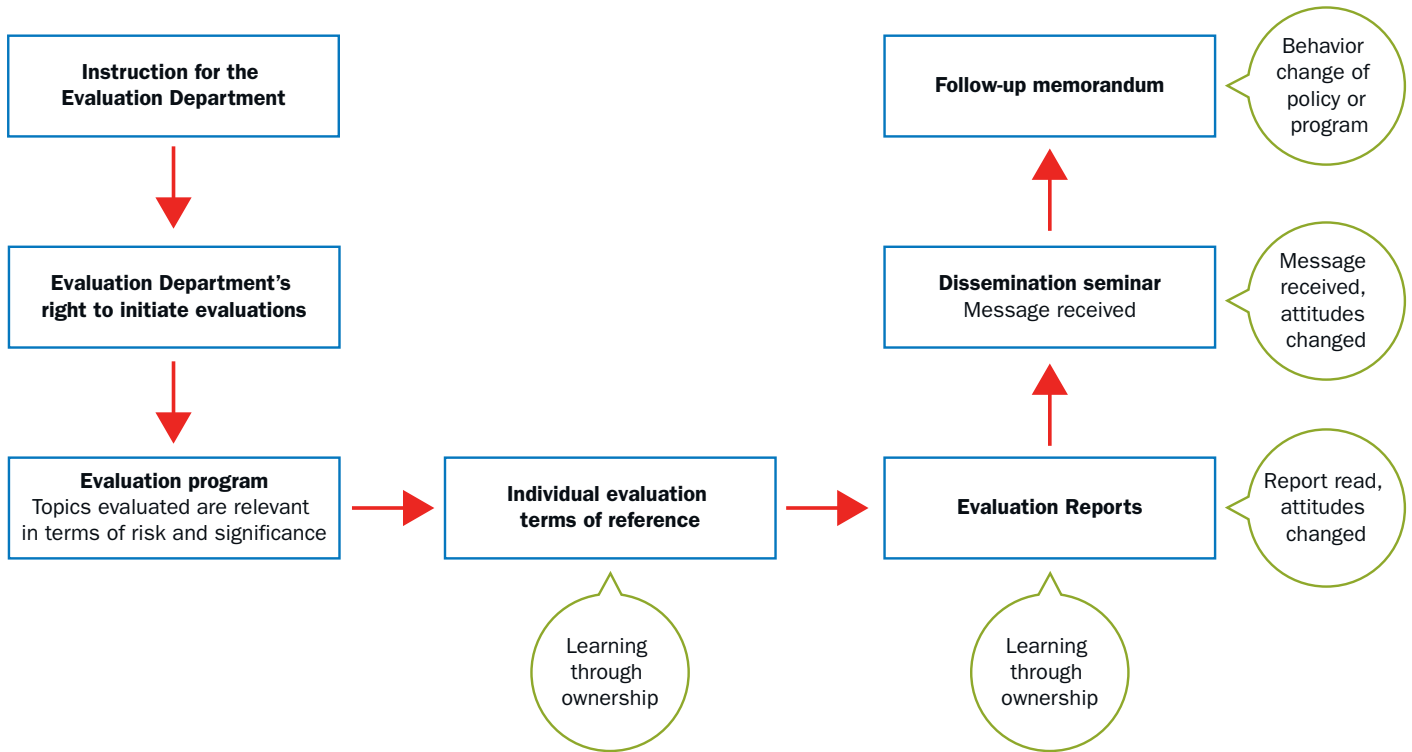
2. Approach and Methodology

This evaluation used a variety of approaches to address the objectives of the study. This mixed-methods approach is ideal for complex evaluations of processes, such as evaluation communication and use, because most of the information is qualitative and no single approach can capture the full picture.[11]

A. Approach and Methods: Individual Evaluations

15. This study has two sets of components. One looks at how individual evaluations have been used, the second at the systemic issues around evaluation use. At the individual evaluation level, data were collected in several ways, allowing for triangulation of findings.

Case Study

16. Case study is a key part of the evaluation as it provides the in-depth understanding of why an evaluation was used or not used. Case study is appropriate as it is "…a method for learning about a complex instance, based on a comprehensive understanding of that instance obtained through extensive description and analysis of that instance taken as a whole and in its context."[12]

Data Collection

17. Data collection consisted of semi-structured interviews and document review. Contents of files on each of the reports were examined. Initial in-person interviews were conducted the week of September 23rd, 2012 and additional in-person interviews the week of October 13th, 2012. Telephone interviews were conducted thereafter as possible during October, and some follow-up interviews in January 2013. Snowball sampling was used to identify those knowledgeable about use/lack of use of the evaluation. This is a technique were the interviewer asks each interviewee to provide the names of others who should be interviewed on the subject of the evaluation.[13] The list of interviews conducted is attached in Annex 1 and the semi-structured interview guide is in Annex 2. Analysis relied on triangulation within cases and across cases as well as findings from other parts of this study, such as a review of the four reports as products (Annex 3), in addition to other document reviews.

---

11 US Agency for International Development, "Conducting Mixed Methods Evaluations," Washington: USAID, 2010.
12   Morra, Linda G. and Amy C. Friedlander. *Case Study Evaluations*. Washington, D.C.: The World Bank,1999.
13 See Rowland Atkinson and John Flint, "Snowball Sampling," in Michael S. Lewis-Beck, Alan Bryman, and Tim Futing Liao (eds.), *The SAGE Encyclopedia of Social Science Research Methods*, Thousand Oaks, CA: SAGE Publications, 2004.

Limitations

18. The case studies had several important limitations. Each case is complex, and the time for examining its complexity was short. First, 30 days in total were allotted for the case study. This meant that while a report might have focused on several countries—for example Petroleum-Related Assistance itself consisted of case studies of Mozambique, Bangladesh, East Timor and Angola—use of individual case studies within a report was not studied. No in-depth look at a country's use of a report was conducted. Second, not all those persons identified as potential interviewees were contacted and not all those who were contacted were available for interview. The interviews were not exhaustive. Given the number of names that were identified, interviewees were prioritized by likely relevance to the report. Those interviewees located in Norad were most accessible. Those individuals within MFA or the embassies were difficult to access because of travel schedules and the frequency of turnover of positions. Third, the main focus of the case studies was not to follow up on each recommendation made in each report, but rather to increase understanding in each of the cases of why key actors perceived the reports as more or less useful. Generally those interviewed did not have recollection of specific report recommendations as much as a general sense of the recommendations and actions.

19. Several case study interviewees asked that direct quotes either not be used or not be personally identifiable to them and we have honored those requests. With the relatively small number of interviews conducted for each case, we have tried to understand the case and to tell the story of the usefulness and use of each report broadly, while respecting and reflecting the variety of perspectives. We have not valued the perspective of one group, for example consultants, more than another, such as Evaluation Department task leaders. We caution that it is not meaningful to compare responses of different groups—e.g. Norad staff vs. MFA—either within a case or across the four cases.

Data Analysis

20. Content analysis software, AtlasTI, was used to help organize the qualitative interview data. Coding categories were developed based on the interview questions and a reading of the interviews. The data were then coded by these topics in order to facilitate review within each case and across cases. Comments on dissemination of reports, those positive and those negative, are an example of a coding topic. A more detailed example of codes used is provided in Annex 4.

Key informant interviews

21. The interviews with Norad and Ministry officials during the initial stages of the study provided a good deal of useful information on a range of individual studies. In particular, interviewees were asked to indicate evaluations that had been used well, or not, and a rich set of responses was developed.

22. The team used the information from these interviews to develop a set of observations on both which evaluations were/were not especially useful, and—importantly—the reasons why, as given by interviewees. This analysis required some follow-up with interviewees, and in some cases involved identifying additional sources.

23. The interviews were analyzed in conjunction with the mapping and case study data in order to triangulate the findings on whether and why/why not individual evaluations are used.

Citations

24. Citation searches also helped to identify the extent to which the evaluations have been used. While citation analysis can be problematic as a measure of influence or use, it does allow insights into how much attention an evaluation has received.

25. For purposes of this analysis, the primary searches were of major Norwegian media. The objective was to determine whether Norad-specific evaluations were being brought into public discourse on Norwegian issues relevant to development cooperation. In addition, selective searches also were conducted in non-Norwegian sources, including countries covered in evaluations, to determine the extent of coverage of Norad evaluations in non-Norwegian contexts, including the literature on international development evaluation.

B. Approach and Methods: Systemic Issues

26. The findings from the analysis of individual evaluations help to inform the analysis of the systemic issues in evaluation use. However, in addition, data specifically about those systemic issues were collected and analyzed.

Document reviews

27. At this level, the formal policies and guidance for evaluation conduct, reporting, response, and implementation of recommendations were the key documents reviewed. As part of the analysis, the processes discussed in these documents were compared with those recommended by international good practices (e.g., those of the OECD), and to other development organizations, as appropriate. The idea was to ensure that they are not seen in a vacuum, but within the context of the development community's knowledge of and experience with such systems.

28. In addition, we gathered information on use for all Norad Evaluation Department reports from 2006-11 to track their use, and a separate review of the 15 reports from 2011-12. These were done through reviews of existing documents, as well as on-line searches in both Norwegian and English.

Key informant interviews

29. Interviews conducted during the team's mission to Oslo provided a wealth of information on these issues. Follow-up interviews were carried out with Norwegian sources as well as with development evaluation departments in a set of peer countries (see below) in order to clarify specific points and fill gaps, mostly through telephone and email exchanges.

Survey

30. The Evaluation Department provided the team with the results of a survey it conducted among Norwegian aid staff in Norad, MFA, and the embassies. The survey was designed to learn more about how staff view the evaluation function and how they use (or do not) evaluation findings and recommendations. The Rand team used the analysis of data from the survey to help inform this part of the study.

31. Survey questionnaires were administered to staff of Norad, MFA, and embassies during the first week of November 2012, with several follow-up requests during the course of the week. Altogether, 640 staff were surveyed, and 232 (36 percent) responded. Those surveyed included all staff involved in development assistance at a substantial or decision-making level, not a sample. Data comparing selected characteristics of respondents to the target population showed no obvious major sources of response bias: there was a slight over-representation of women (51 percent of respondents versus 46 percent of the target), which could have had at most a small effect on some findings, given that men and women sometimes differed in responses. However, the distribution of respondents by age and agency was almost identical to the target, eliminating those as potential sources of response bias.

32. As might be expected, respondents were highly-educated, with nearly 90 percent having attained at least a Master's Degree. More than one-third (37 percent) had worked in the Norwegian development cooperation system for five years or fewer, and more than 82 percent had been in their current positions for no more than five years. So responses likely represent more recent experiences, which is helpful for this study; in fact, 82 percent reported that a program or policy for which their unit is responsible had been evaluated within the past five years, and 58 percent that such an evaluation had been conducted by Norad. In addition, nearly two-thirds of respondents (64 percent) held positions as Advisors, Senior Advisors, or Special Advisors, and most of the rest (29 percent) as mid- to top-level Managers. These groups should be well-positioned to report on the use of Norad evaluations. Indeed, nearly a third (32 percent) reported spending between 11 and 30 percent of their work time reading evaluations, research reports, and policy studies.

33. Many respondents also had had some level of exposure to monitoring and evaluation (M&E). Nearly three-fifths (59 percent) had received some M&E training. Of those, 70 percent had received internal training, more than half

(51 percent) had participated in external workshops or seminars on M&E, and nearly one-fourth (24 percent) had had training through university programs. And almost three in five (57 percent) had taken part in at least one evaluation. However, respondents were not necessarily well-positioned to use evaluations for making policy decisions. On a scale measuring their ability to influence Norwegian development cooperation policy decisions within their own regions or areas of work, about one-third (33 percent) rated themselves on as 7 or higher on a scale of 0 to 10, where 10 represented the most influence.

Organizational analysis

34. The Norad evaluation process was compared with that of peer organizations to identify significant similarities and differences that help illuminate variations in effective use, and to inform possible recommendations concerning the Norwegian system. This analysis used information available through the OECD/DAC Evaluation Network.[14] More in-depth analysis was conducted on four units through telephone interviews with the evaluation unit heads or their designated representatives. At the suggestion of Norad staff, Sweden, Denmark, the Netherlands, and the United Kingdom were selected. These cases vary widely in organizational structures and processes, making them particularly useful for identifying possible opportunities for strengthening Norad's evaluation system, as well as for identifying cautionary lessons on initiatives that likely should be avoided.

---

14   Op.cit OECD, 2010.

3. Finding on Individual Evaluations: Case Study on Use

The Case Study on Use is intended to provide some in-depth understanding of why specific evaluations were or were not used and, as appropriate, how they were used. Case study is "...a method for learning about a complex instance, based on a comprehensive understanding of that instance obtained through extensive description and analysis of that instance taken as a whole and in its context."[15]

A. Nominations and Selections

35. The nomination approach for the case study was extreme case, a form of bracketing. In interviews conducted in Oslo the week of August 13, 2012, key informants were asked to nominate evaluation reports that were examples of particularly strong use of evaluation and conversely, those that were examples of particularly weak use of evaluation. Some 19 evaluation reports were nominated across the interviews.

36. Bracketing implies multiple case studies, but more is not necessarily better when using case study methodology. The key is obtaining sufficient understanding of each selected case in its context within the time available. The results are not generalizable to the universe of evaluation reports. We determined that within the time available we could undertake four cases and that in selecting the cases, we would take into account several factors of importance: issuance date, type of evaluation, and variety of perspectives. Each is discussed below.

37. **Issuance Date:** We knew that our case study would rely heavily on data obtained through interviews and that the information would be based on recall of events and perceptions. It is well known in the social sciences that the longer the period the respondent is being asked to recall, the less reliable the data obtained. However, we also know that sometimes it takes years to see the full impact of a study in terms of implementation of recommendations. With regard to the case study on evaluation use, it meant that while sufficient time must have passed since a report issuance for use to occur, generally we preferred more recent reports to older reports. Therefore we determined that cases from reports issued in 2010 and 2011 were preferred with one older report to be selected.

---

15   Morra and Friedlander, 1999, p.3.

38. **Range of Report Types:** The Evaluation Department commissions different types of evaluations including broad thematic evaluations that cut across organizational lines, country or region specific evaluations, sector specific, program specific, and process specific evaluations. It has tried real-time evaluation as well as evaluability assessment. While the case study sample of four reports cannot represent all the evaluation types used, we gave preference to those reports reflecting different types of evaluations in the case selection.

39. **Variety of Perspectives:** All else being equal, we wanted the case study sample to reflect the nominations of different key informants. Cases where the same evaluation was nominated as both a good use example and a poor use example were of particular interest as they were likely to reflect different understandings of "use".

40. These criteria were generally applied. The cases selected are illustrative and not representative of the Department's issued reports or of the 19 reports nominated. Based on these considerations, the following four nominated evaluations were selected for case study:

- *Evaluation of Norwegian Development Cooperation to Promote Human Rights*. Report 7/2011.
  This report is an example of a thematic evaluation. It was nominated as an example of poor use. While there were other examples of thematic evaluations, this one was under management discussion at the time of our study and a source of frustration in terms of responsibility for follow up.

- *Pawns of Peace – Evaluation of Norwegian Peace Efforts in Sri Lanka, 1997-2009.* Report 5/2011.
  This report is an example of an unusual effort to evaluate a peace process. This is not a topic with a commonly used evaluation methodology. It was also of interest in that it was a report requested by MFA and a report with high visibility. Additionally, it was the only report nominated as an example of both good and poor use.

- *Evaluability Study of Partnership Initiatives – Norwegian Support to Achieve Millennium Development Goals 4 & 5.* Report 9/2010 – Study.
  This study used the relatively new approach of evaluability assessment. It was undertaken in response to the Evaluation Department's often-heard complaint that their evaluation reports almost always indicate a lack of good data limited the findings. The report was nominated as an example of poor use. Norad staff had high hopes for this evaluation and had encouraged and supported its conduct. They were disappointed by what they perceived as low impact of the report and wanted to know why it did not have greater use.

- *Evaluation of the Norwegian Petroleum-Related Assistance - Case Studies Regarding Mozambique, Bangladesh, East Timor and Angola.*

Report 1/2007.
This report was selected as a good example of use that occurred over time. It is also a sector report and reflects a long-term line of effort for the Evaluation Department. The Evaluation Department was particularly interested in including it as a case.

B. Findings Case 1: *Evaluation of Norwegian Development Cooperation to Promote Human Rights.* Report 7/2011

41. This report was often discussed as part of a series of three evaluations related to human rights—one generally on human rights and the subject of this case study focus, one on child rights, and one on the disabled. The report addressed questions such as: Is there a clear policy behind the Norwegian support to human rights? Is there a specific Norwegian profile in the support to human rights? What results have been achieved? The three main components of the evaluation were mapping and analysis of the human rights portfolio, an evaluation of the support to freedom of expression with particular emphasis on free and independent media, and an evaluation of the country programs of the Norwegian Centre for Human Rights in Indonesia and South Africa. A major finding of the study was that despite Norway's significant policy and funding commitment to Human Rights, there was no mechanism in place to ensure the systematic incorporation of human rights in the prioritization, allocation and formulation of Norwegian development cooperation. In an illustration of conceptual learning, one interviewee stated that the report was useful just in its finding that "no one place in MFA has responsibility for Human Rights strategy or its mainstreaming."

42. This evaluation began with issues of its scope. As the report indicates, support to the protection of human rights around the world has consistently been a major objective of both Norwegian foreign policy and development cooperation. In the 2009 *Report on the Role of Human Rights in Norwegian Foreign and Development Policy*, the Ministers of Foreign Affairs and Development Cooperation gave a joint statement reinforcing that one of the Norwegian government's primary objectives is to promote an international legal order based on democracy and human rights. They indicated that not only do all states have a legal obligation to advance human rights, they also have a moral obligation to protect individuals from human rights abuses. Within the Ministry of Foreign Affairs, the Department of UN, Peace and Humanitarian Affairs has a section on Human Rights and Democracy. This unit administers a specialized human rights budget line. This targeted allocation is seen as an important tool to emphasize human rights priorities in Norwegian foreign policy. However, its allocations were found to be less than one tenth of the total funding for human rights and the Evaluation Department wanted to look more broadly in this study at those activities in Norway's development cooperation portfolio considered to have a dominant human rights aspect. Although the Ministry of Foreign Affairs had commented on and agreed to the terms of reference for this study, there

were changes in key personnel in both the Ministry and the Evaluation Department early in the process. The new staff on the ministry side cautioned that the scope was too broad. As a result, initial ownership of the evaluation within the Ministry was not strong, heralding later problems in obtaining a management response on it.

43. This one major issue was behind the report's nomination for this case study. Those interviewed repeatedly stressed that thematic evaluations such as this one are a major problem for the organization in that they cut across the areas of responsibility for the Ministry's various departments. Therefore, assigning and getting acceptance of responsibility for management follow-up on the evaluation is a major problem. However, no one argued that thematic evaluations should not be done or that they were not important to do.

44. As indicated, the Section on Human Rights and Democracy within the Department for the United Nations, Peace and Humanitarian Affairs clearly does have some responsibility for follow-up. But it was also suggested, for example, that the Section on International Development Policy in the Department for Regional Affairs and Development should take some responsibility as development cooperation projects should be rights based. To illustrate, a question might be whether a children's project in Kenya is adhering to the child labor convention to which Kenya had agreed. Indeed, the report finds a lack of clear human rights focus in development cooperation projects.

45. In addition to the various units within the MFA, Norad also has a role in the human rights area. In mid-2004, the responsibility for state-to-state official development assistance transferred from Norad to the Department for Regional Affairs and Development, but Norad, among other responsibilities, continues to fund Non-Government Organization (NGO) activities in developing countries. NGOs are key in-country implementers of development projects. Norad has a special human rights advisor who offers technical advice when requested. In sum, Norwegian support for human rights is given through many budget items and channels, with different goals and guidelines and assignment of overall responsibility has no easy solution.

46. The Evaluation Department submitted its follow-up memo on January 17, 2012 to the Ministry of Foreign Affairs. Specific use of the report findings is not yet under discussion as even after much public debate, no unit sees that it can or should take the major responsibility to develop an action plan for the recommendations. As of October 2012, the report had bounced around departments and sections for about 6 months. Despite some very high-level meetings, agreement had not yet been reached as to which unit should be assigned the follow up. With these thematic reports several sections often have some kind of interest, but they do not think they have the whole responsibility (or authority) and are unwilling to take it.

47. The Section for Management of Subsidiary Agencies and Development Funds in the Ministry of Foreign Affairs has responsibility for Norad (where the Evaluation Department sits) and so follow-up assignment on evaluations is a routine part of their responsibilities. The Section believes it spends too much time trying to negotiate and place the responsibility for these types of thematic studies. Vacancies and turnover of personnel also play a complicating role.

48. More broadly, a strategy under discussion for the future is asking the Evaluation Department to obtain ownership for follow-up responsibility before the evaluation is started, or even as part of choosing issues to evaluate. Some perceived the topic of human rights as not a high priority topic for Norway at this time and this may be some of the motivation for the suggestion. While relevance and ownership are key issues for evaluation and evaluation use, we heard some acknowledgement that obtaining this upfront ownership for thematic evaluations might not be possible. Certainly it could limit the ability of the Evaluation Department to plan and undertake thematic evaluations in the future because either they cannot obtain upfront ownership or the time spent trying to obtain ownership becomes prohibitive and the study idea is dropped. As discussed in Chapter 4, ownership is an important factor in promoting evaluation use, but in selecting topics for evaluation the Department also must consider other values, such as holding public agencies accountable and maintaining the independence to decide what areas are in need of evaluation.

49. On December 12, 2012 when work on this evaluation of use was substantially complete, the management response to the report was delivered. We noted the table format used with columns for recommendations, actions or rationales for no action, responsible entity, and timeframe made it easier to understand what would be done and not done. Recommendations focusing on preparation of guidance for human rights activities were accepted. Others calling for new procedures for promoting cooperation between MFA departments for development and human rights were not taken up, although closer contact and dialogue were supported.

50. We found a strong case of indirect use of the report outside of Management. The Norwegian Centre for Human Rights is both the Norwegian human rights National Institution and a multidisciplinary University of Oslo Centre. Its international programs are funded through agreements with MFA and Norad. The Centre had wanted to be evaluated for several years prior to the study. Although only two of their international programs were evaluated in the human rights study, the report has had impact on their overall program. They report that just the process of being evaluated was good for their thinking and planning. They found that they had to think more strategically about how they present their work—what they are doing and why they are doing it. Now they have a better framework—a results-based management framework for developing annual plans and activities. They note that they would likely be using a results framework regardless of the evaluation, but

the evaluation got them to start thinking earlier than they would have otherwise.

51. In sum, this report illustrates the importance of up-front buy-in and ownership of an evaluation, in this case a broad thematic evaluation, for later use of findings. It also shows promising use of a table format for report recommendations and management actions. And a final point is the report's illustration of indirect influence on others outside of program management.

C. Findings Case 2: *Pawns of Peace – Evaluation of Norwegian Peace Efforts in Sri Lanka, 1997-2009.* Report 5/2011

52. One interviewee commented that this is the only evaluation that MFA is really aware of. While this may be an exaggeration, the evaluation of Norwegian peace efforts in Sri Lanka had high visibility from the start—it was requested by the Ministry of Foreign Affairs—and the Evaluation Department was excited to undertake it and subsequently put considerable resources into it. Taking on an internal but independent evaluation of a political issue was a big decision for the Ministry in requesting it and the Evaluation Department in accepting to undertake it. Most evaluations focus on the performance and/or results of policies, programs, or projects for which there are more established methods. Before the study proceeded considerable discussion took place on whether this was an issue that could be informed by evaluation. The emphasis was put on learning from the evaluation. The Ministry had considerable input on the terms of reference for the study, although the Evaluation Department did not accept every one of its suggestions. The report remains controversial in some of its conclusions and recommendations, but almost all interviewed admit that at a minimum it made them reflect on the issues—an example of conceptual use wherein the evaluation changes understanding of an issue rather than leading to direct action. The report has been widely praised for the nature and extent of the dialogue that it provoked. Most praised the Evaluation Department's handling of the study as hands-on, open, inclusive, and transparent. The Ministry also received praise for studying their own process and exposing themselves to scrutiny.

53. There is evidence from notes and interviews that the study is an input in MFA decision-making about new engagements. We were told that it is being used as justification for preparing guidelines for future use in reaching decisions about Norway's engagement in future peace efforts.

54. A minority we interviewed said that the evaluation was not helpful and did not address the right questions. They disagreed with the report's conclusion that Norway should have withdrawn earlier from its role as mediator in Sri Lanka when the peace process became deadlocked. Their view was that Sri Lanka was a unique situation for Norway and not generalizable, and that Norway could not have withdrawn earlier when all in the international

community asked them to continue. Also it was important to have a party that could be an intermediary, even when the parties were going back to war. This minority also said that the report not only reached the wrong conclusions, but that it asked the wrong questions. We were told that the Ministry did not want an examination of Norway's role in Sri Lanka, but rather a limited examination of Norway's involvement in the political process—how Norway facilitated that process, how Norway maneuvered within its framework, and how it might have done better. The Terms of Reference for the study call for a focus on the role and performance of Norway from 1997-2009 as a facilitator of dialogue in the peace negotiations between the parties. However, the TOR presents some 36 questions to be addressed that yield a broad range of issues within this focus. A chapter in the report on aid and peace particularly generated controversy, but it is reflected in one of the TOR questions which asks: "to what extent was the broader Norwegian aid portfolio geared towards supporting Norway's role as facilitator of the dialogue?"

55. Some of those we interviewed argued that the report was too academic, not grounded in practical experience, and reflective of the authors' academic goals, framework, and ambitions. However, documentation shows that the report draft went through the Evaluation Department's rigorous quality assurance process. Moreover, it was viewed as a positive that the case would be taught in academic settings such as universities. We note that the Evaluation Department has no restriction on the authors of a report turning the evaluation into a book or other publication after its formal release and presentation, other than the obtaining of its consent and the consent of those interviewed to have their interviews used. It views this as leading practice. In general in the evaluation community, external dissemination in the form of presentations to evaluation associations, papers for journals, edited books and the like are encouraged and considered good practice.

56. There was some surprise that the report did not get more Norwegian media attention. The perception was that the report seemed to get more coverage from Sri Lanka newspapers than inside Norway. Also, we were told that the day the report was published, there was a shake-up in the Norwegian government that made media headlines and this may have taken some attention from the report. When the report was published a seminar was held with the Evaluation Department at an important Oslo location. The event was streamed for the public to encourage debate. A high level panel responded to the presentation. There was limited media coverage following the seminar. Additionally, the Team Leader for the study was invited by MFA's Section on Peace and Reconciliation to a lunch seminar in January 2012 to speak on lessons learned, but the seminar was lightly attended. He has contributed a chapter on the Sri Lanka evaluation to a forthcoming edited book that is devoted to evaluation methods for aid in conflict situations, and intends to write something more for a Norwegian journal, as well as perhaps participating in more discussions.[16]

---

16   "Evaluation Methodologies for Aid in Conflict." Eds. Ole Winckler Andersen, Beate Bull and Megan Kennedy-Chouane, Routledge, 2013 (forthcoming).

57. The Section for Peace and Reconciliation of the Ministry of Foreign Affairs submitted an action plan at the end of March 2012. On the one hand, they pointed out that the report already had been used actively for decision making in the Philippines where Norway has been involved. On the other hand, the note indicates that each peace process will be affected by the national and international context and that such processes are difficult to predict. At the same time they agree with the Evaluation Department on the benefits of a general strategic framework for these efforts, embedded in the MFA's Peace and Reconciliation Section. A possible Parliamentary report in 2013 on peace and reconciliation is indicated as a potential vehicle for developing this strategic framework, but its production could not be verified.

58. In sum, this case illustrates the political nature of MFA and MFA's interest in evaluation when it addresses an issue perceived as highly relevant. It also shows the challenges of designing and conducting an evaluation to address what is essentially a political issue. While this was a high risk evaluation, it was also high reward in terms of showing the Evaluation Department as transparent, inclusive, and open to comment, as well as in use of the evaluation findings.

D. Findings Case 3: *Evaluability Study of Partnership Initiatives –*
Norwegian Support to Achieve Millennium Development Goals 4 & 5.
Report 9/2010 – Study

59. This study had its genesis in the realization that baseline data, indicators, and reporting systems were not yet developed two or more years into implementation of Norway's five large programs aimed at achieving Global Millennium Development Goals 4 (to reduce child mortality) and 5 (to improve maternal health) through the Partnership Initiatives, threatening the ability to evaluate them in future years. The Department had produced a number of prior evaluation reports on other programs with the message of "not enough data are available" for analysis of program impact. Such report messages were perceived not to be helpful other than in providing an excuse for lack of evaluation findings. The problem identified was that the programs often lack adequate the correct baseline data, indicators, monitoring systems and processes, and regular results-based management (RBM) reporting and documentation. It was hoped that the RBM framework developed for the Partnership Initiative also would serve as a model for other programs.[17]

60. Ideally, program management develops a results-based management framework early as an annex to the legal agreement. The framework articulates the logic behind a program, defines the key outcomes the program is to achieve and identifies indicators for measuring progress to their achievement. Management collects baseline data on the indicators and uses that information to set realistic targets and dates for indicator achievement. Management then monitors progress on the indicators and

---

17   Ray C. Rist and Jody Zall Kusek. (2004). "Ten Steps to a Results-Based Monitoring and Evaluation System". World Bank: Washington, D.C.

makes program adjustments indicated. The frameworks facilitate before and after comparisons and provide data for evaluations.

61. It is a requirement in the Ministry of Finance's Financial Management Regulations that program or grant managers obtain reports from the recipient that make it possible to establish the degree of target achievement[18]. This means that the Embassies and Ministry Departments must take responsibility for assessing the extent to which program results have been achieved. Norad's Department for Quality Assurance works to promote good results frameworks, and using this requirement states that to do this objectives, indicators, and baseline data are needed. However, neither embassies nor Ministry departments are required to get comments on their results frameworks or to use a template that has been developed.

62. Some of the Norad staff became familiar with an evaluation approach called evaluability assessment. An evaluability assessment is a systematic process used to determine the feasibility of a program evaluation.[19] Its purpose, however, is not only to conclude whether the evaluation is to be undertaken or not, but also to prepare the program to generate all the necessary conditions to be successfully evaluated. This may involve steps such as clarifying the logic model, helping to identify key program outcomes and indicators, and the like. The Norad staff was enthusiastic about trying an evaluability assessment.

63. Conversation with the Evaluation Department generated support for doing an evaluability assessment as part of Norad's quality assurance and advisory roles. This program was launched under the leadership of the Prime Minister in 2007. The Partnership Initiatives take the form of bilateral cooperation agreements with countries with high child mortality rates, including India, Pakistan, Malawi, Nigeria, and Tanzania. The study team assessed in each of the five countries the extent to which the Partnership Initiatives could be evaluated in a reliable and credible manner, and made recommendations and proposed action plans for impact evaluations to be conducted at a later stage. It was hoped that with the study, Norway's embassies and the Ministry of Foreign Affairs would better understand the situation and put on more pressure for strong data.

64. Those we talked with indicated that the report quality was good and the message clear. To some, however, the use of the report has been disappointing. The study's Team Leader gave a seminar presentation on the report and there were open discussions on it, but few from the Ministry attended. Within twelve months, the Ministry of Foreign Affairs should have provided information on their uptake from the report. However, there has not been an action plan on this 2010 report. We were told the report suffered from the frequent lack of clarity on whose job it is to do follow up. The Ministry, embassies, and Norad are not clear about who has this

18   Ministry of Finance. Financial Management Regulations. Paragraph 6.3.6 (page 55).
19   M. F. Smith, *Evaluability Assessment: A Practical Approach*. Boston: Kluwer Academic Press, 1989

responsibility. We were told that embassies in particular are understaffed and existing staff are swamped. Building follow up into their role is unlikely "as each evaluation report adds to the burden." The Ministry throws back the question to Norad of how do we address this? Norad indicates it is up to the Ministry and the embassies.

65. Others said that the study was relevant, but that it did not give sufficient guidance to change what is being done. There was a steering committee in effect for the program with representatives from the Ministry, the Health Directorate, the Norwegian Knowledge Center for Health, and Norad. This group might have made decisions on how to proceed with the recommendations, but little came out of it. The steering committee is now being restructured.

66. Another issue that surfaced is the number of other players undertaking often comprehensive and sophisticated evaluations of the Millennium Development Goals, often at the same time. An example was given of a small Pay for Performance pilot that is part of the program, but funded with global monies. Norway indicated that it would take responsibility for process and impact evaluation, but found that USAID already had major funding for this evaluation activity and impressive partners.

67. Others indicated high use of the report. At least one Norad advisor reported that his unit has used the report to "market" their role as advisors on helping to structure these kinds of programs for later evaluation. The Quality Assurance Unit indicates that they have used this report more than any other evaluation—it is highly relevant to their work. They gave comments on the study and report draft and know it quite well. Examples of use include a training seminar on RBM that was conducted together with the Evaluation Department. The seminar included a presentation on the relationship between monitoring and evaluation. They also used the report to prepare for a grants management review at the embassies and to dialog with them. Finally, they have used it for discussions within Norad of weaknesses and room for improvement in evaluation.

68. The above focuses on Norad use and a question remains of use by the five countries involved in the Partnership Initiatives. There is a sense that the real actions that need to happen are in the purview of the embassy staffs in the five countries. The ability of headquarters staff to effect change was perceived as low. At least one closely involved Norad advisor pointed out that despite sitting with the people working on the different country initiatives and sharing the report findings, in the end the goal of helping embassies improve their monitoring and evaluation was not achieved. The key is increasing ownership at the country level. As one Norad interviewee stated: "Work to make the situation better has to be at the country level. It cannot be done from here."

69. In India where monitoring and evaluation were considered to be more pressing because of a mid-term review, the report did elevate the discussion. The report has been used subsequently in relation to the program planning of Phase 2 of the Norwegian Partnership Initiative in India. The Evaluability Study's recommendations for monitoring and evaluation were used to help set up a monitoring and evaluation framework for Phase 2. For India the timing of the Evaluability Study was good.

70. Nigeria's program had been delegated to DfID to manage. Nigeria has a large technical assistance program to build and support national information systems. Site monitoring is done to check that the data are of sufficient quality and local and state government systems are also being strengthened. Nigeria was used as a comparison best case.

71. Pakistan and Tanzania were major contrasts to India in terms of use of the Evaluability Study. In Pakistan, where the program was delayed while baseline studies were undertaken, it proved difficult to obtain the raw data in order to verify the baseline findings. (It later developed that the data had been discarded prematurely by the local consultants who had carried out the baseline surveys.) The study was not considered particularly useful there. In Tanzania, the study was viewed as not providing new findings, but simply confirming what was already known. For example, it confirmed that maternal mortality is not a good measure for Tanzania as maternal death is rare given the total number of births. In addition, when the Evaluability Study was carried out, the program elements were not yet started up in Tanzania. The timing was poor—it was not in tune with where the program stood in Tanzania. A final issue was that early on there should have been more consultation at the country level about what was really needed. As one informant stated: "If it had been country-specific in asking what was needed for a good evaluation, it might have been more useful. The study was more for Oslo's needs—not either the embassy's or the program's."

72. A theme running through this case was the role of the Evaluation Department. Points were made that for use of reports, there has to be more than a one-off seminar. Targeted briefings in country are needed. One suggestion was for the Evaluation Department to rethink its role and put more emphasis on marketing. For example, it could initiate on-line discussions and invite people to come and add their stories. Another related point was how studies are titled: a report titled "evaluability assessment" is not going to draw readers outside of evaluation professionals. Additionally, topics such as this one are too narrowly focused for broad interest.

73. In sum, the overall picture is of a report that has had more internal use than country use, but even then issues of unclear ownership limit results. In the one country where there was good use of the study, the timing of the study was excellent and met current needs. The case also illustrates the need for a marketing strategy for the evaluation report.

E. Findings Case 4: *Evaluation of the Norwegian Petroleum-Related Assistance - Case Studies on Mozambique, Bangladesh, East Timor and Angola.* **Report 1/2007.**

74. This report is on the topic of petroleum-related assistance which represents a long-term line-of-effort for the Evaluation Department rather than a "one-off" report on a particular program or thematic area. It is the only one of the four cases reviewed in this case study for which there was an action plan and a follow-up report on implementation of the action plan. However, it must be noted that the follow up action plan for the petroleum evaluation dated 2007 arrived only in 2011 (24/06) and the follow-up report arrived in 2012 (3/1). The 2007 report was followed by a 2009 management review of the Oil for Development program.[20] Norad's 2010 Results Report highlighted the 2007 report and discussed results and lessons learned in the sector from the study, illustrating with Bangladesh.[21] Another report related to the earlier effort, *Facing the Resource Curse: Norway's Oil for Development Program*, was in the final draft stage at the time of this case study effort. Two of the four cases followed in this 2012 study are the same as in the earlier Petroleum-Related Assistance study, which makes it highly relevant for this case study.

75. The 2007 evaluation examined Norwegian assistance to four selected countries in the early 1980s until July 2006. The primary case was Mozambique, with comparisons to East Timor (now Timor-Leste), Bangladesh, and to some extent Angola. The study focused on the results of the Norwegian assistance, the content and quality of the assistance, and the successes and challenges between Norway and the partner countries/institutions. The 2012 study re-examined Mozambique and Timor-Leste, among other countries.

76. Not surprisingly given the extent of the follow-up, those interviewed in connection with the 2007 report found it to be a useful and used report. They pointed to the relevance of its recommendations, especially its message on the need to give more attention to governance issues and that it is not about just running a concession well. Also viewed as important were the report messages that capacity building efforts require a long-term perspective, and also that if development is the goal, starting with countries that are petroleum "newcomers", or have new regimes, is advantageous. Petroleum-Related Assistance was described as a technical assistance project with a

---

20 "Rapport fra forvaltningsgjennomgang i Olje for Utvikling (OfU),"Oslo: Norad, 2009.
21 *Capacity Development: Building Societies Capable of Sustaining Themselves*, Results Report, Oslo: Norad, 2010.

development emphasis. The 2012 report specifically revisits and reviews the earlier evaluation as specified in its Terms of Reference.

77. Of course, not all 2007 report recommendations were implemented uniformly. The 2007 report covered work that preceded the Oil for Development program that started in late 2005, but the issues the report addressed were generally viewed as even more important as the size of the assistance grew tremendously over the years. The 2007 evaluation divided its recommendations into six parts. In our case study interviews, we heard some comments criticizing parts of the recommendations as being sweeping recommendations based on weak evidence and as often being too general and "politically correct". There was also call for recommendations to be more operational. We found Management's Action Plan in response to the report recommendations commendable for identifying specific follow-up actions and date they were to be achieved, and for assigning responsibility for achieving them. Nonetheless, it is difficult to track a one-to-one correspondence of response to recommendation because of the multi-part nature of the recommendations that is not closely paralleled in the action plan.

78. The same issue surfaces with Management's Follow-Up report. A table format listing the theme—for example, good governance or local ownership—and then numbering the specific recommendations, the action(s) planned in response to each (or none, if this is the Management's position), and actual action(s) taken would be helpful in assessing response.

79. Indeed, many of the recommendations in the 2007 report are not operationalized. For example, under the theme of Good Governance, one recommendation is that "Good governance should be pursued at the highest political level." This is one example of a recommendation where it is not clear what the action would look like to implement it. How a recommendation is implemented is up to Management, but what needs to be done—the action that is needed to satisfy the recommendation—should be clearly stated for it to be operationalized.

80. It is clear that many recommendations take more than 12 months to achieve. The 2007 report recognized this in recommending that there is discussion and analysis annually on deviations from implementation plans during the previous year and agreement on realistic action plans for the coming year. An example of a recommendation that could not be achieved within 12 months is the 2007 report's recommendation for embassy staff to work with management of the local partner institution to develop detailed descriptions of objectives, outputs, indicators, milestones, etc. quantified and set in time for the program. The 2012 report still indicates a weak results management and framework for the Oil for Development program to be able to report performance against targets consistently.

81. In sum, this case illustrates the need for recommendations to be operationalized and also for a multi-year annual framework for follow up on implementation of the recommendations. It shows value in an evaluation approach that is not "one-off", but rather a long term line of effort involving multiple reports.

F. Cross-Case Analysis of Comments

82. The set of interviews we conducted produced a rich trove of thoughtful comments related to the Evaluation Department—its independence, the quality of reports, its use of consultants, its processes such as Terms of Reference, and the like. They are summarized below. No attempt is made to provide frequency counts as the interviewees are not a representative sample. However, this section discusses themes that surfaced across two or more of the cases—a common form of triangulation.

Evaluation Department: Quality of Reports

83. Most reported that the quality of the reports was generally good. But there was a strong sentiment that they were not user friendly. The reports were characterized by many interviewees as too long, unnecessarily full of acronyms, too technical, and academic in style. Many reported that it is not clear for whom they are written. One interviewee commented: "These reports should be written so that my mom and son could read and understand them! The only acronyms that should be allowed are the UN and USA! Half of the civil service doesn't know the acronyms being used!" Several commented that very few people could be expected to read through any of the evaluation reports. There was a sense that they were now written for a small group of academics, professional evaluators, and respondents.

84. Many interviewees commented that they themselves did not have time to read reports and, at most, read executive summaries of relevant reports. There were comments calling for improvements in executive summaries, such as limiting them to 3-4 pages, written so as to engage the public and the media, and more broadly disseminated. Support was also expressed for short Policy Briefs, written in user-friendly fashion. In our four -report sample, the average length of the reports was well over 100 pages. Three of four reports needed two single-spaced pages to list acronyms used. Executive summaries were 5 or 6 pages. These findings support the call for more user-friendly reports.

85. There was some comment that recommendations were often too general and needed to be more operational. Analysis of the recommendations made in the four cases substantiated this point. (See Annex 3.) The analysis found that the recommendations presented by the consultants have tended to be of a quite general nature, in presented passive voice/tonality, not directed at anyone in particular, and for these reasons not necessarily easy to convert into action- and future-oriented tasks.

86. Another issue that surfaced was the need for broader dissemination and an over-reliance on the "one-off" formal seminar as the dissemination vehicle for evaluation reports. The seminars were well known and appreciated, but they were viewed as not sufficient for organizational learning, in terms of helping to make learning broadly available and readily accessible to those who need it to carry out their development responsibilities. The sense was that with the seminar, the consultants' formal role is finished and the report has been published as a Norad product. More extensive dissemination and promotion of the report, especially with in-country embassy staff, should be considered by the Evaluation Department. Many mentioned that it was the dialogue about the issues that made the evaluations particularly valuable.

87. The importance of these views is not necessarily that they are "true" in an objective sense. Rather, it is that at least some of our interviewees perceive Norad evaluations to have these issues, and such perceptions can influence whether and how evaluations are used. (Some of these issues, as well as others, are discussed in the review of the four reports undertaken as part of this evaluation, and discussed in Annex 3.)

Evaluation Department: Use of Consultants

88. There was a suggestion that the Evaluation Department needed to think outside the box and break away from using the small circle of development specialists. Ways need to be found, it was suggested, to galvanize the public. Several indicated that the Ministry was interested when the public was interested; otherwise true interest in evaluation was low.

89. Several mentioned what is in effect a "catch-22" in using consultants. The Evaluation Department told us that to maintain its independence it needs to use external expert consultants, selected using transparent international tendering processes, to carry out its studies. The experts appropriate for carrying out a given evaluation may be relatively few and often are academics who have published or lectured on the issue (or related issues) under study, suggesting they have a predisposed position. It is difficult for them not to have their own academic goals and to use the study to help advance them. It is not surprising given this situation that some respondents suggested that a report's conclusions could have been written in advance of the study. Evaluators are hired for their expert knowledge and research; this can mean they adopt theoretical perspectives seemingly set in advance or express opinions in reports without clear evidence behind them. One interviewee put it well: "it is difficult—if you know a lot on an issue, chances are you have a stake in the outcome." This is not to say that evidence is lacking for evaluation report conclusions. The Evaluation Department conducts rigorous quality control itself and through its broad invitations for review of the draft report, even if, in a few cases, they have had a difficult time assuring quality. It is to say that the prior involvement of the evaluators in the issues can lead to a perception—justified or not—that the evaluation reflects previously-held views and commitments, not just the evidence gathered during the evaluation.

The Evaluation Department's Processes

90. It is clear from our observations and interviews that the Evaluation Department takes substantial time to consult on what evaluations should be conducted. They confer with all the Departments, embassies, and Ministries and it is a long process. But in the end, several interviewees indicated a sense that the reports still may come at the wrong time and not be relevant. One put it this way: "At times I find the evaluations do not seem relevant. They seem to be at wrong time in the process or they don't address the questions that need to be answered. It is tough as issues and priorities are changing all the time."

91. Two of the Department's consultants argued that the problem of relevance was evident in how the Evaluation Department handles Terms of Reference. Too many and too detailed questions in the Terms of Reference were viewed as a serious problem contributing to reports eventually lacking relevance. One compared the 11 page Terms of Reference with 53 questions posed by the Evaluation Department for the study on which he was Team Leader to the Terms of Reference for the Committee on Terrorism which was 2 pages in length and a total of 3 questions. The four reports for our case studies actually averaged about 20 questions each—not 50. The message was that the Evaluation Department is too concentrated on technicalities and highly detailed questions and poses too many of them. The greatest difficulty is that some of the evaluators perceive a lack of flexibility in the questions: If it wasn't in the Terms of Reference it cannot be addressed and if it was in the Terms of Reference, it has to be addressed. This would mean that highly relevant promising issues that emerge in the course of the evaluation—and this is typical in the course of evaluation studies—cannot be followed up and questions that turn out to be of little interest or relevance, cannot be dropped. However, the Evaluation Department has told us that adherence to original Terms of Reference is a tool it uses to meet its responsibility for overseeing the quality of the evaluations it commissions. In particular, it is concerned that allowing too much latitude makes it hard to ensure that evaluations meet high methodological standards. Moreover, the evaluations are governed by procurement laws and regulations, which limit the degree of latitude the Department can accept.

92. Several of those interviewed also told us the Evaluation Department was getting too prescriptive on methods. So, for example, they want to do impact evaluations that use randomized control trials. This, they argued, drives the evaluation but sometimes the randomized control trial is on a narrow component of a project or program and does not reflect the most interesting questions. In fact, the Evaluation Department reports that they have done only one impact evaluation, and it was not a randomized control trial, but a quasi-experimental study, so the reasons for this concern are not fully explicable.

93. In general, the Evaluation Department was praised for reaching out and asking for input into Terms of Reference. Similarly, the Evaluation

Department's allowing of broad reviews of draft reports was believed to work well, and while perhaps lessening the appearance of the unit's independence to some extent, really worked to catch errors and improve reports. The Evaluation Department was consistently praised for being open and transparent and involving of others.

94. There is a clear relationship between these two observations. Having many stakeholders involved in determining what an evaluation should cover means that it is likely more questions will be raised. That, in turn, could require closer oversight of the evaluation itself, and thus more attention to the methods being used. There are trade-offs here that require a balancing act by the Department.

95. As noted above, the seminar as a vehicle for sharing the Evaluation Report was not viewed as particularly effective. The quality of the comments is not thought to be high, particularly if the report is released the day of the seminar so that few have had the opportunity to read it. The seminars are not particularly well-attended. Typically perhaps 80-100 attend the seminars on high interest evaluation reports, with 20-50 for those of more narrow interest. The Evaluation Department and consultants use their contacts to promote attendance, but seminar attendees tend to be disproportionately academics.

96. In several cases, the seminar was not held at the time the report was released, but later, in some cases much later. Because consultants have other obligations, it is not always possible to have a timely event. Under these circumstances, the seminar, as one might expect, often draws very low interest. The most serious consequence is that it also raises suspicion of the motives for the delay.

Evaluation Department - Independence

97. The Evaluation Department works hard to maintain its independence and to be seen as an independent unit. It uses consultants for its evaluations and views itself as in an independent quality assurance role. Evaluation reports are seen by the Department as the products of the consultants, although drafts are widely reviewed, and the Department challenges statements that do not seem evidence-based. The Department gives its opinion of the report and the recommendations to MFA in a formal memo following the dissemination seminar. We were told that the Evaluation Department sees it as outside their mandate to follow-up on recommendations. Their involvement generally ends with the writing up of the formal Memo to the Ministry of Foreign Affairs following the seminar. This has implications for promoting use of the evaluation. Those consultants who were Team Leaders for the studies told us that for the most part they have no idea how the study they led was used as their involvement ended with the formal seminar. The dilemma, raised by one former Team Leader, "is that if the unit is hands-off with the evaluations in their efforts to try to keep their independence, how can they promote use? And if they want use, they need to have more than a

one-off seminar." In addition, we were told that for many studies the actions that need to happen are in the purview of the in-country embassy staffs.

98. Promoting use of a report could be viewed as promoting its recommendations and that presents a dilemma for the Evaluation Department. Department leaders told us they believe the unit's independence would be compromised if it were to push for implementation of the recommendations. They argued that if they did follow-up on evaluative work it would be viewed as the Department evaluating its own recommendations. But not pursuing implementation of recommendations has its own downside. Getting ownership of the findings and recommendations has been a major problem for the Ministry. As stated by one informant: "You don't get measured on how well you follow up on an evaluation. Ministers have lots of other fires and pressures that they deem more urgent." Many referred to the political nature of MFA as a constraint on its ability to implement evaluation recommendations; however, this constraint is common to all evaluation of public policies, program, and projects, not at all unique to MFA.

99. Organizational placement of the independent Evaluation Department was discussed by our interviewees. On the one hand, some recalled that when the Evaluation Department sat in MFA, its placement was problematic. Being co-located in the Ministry was perceived as creating an appearance of conflict of interest for the Evaluation Department and potentially less hard-hitting reports. On the other hand, the current placement in Norad was viewed, and confirmed by this study team, to contribute to excellent access to and relations with Norad staff. Norad staff appreciated it greatly. But it was also viewed as not leading to ownership of the evaluations by MFA and as still leading to perceptions of lack of independence. As one informant stated: "Placement is a dilemma. MFA is responsible for Norad but Norad is not an independent entity and so how can the Evaluation Department be independent?" A few thought that if independence is the paramount issue, the Evaluation Department should report to Parliament, like the Office of the Auditor General.

100. Several encouraged "out-of-the-box thinking" for solutions to these issues. For example, the Department might consider producing a smaller number of reports, freeing up resources to better disseminate those reports to a wider audience. The interviewees claimed this would be a worthwhile tradeoff— that it was essential that the Evaluation Department reach out to broader community. They argued it was time that Department looked at options to do this kind of outreach.

101. Specific suggestions were to identify and address bigger picture issues that would guide future areas of development and to think more about marketing and appeal—at least for larger evaluations. Targeted briefings on the study findings—especially in-country—were mentioned as a good direction. Other suggestions were to perhaps try having a developing country do the

evaluation, to do an evaluation where the baseline is recreated, and to collaborate with another organization's evaluation unit on evaluations. The Swedish International Development Agency (Sida) was specifically mentioned in this regard. We note that the Evaluation Department did collaborate recently with Sida on the evaluation related to child rights, and that it also often uses developing country evaluators to carry out studies.[22]

102. A major issue continues to be that the organization still lacks a good standardized procedure for assigning follow-up responsibilities on evaluation recommendations. Across the Ministry, embassies, and Norad, it is not clear who has responsibility to take the lead on responding to recommendations. This was particularly a concern for thematic evaluations that cross-cut different departments and sometimes agencies.

G. Conclusions

103. Overall, the four cases provide a rich array of issues and themes related to the use of Evaluation Department's reports. Key among them are:

- The importance of promoting open, inclusive, and transparent procedures for the design and review of evaluations;

- The importance of early buy-in and ownership to later use of evaluation findings;

- The need for evaluations to focus on a limited set of well-defined questions, avoid acronyms, be short with strong executive summaries, and written so as to engage the public and the media in a more user-friendly fashion;

- Need for a marketing strategy for evaluations—even if it means fewer evaluations are conducted;

- Need for recommendations to be operational and for a standardized multi-year annual framework to follow up on them;

- Benefits of a line of effort resulting in multiple reports versus "one-off" reports;

- Importance of considering and promoting use of evaluation findings not only within the Norwegian government, but also with a broader community of indirect users;

- Challenges to an Evaluation Department's efforts to be independent of its own commissioned evaluations but still promote their use;

---

22   Presumably, some interviewees were unaware of these efforts.

- Difficulty of raising interest in evaluation in an essentially political institution such as MFA.

104. As a case study, these findings are illustrative. Had we selected four other evaluation reports to review the issues and themes identified might be different. Nevertheless, the case study raises important issues and themes that can be examined, considered, and perhaps validated through a survey approach.

105. The case studies presented in this chapter provide a set of rich insights into the extent to which and ways in which Norad evaluations have been used. It is clear that factors affecting use include timeliness, quality, and targeting. The role of the Evaluation Department in promoting use also comes into focus, especially its reluctance to promote recommendations actively. Some of the issues raised in these case studies are reflected in the broader view presented in Chapter 4, so a fuller discussion of conclusions is presented in Chapter 5.

4. Findings on the System for Evaluation Use

While information on the use of individual evaluations is instructive, this study also examined how well the overall evaluation system works to improve evaluation use. This analysis takes account of the broader context in which the evaluations are conducted, the structures within which the evaluation function operates, and the varied uses to which evaluation may be put. The conclusions come from triangulating data from document reviews, key informant interviews, and the survey described in Chapter 1.

A. Instrumental Use: Formal Responses to Evaluation Recommendations

106. The most obvious kind of instrumental use is in adopting evaluation recommendations. As noted in the Chapter 3 on the case studies, Norad has a well-articulated process for responding to evaluation recommendations.[23] When a draft evaluation report is submitted from the consultants, it is circulated for comment to relevant stakeholders following a rigorous quality review by the Department. Once a final version is completed and approved, the Evaluation Department arranges for publication of the report and organizes a seminar for the publication date selected. Following the seminar, it then writes a follow-up memo summarizing the report and seminar comments and sends it to MFA. Additionally, the Evaluation Department indicates in the memo which report recommendations it supports. This generally is done within 2 weeks of the seminar, but sometimes for various reasons it takes longer.

107. The Ministry of Foreign Affairs is responsible for preparing a formal response to the findings and recommendations. This is to be submitted within six weeks of report submission to the Ministry by the Director of Evaluation through the Director General of Norad. Where specific actions are called for, the Ministry must prepare for the Secretary General of MFA an action plan outlining what actions are to be taken. After one year, there is expected to be a follow-up report by the Ministry on what actions were taken. The Evaluation Department does not carry out its own follow-up review of actions taken in response to the evaluation, however.

108. That is the formal process. However, practice does not always follow this process. For this review we mapped the response to evaluation reports in order to examine data on the extent to which evaluation plans were

---

23   OECD, 2010, p. 37.

developed and follow-up reports submitted. Overall, there were 47 evaluations completed between 2006 and 2011, of which follow-up memoranda for 41 have been produced. Among these we found that 37 had recommendations that required response from program officials. However, we could identify only 28 for which action plans were available, and 4 for which there was a decision that no follow-up was needed. Of those 28, the Evaluation Department reports having follow-up reports for only 19. In part, this discrepancy reflects the lack of action plans or follow-up for the 2011 cohort of evaluations at the time of data collection. Nonetheless, even taking that into account, the Evaluation Department reports that of 35 evaluations completed during 2006-10 there were 32 follow-up memoranda, leading to 23 follow-up plans (plus the 4 for which no plan was needed). Of the 31 evaluations over that period for which there should have been follow-up reports, there were only 19. This disconnect between the formal process and actual practice is recognized within Norad and MFA. The Evaluation Department has taken steps to address this issue. One initiative it has undertaken involves creating a table to track follow-up actions on evaluation recommendations. Initial feedback from MFA appears to have been favorable, but it is too early to determine whether this will increase actual compliance with the formal process.

109. This is not to say that evaluation recommendations are not taken up by the responsible offices. In Chapter 3, for example, we saw that recommendations in the evaluation of petroleum-related assistance were taken on board. Thus, the picture hardly is one where evaluation recommendations are largely ignored. Rather, the issue here is that the system for responding to evaluation recommendations, while robust in principle, is somewhat lacking in practice. In part, this may reflect something about how the system is designed.

110. In that regard, we examined practice in other development organizations similar to Norad. Perhaps the nearest parallel is DANIDA, in which the Evaluation Department reports to the Minister for Development Cooperation through the State Secretary for Development Policy. As with Norad, the Evaluation Department is independent of the management of development work, and relies on external consultants to carry out its work. DANIDA has developed a system that requires a follow-up memorandum that is discussed at the Programme Committee, out of which comes a management response. Crucially, a Quality Assurance Department in the Ministry—not the Evaluation Department—is responsible for monitoring the implementation of follow-up activities on a regular basis. That department checks and reports on what Management has done in response to recommendations. This system is relatively new, so long-term effectiveness cannot be established, but the initial report found nearly 100 percent compliance with the requirement, according to an official with DANIDA.

111. On the other hand, practice at the Policy and Operations Evaluation Department (IOB) of the Netherlands Ministry of Foreign Affairs is quite

different. IOB reports go directly to Parliament through the Minister, with a policy response attached. However, IOB does not track follow-up actions, which is seen as a matter for the operational units of the Ministry. At one time there was a requirement that the relevant Department develop an action plan with follow-up by the Audit Committee, but that proved ineffective, according to our interviewees. What seems clear is that, at least in the cases of Norad and IOB, the formal requirement for follow-up is insufficient to ensure that it happens. This leads to a consideration of what factors may inhibit adoption of evaluation recommendations.

B. Factors that Inhibit Adoption of Recommendations

112. In interviews with MFA staff, the survey, and document reviews we identified several reasons for less than full compliance with the formal process of response and follow-up. As figure 2 shows, nearly three in five survey respondents indicated that in the most recent evaluation report on their unit's activity the recommendations were useful overall, while only one in five said they were not relevant. Nonetheless, only about one-third reported that they had adopted many of the recommendations. What explains the apparent discrepancy between the overall view of recommendations and their actual adoption?

Figure 2: Recommendations in most recent evaluation report (n=129)[24]
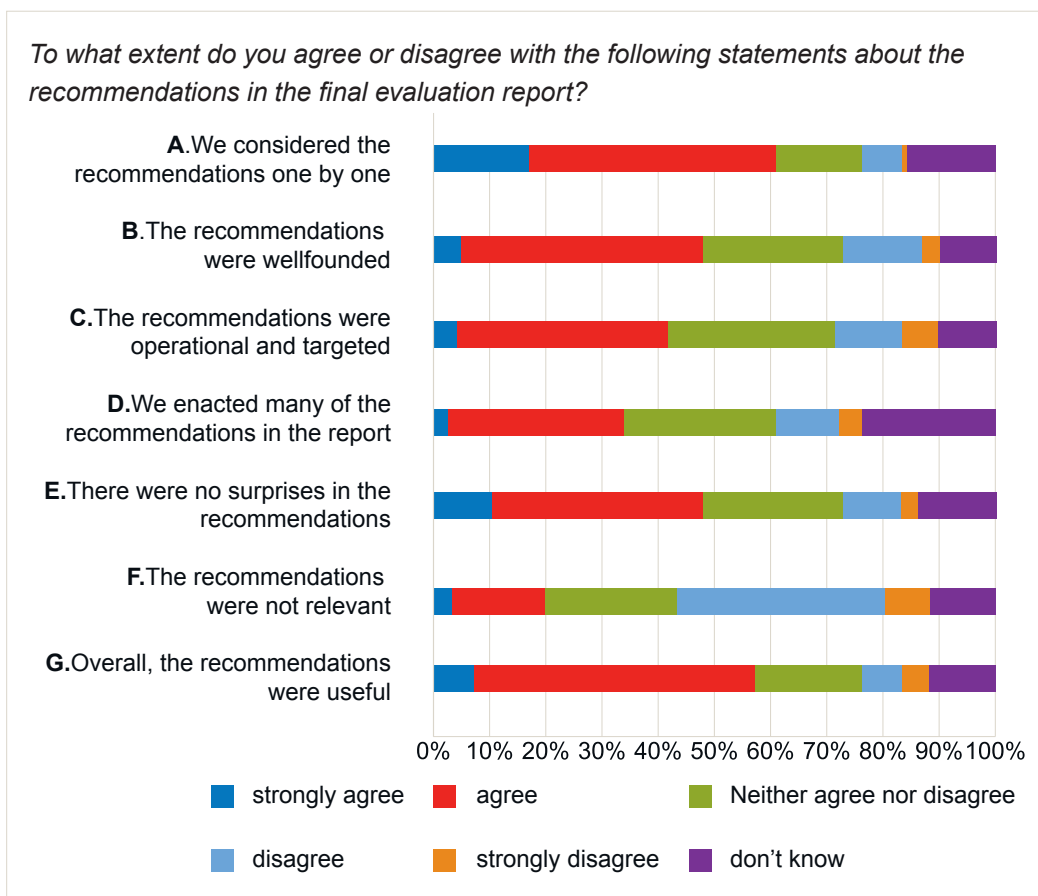
24   Indicates for each figure the number of respondents who answered that particular question.

113. One major reason, highlighted in interviews, is that many reports cut across several areas of responsibility or multiple countries. For example, in the case of the evaluation on human rights, the responsible units for countries were in regional departments, while thematic responsibility lay with the of UN, Peace and Humanitarian Affairs, and the policy issues were in the domain of a number of units. In this complex environment, accountability for acting on the evaluation's findings and recommendations was diffuse, so it was not clear which unit would need to take what actions, nor could any be held accountable.

114. Second, reports sometimes are no longer relevant at the time they are issued. Such reports may be seen as backward-looking, essentially historical, rather than a guide to future actions. The survey findings do not indicate that this is the major reason for non-adoption (only about 20 percent of respondents cited this as an issue), but it did emerge as a problem in a few cases.

115. Third, the survey, informants, and our own analysis of evaluation reports agree that recommendations sometimes are not well-developed or supported. A common observation is that recommendations are made without an understanding of the context in which they will be considered. Fewer than half the survey respondents found the recommendations well-founded, and only a bit more than four in ten described them as operational and targeted. Especially troubling are recommendations that call for expending more resources to fix identified problems without taking account of competing demands and broader priorities. So, part of the problem with adoption of evaluation recommendations may be that they are not adequately supported in the evaluation or are not sufficiently operational to be used. This implies a need for more work on how recommendations are developed in the evaluation process.[25]

116. Finally, in key informant interviews we were told that there is no culture within the system for saying "no" to recommendations, even if they are not seen as relevant or practical. Instead, those responsible for taking action may simply agree in principle, but then take no follow-up actions.

C. Factors Promoting and Inhibiting Broader Uses of Evaluation

117. Of course, evaluations are not just about developing and implementing recommendations, as important as that may be. As outlined in chapter 1 and discussed in the case studies in chapter 3, evaluation can be useful for a wide range of purposes. Other forms of **instrumental use** also are important. For example, an evaluation of Norwegian Centre for Democracy Support found that it was not achieving its objectives.[26] This led to a reorganization of the way such aid is provided.
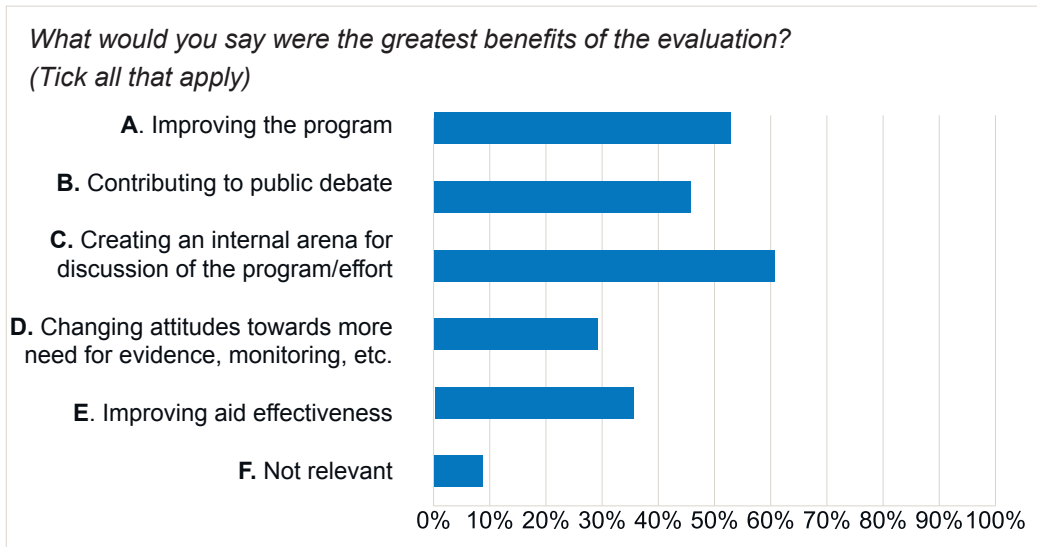
---

25   See Michael Hendricks and Elizabeth A. Handley, "Improving the Recommendations from Evaluation Studies," Evaluation and Program Planning 13:2, 1990, pp. 109-17.
26   Evaluation of the Norwegian Center for Democracy Support, 2002-2009, 2010.

118. Figure 3 shows how Norad, MFA, and embassy survey respondents reported on the range of benefits they saw from the most recent evaluation in their units. The response categories broadly reflect the typology introduced in Chapter 1. More than half the respondents cited program improvement as a major evaluation benefit, an example of instrumental use.

Figure 3: Evaluation benefits (n=127)

What would you say were the greatest benefits of the evaluation?
(Tick all that apply)

A. Improving the program	
B. Contributing to public debate	
C. Creating an internal arena for discussion of the program/effort	
D. Changing attitudes towards more need for evidence, monitoring, etc.	
E. Improving aid effectiveness	
F. Not relevant	

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

119. But while most of the focus in evaluation tends to be on instrumental use, **conceptual use** often is more important, even if the resulting actions are not what the evaluation recommended. An even larger share of respondents— about 60 percent—cited creating an internal space for discussion of programs as a benefit of evaluation. Key informants provided instances of this kind of use. For example, the evaluation of providing aid through non-governmental organizations (NGOs) was credited with engendering dialog on how to work with such groups effectively.[27] The evaluation of Norway's development work in Afghanistan has been a particularly acute example of how evaluation can help to redefine issues in a broader context.[28] It was referred to specifically in the Foreign Minister's statement on Afghanistan to Parliament in May 2012, which cited its main findings and also noted the difficulties of measuring results highlighted in the report. In addition, the report was picked up by the media, including newspapers and television, and apparently helped spark calls for a more thoroughgoing assessment of Norway's role in Afghanistan.

120. More generally, the survey also found relatively strong support for the proposition that evaluation reports are an important source of knowledge, relative to most other sources (see figure 4).[29] Indeed, nearly eight in ten respondents indicated that evaluation reports are important sources of

---

27   Results of Development Cooperation Through Norwegian NGOs in East Africa, 2011.
28   Evaluation of Norwegian Development Cooperation with Afghanistan, 2001-2011, 2012.
29   All figures in this chapter come from "A Summary of Results from the NORAD Survey: How do you engage with evaluation knowledge in your day to day work," November, 2012.
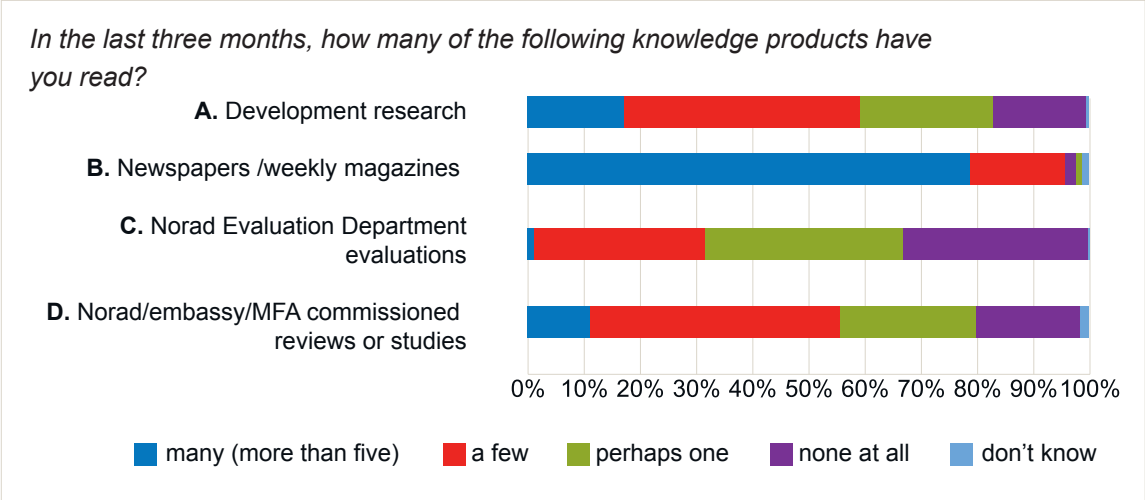
information for acquiring new knowledge about their areas of work, about equal to research and policy studies. Only informal discussions with colleagues and friends rated slightly higher among respondents. This confirms the potential for evaluations to contribute significant new knowledge to development organizations.

Figure 4: Tools for acquiring knowledge (n=261)

In your opinion, how important are the following tools for acquiring knowledge about what works in your field?
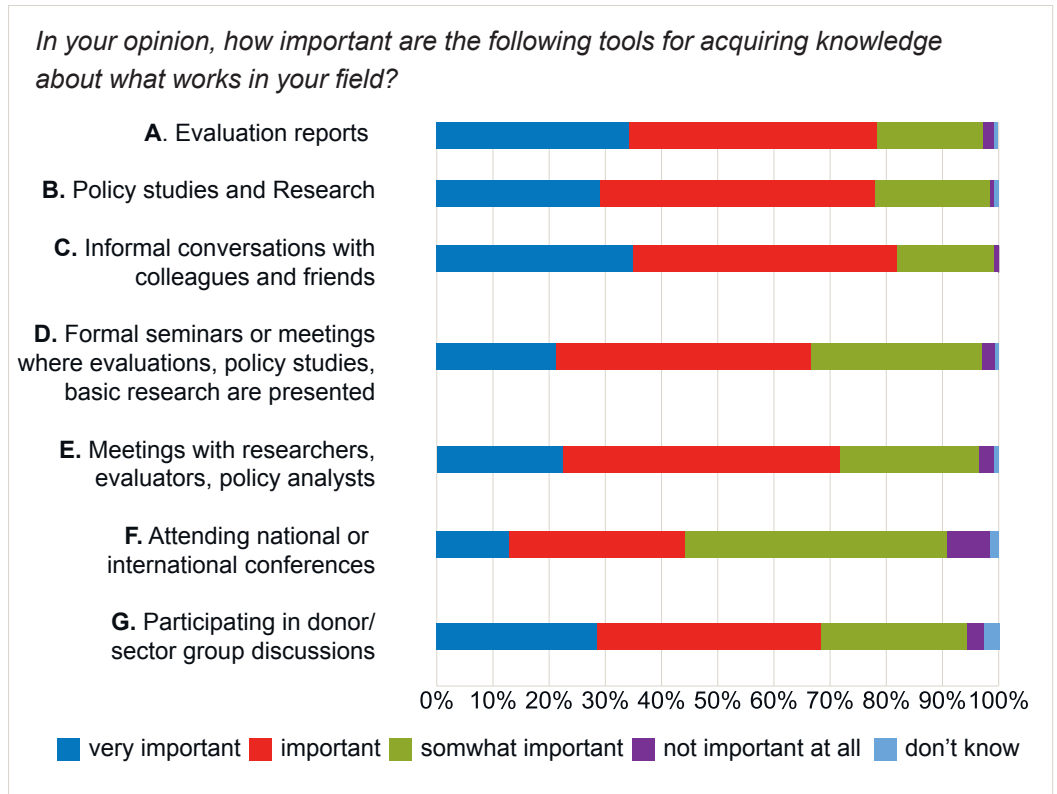
121. However, despite these high ratings on evaluations as a source of knowledge, the reports are not as well-read as other sources of development knowledge, as shown in figure 5. In part, this may reflect the fact that the question asked respondents to consider only the previous three months. It is not likely that the Evaluation Department produced a report on a topic of professional interest for most respondents during that relatively short period. But that likely is true for the other kinds of reports shown, except for newspapers and magazines. This suggests some disconnect between recognition of the value of evaluation reports and actual use.

Figure 5: Use of knowledge products (n=262)

In the last three months, how many of the following knowledge products have you read?

A. Development research

B. Newspapers /weekly magazines

C. Norad Evaluation Department evaluations

D. Norad/embassy/MFA commissioned reviews or studies

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

■ many (more than five) ■ a few ■ perhaps one ■ none at all ■ don't know

122. One reason for this disconnect may be seen in figure 6. There it becomes obvious that time is a major constraint. The more relevant and high quality the report is, the more likely it is to be used. But that clearly depends on staff having the time to read evaluation reports in the face of multiple and conflicting pressures. Indeed the strongest source of support for using evaluations is personal incentives. This is a key point because it highlights that in order to get staff interested in becoming familiar with and using an evaluation it is important to recognize and respond to their needs on a fairly personal basis. In particular, providing enough benefit to overcome their time constraints is particularly important from this perspective. This implies a need to ensure both that evaluations address issues that are important and compelling enough to make the time investment worthwhile, but also to find ways to make the findings, lessons, and recommendations from evaluation accessible in a variety of formats that meet various users' needs, as discussed below.

123. Another negative influence in some cases is the quality of the evaluation reports themselves. Figure 6 shows that about one-quarter of respondents identified quality as a factor inhibiting use. We examined how some forms of evaluation communication and dissemination affect use. One part of this was an analysis of the reports themselves to examine how effectively they communicate their messages. To do this we did an in-depth examination of the reports on the evaluations covered in the case studies. A full discussion of this analysis is found in Annex 3.

Figure 6: Factors supporting or inhibiting use of evaluation (n=254

To what extent do the following factors act to support or inhibit the use of evaluations/research in decision making/to improve your work?

Legend:
- strong driver for use
- moderate driver for use
- slight driver to use
- don't know
- slight barrier to use
- moderate barrier to use
- strong barrier to use
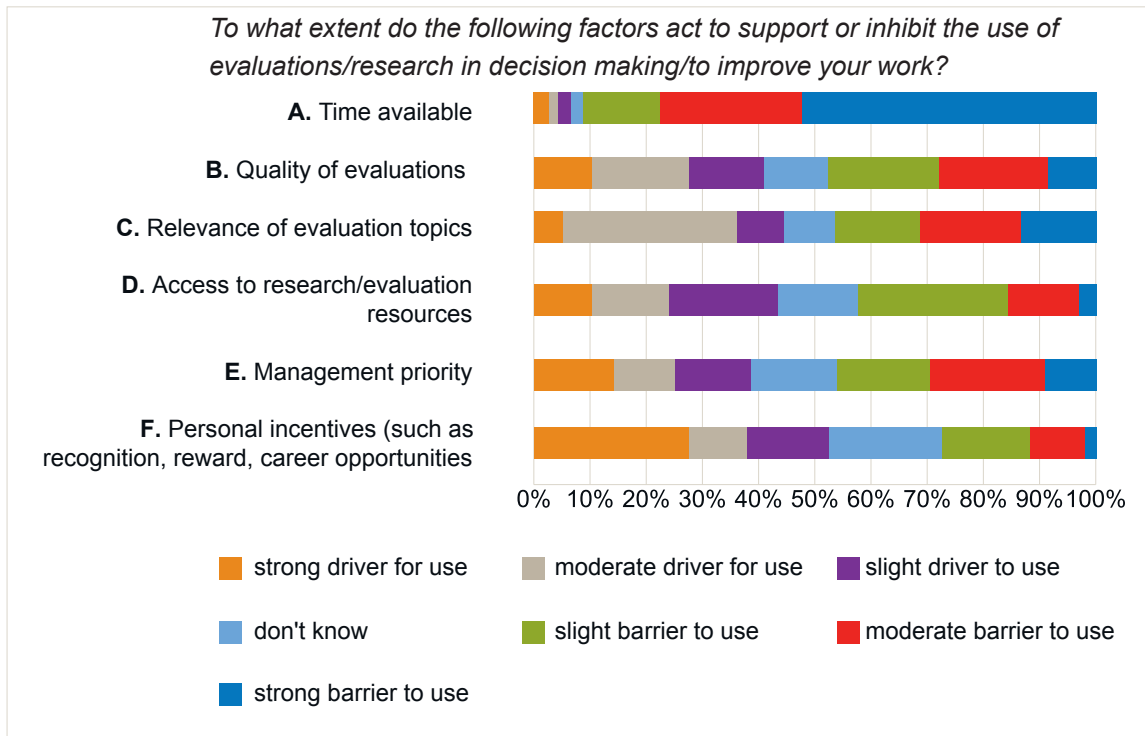
124. The major findings from that analysis were that the reports as communications products could be improved in easy but important ways. One is to avoid overly-long reports. Recall from the survey that one of the major impediments to reading and absorbing evaluation reports is time. Long reports exacerbate this problem. Perhaps more importantly, it can be helpful to readers to provide easy-to-find markers that highlight the key points, so that even busy readers who cannot read the entire document are able to find the important information they need. Some tools that have been used include side-bars or marginal highlights that make it possible to follow the major argument of the report without reading the main text. Alternatively, highlighting the key points through the use of bolding or text coloring can serve the same purpose.

125. Another use for evaluations is **enlightenment**, which is the propagation of knowledge gained through evaluation to communities of interest outside the relevant decision-making units of government. One way to promote such use is by making evaluation reports broadly available. The Evaluation Department publishes all its reports on its Website. Because reports are published in English, they are accessible to a wide international audience.

126. As part of this study, we conducted a review of coverage of Norad evaluation reports, both within Norway and internationally. We found that most references to Norad reports were on non-media Websites, particularly those associated with research organizations and NGOs. A good recent example is the forest initiative evaluation, which has been discussed by the

World Wildlife Fund and picked up on international Websites relevant to the issue.[30] In fact, these can be important audiences for Norad evaluation, perhaps more than those reached by general circulation media, such as newspapers. That is because the sites tend to be focused on substantive issues in evaluation (e.g., themes such as gender and governance), or research and evaluation methods.

127. Still, there is not a great deal of public coverage of evaluation reports. This is not unique to Norad; most development organizations face the same limitations on coverage. But there are things that the Evaluation Department could do to get consistent media coverage of its reports. For one, it would have to develop relationships with reporters and editors. Currently, such public communications are handled through the Norad communications office, not the Department itself. There are examples of evaluation units adopting an active media presence to enhance the resonance of their messages, but this comes with clear risks, both in terms of heightened public scrutiny and perhaps controversy, and of negative internal reactions to what might be perceived as public relations rather than dissemination.

128. We also found evidence of **process use**, that is, learning through participating in the evaluation process itself. Figure 7 shows that many Norad, MFA, and embassy staff have participated in the last evaluation involving their units. Participation was heaviest in the late stages of the process, but four in ten respondents reported being engaged throughout, which is regarded as good practice. Figure 8 shows that about 40 percent of those subject to evaluation learned to think more evaluatively as a result, a good example of process use.

Figure 7: Participation in the evaluation process (n=132)

Thinking about the same evaluation, to what extent were you involved at different stages in the evaluation process?

A. At the beginning of the evaluation process (e.g. in development of the ToR/inception report)

B. Throughout the process (e.g. as you interacted with the evaluators)

C. Towards the end (e.g. as you received the draft final report/seminar)

D. After the process is completed (e.g. as the follow up memo was received)

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

■ a great deal ■ somewhat ■ a little ■ not at all ■ not relevant

---

30 *Real-Time Evaluation of Norway's International Climate and Forest Initiative*, 2011; and the follow-on *Lessons Learned from Support to Civil Society Organizations*, 2012.

Figure 8: Uses of evaluation (n=131)

With regard to the same experience, did you make use of the evaluation evidence to do any of the following?

A. To change, confirm or modify an existing programme/intervention (i.e. to influence other people to change, confirm or modify an existing programme)

B. To for ex. influence a new programme design

C. To learn to think evaluatively (for ex. to more critically examine the programme/intervention logic, indicators and/or alternative ways of doing things)

D. To understand a programme/intervention in a new way (for ex: become aware of its strengths and weaknesses)

E. To acquire broader learning (i.e. when moving on to future programmes/ new areas of work/new colleagues

Legend: yes, to a large extent | yes, to some extent | yes, only to a limited extent | no, not at all | not relevant

129. Finally, evaluations may have **symbolic use**. The mere fact of the evaluation creates its own value. A number of key informants and case study interviewees validated this use, without being prompted. One argued that the simple fact that an independent evaluation of an activity takes place is an important source of discipline for the entire system. Others noted that they found evaluation useful as a way to validate their work. This is an often underappreciated function of evaluation.

D. Structural Factors in Evaluation Use

130. We also examined some of the structural issues that could affect evaluation use. Here we also drew comparisons to some alternative models used in other development organizations.

131. As noted above, the Evaluation Department is housed in Norad, but reports to MFA. Nominally, the Department's Director reports directly to the Secretary General, but in practice the reporting is to a Deputy Secretary General. It is not clear that the Evaluation Department's work ranks high on the priorities of this official, given many other responsibilities. He and the Director do not appear to have regular meetings, for example.

132. There are some advantages to the current arrangement, however. It does allow Evaluation Department staff ready access to Norad staff, in what appears to be a generally constructive working relationship. At the same time, it provides opportunities for the Evaluation Department to inform Norad staff on findings. Several key informants cited these exchanges as an important form of influence because of effects on how development issues

are conceptualized and issues framed by program staff as a result. One example cited specifically was the work on higher education policy.

133. Nonetheless, the Department does occupy a rather ill-defined position within the overall organizational structure. Based on our observations and interviews, this anomalous positioning does not provide a strong base from which the Department can promote use of its evaluation products. The Ministry staff tend to see it as part of Norad, while the Norad staff note that it actually reports to the Ministry. In neither organization is there a strong incentive to pay close attention to the Department's work, as evidenced by the lack of consistent follow-up to evaluation reports.

134. There are alternative models. Sweden has experimented with a fully autonomous evaluation unit, the Swedish Agency for Development Evaluation (SADEV), mandated to evaluate all development cooperation. SADEV reported directly to the Government through the MFA, but is institutionally independent. However, its autonomy contributed to its undoing. It did not gain traction within the governing system, and its operations were discontinued at the end of 2012.[31] The Swedish International Development Agency (Sida) does have its own evaluation unit, however, which will continue operations, though these cover only Sida activities, whereas SADEV had a broader mandate.

135. Another model has been developed for the United Kingdom. There, much of the evaluation work has been taken over by the Independent Commission on Aid Impact (ICAI), which is completely outside the structure of the Department for International Development (DfID). ICAI commissions evaluations and expects to do checks on follow-up on a regular basis. But the organization is new so it is too early to tell how well this will work. DfID itself has acted to embed evaluation deeply into its processes. The business case for DfID projects requires a specific evaluation plan.

136. There is no one perfect model. Any organizational structure for evaluation requires balancing independence with relevance, since the more distant from the decision-makers the organization becomes, the harder it is to be sure what kind of information is needed and when. That said, the current arrangement of Norad's Evaluation Department between Norad and MFA implies some trade-offs that may not be optimal.

E. Learning from Evaluation

137. In general, evaluation plays two distinct roles in development organizations: providing accountability for achieving results, and promoting learning from experience. These roles most often are seen as complementary. Thus, an OECD-DAC synthesis report argued: "this is not an either/or situation. Both accountability and learning are important goals for evaluation feedback. Of

---

31   Beginning in 2013, a new organization, the Expert Group for the Evaluation and Analysis of Swedish International Assistance, was created, but this took place after the work for this evaluation was completed.

the two, learning is the area where agencies recognise the greatest challenges."[32] But others take a different view; Serrat, for example, concludes from an analysis of the different purposes, modalities, and uses of evaluation: "the two basic objectives of evaluations—accountability and learning—are generally incompatible."[33] And, Serrat argues, because parliaments, treasuries, taxpayers, media, and interest groups tend to prefer an accountability focus for evaluation, learning often is of secondary concern.

138. Thus, using evaluation for learning can be challenging. Norad has been concerned about how to improve learning from evaluation for a long time; a major study was commissioned 20 years ago.[34] A major finding from that study was that staff "rank learning from project reviews highly, but learning from evaluations does not appear to be so significant." Figure 4, above, implies that evaluation is more highly valued as a source of learning now than it was in the early 1990s.

139. In a review of recent literature on learning from evaluation and research to affect policy, Jones and Mendizabal found that such use is strongest when four conditions are met:

- Questions are set by users, relevant to their needs, answered when needed, and credibly answered.

- Credible outputs are packaged, formatted, disseminated for ready use.

- The process includes face-to-face meetings with producers and users of knowledge, with engagement throughout and after the research process.

- Intermediaries, brokers, and networks are developed to support dissemination.[35]

140. As noted in previous sections of this report, Norad's Evaluation Department conducts extensive consultations to determine its work program, and to design the Terms of Reference for individual evaluations. It also employs a robust review process to ensure the quality and credibility of responses. In some cases, timeliness has been an issue, although this is a common concern across development evaluation generally.

---

32   OECD-DAC, *Evaluation Feedback for Effective Learning and Accountability*, Paris: OECD, 2001, p. 11.
33   Olivier Serrat, Learning from evaluation, Washington, DC: Asian Development Bank, 2010, p. 3.
34   Knut Samset, Kim Forss, and Otto Haughlin, A Study of the Feedback from Evaluations and Reviews in Norwegian Bilateral Aid, Oslo: Norwegian Ministry of Foreign Affairs, 1993.
35   Harry Jones and Enrique Mendizabal, Strengthening Learning from Research and Evaluation: Going with the Grain, Overseas Development Institute, 2010.

141. Where there may be room for improvement is in the other areas. As noted in several of the case studies, Norad evaluation reports often are long and highly technical. This kind of report does not lend itself to ready use. As discussed elsewhere in this report, there are steps Norad could take to improve on the form of presentation so as to encourage more use.

142. In addition, experience in other development organizations points to the value of synthesizing information from a variety of studies on related problems. This can be a good way to package knowledge developed through evaluation in a more digestible form for potential users. Beyond that, the process of synthesis often leads to new insights and knowledge that the individual reports cannot provide. The Department has done several such synthesis reports.[36]

143. More regularly, it summarizes evaluation lessons from across each year's evaluation studies in its Annual Reports.[37] The one- or two-page summaries of key learning from each report are potentially useful products, but packaging them in a report that cuts across many sectors and regions makes it less likely that the information will reach audiences with specific information needs. But experience in a number of development organizations suggests that these individual pieces could be extracted from the Annual Report and used as free-standing evaluation notes on the Department's Website, where they will be more readily found by target users. A number of development organizations, including the World Bank, IFAD, and others, already provide such brief notes as a way of supporting learning from evaluation.

144. One of the key messages from Jones and Mendizabal is that most learning takes place in face-to-face communication. In the survey done in collaboration with this study, nearly nine of ten Norad respondents cited informal discussions with friends and colleagues as a frequent form of knowledge sharing. In addition, three-fourths mentioned division or team meetings. By contrast, only about one in nine named intranet discussions. This highlights the importance of face-to-face dissemination for propagation of learning from evaluation. The Evaluation Department's seminars certainly are one way to conduct such in-person dissemination, but as this report notes these are not always effective in reaching key audiences, such as Norad and MFA decision-makers. This suggests that the Department may need to explore other avenues to conduct in-person dissemination of findings. For example, the Department might wish to make presentations to formal meetings of relevant MFA or Norad units that bring the results of evaluation directly into a forum where the key internal audience is available.

145. Perhaps even more important is the need for communication throughout the evaluation process. As in most organizations, it is difficult to maintain regular

---

36   See, for example, Espen Villanger, *Cash Transfers Contributing to Social Protection: A Synthesis of Evaluation Findings*, Oslo: Norad, 2008.
37   *Evaluation of Norwegian Development Cooperation—Lessons Learned in 2011*, Oslo: Norad, 2012.

communications between evaluators and the units whose work is being evaluated. Typically, such communication takes place at the start of an evaluation, episodically during data collection, and then when draft reports are being reviewed. However, research supports more continuous interaction as a more effective way to promote learning. A conundrum for the Evaluation Department is that evaluators are external consultants who normally would not have regular interactions with Norad and MFA staff. That puts the burden on the Department, but there is a concern with maintaining independence that can inhibit such regular communication. Still, promoting learning could require additional efforts along these lines.

146. Finally, the literature supports the use of intermediaries to help get out the key learning from evaluations. The Department already does a number of things along these lines, including participation in workshops, conferences, and international meetings. But these generally are episodic events. There is no articulated strategy for developing and using external networks and social media to promote the learning from Norad evaluations.

F. Conclusions

147. Overall, the Norad Evaluation Department has worked hard to promote the use of and learning from its evaluations, and to provide recommendations that can improve the effectiveness and operations of Norwegian development assistance. However, its efforts are limited by a number of factors. Some of these, such as the quality and realism of recommendations, it can address directly. Others it can influence through its actions but cannot control directly, such as its efforts to encourage greater compliance with follow-up processes by operational staff. Still others are beyond its control; this includes the organizational structure, which may limit its effectiveness. Possible actions are included in the next chapter.

5. Conclusions and Recommendations

A. Conclusions

148. This evaluation reaches several major conclusions. First, it is clear that the Norad Evaluation Department has made a good-faith effort to encourage use of its evaluations. It has consulted widely in deciding on its work program, and has emphasized the need for communication before, during, and at the end of the evaluation process. It has organized seminars and other activities to promote learning and has made its evaluation products widely available. There have been many positive results from these efforts.

149. However, there also are areas where the Department can improve. For example, despite work program consultations, it still is the case in the view of many informants that a number of evaluations are not targeted on high-priority issues or are not delivered in time to affect decisions. In the views of some informants there have been instances in which methodological choices have limited the utility of the evaluation itself, or in which a rigid adherence to highly-detailed terms of reference or methodological choices has prevented evaluators from following more promising analytical trails. As the Department notes, there is a tension between allowing evaluators great latitude and maintaining quality standards; too, the Department has an obligation under procurement law to ensure that deliverables meet agreed specifications. These tensions are not at all unique to Norad specifically, or to evaluation services more generally, but they do require the Department to balance among competing needs.

150. Even with the Department's robust quality control system there are questions about the quality of some reports, both in substance and presentation. Substantive issues include the quality of the data and analysis underlying some reports. Recommendations sometimes are not well-targeted or practical for implementation. There is a need for recommendations to be operationalized[38] and also for a multi-year annual framework for follow up on implementation of the recommendations. It shows value in an evaluation approach that is not "one-off", but rather a long term line of effort involving multiple reports.

---

38   See Hendricks and Handley, 1991, who argue evaluation recommendations should be clear, targeted, and monitorable.

151. In terms of presentation, some reports have been criticized for being too long and too academic, limiting their utility for busy decision-makers and program staff. The Department, however, faces a conundrum in dealing with this, in that it needs to balance its concern with readily-accessible presentation against the perception that it may be trying to impinge on the evaluator independence required under its instruction.

152. At a broader level, this review suggests that the theory of change with which we started (see figure 1) was too narrow and focused on the processes of Norad, and did not take adequate account of the broader arena within which the Evaluation Department works. As shown in figure 9, the independence of the Evaluation Department should allow it to select and carry out a relevant evaluation program, one that helps affect policies, programs, and projects not only directly, but indirectly by informing other sources of knowledge, as well as the political actors.

153. This broader theory of change is based on the premise that the purpose of Norad evaluations is to help improve the outcomes and impacts of Norway's development work. Thus, the Department's independence, derived from the Instruction, allows it to formulate and execute a relevant work program that can influence decisions on development policies, programs, and projects, both directly through its recommendations, and indirectly through its knowledge generation and conceptual influences. Figure 1 focuses on direct effects, primarily through recommendations. But it is clear that evaluation can have indirect effects, as well, through the other uses of evaluation defined in Chapter 1 and discussed in Chapters 3 and 4. For example, it can lead to a reconceptualization of a development issue as it is taken up by think tanks, the media, and others who are able to shape perceptions of issues, and as well as by political and other policy makers.

154. This potential influence on thought leaders and policy makers highlights the importance of communication in the evaluation process, and therefore the need for the Evaluation Department to ensure that its communication and dissemination activities are targeted for effectiveness. In this regard, the Department might consider efforts made by some other development evaluation departments to systematize and expand their use of knowledge sharing techniques. For example, the Independent Evaluation Department at the Asian Development Bank has used a range of up-to-date technologies, such as social media and podcasts, to disseminate its messages to a wide array of development stakeholders.[39]

155. The above makes "visuablity" of the evaluation reports and related-products of increased importance, as well as the use of multiple channels for communications. For example, more use might be made of social media to target particular groups.

---

39   See, for example, the podcast "Successful Engagement with Civil Society Organizations" at http://www.adb.org/site/evaluation/podcasts/successful-engagement-civil-society-organizations.

156. Finally, the Evaluation Department currently does limited tracking of whether, how, and to what extent its recommendations have been taken on board. Many evaluation departments, including those of the World Bank and the African Development Bank, as examples, routinely track the extent of implementation of their major recommendations over a period of several years. This longer view takes into account the fact that many important changes take time to implement, so that a one-year follow-up is insufficient.

Figure 9: Revised theory of change of the evaluation function in Norad
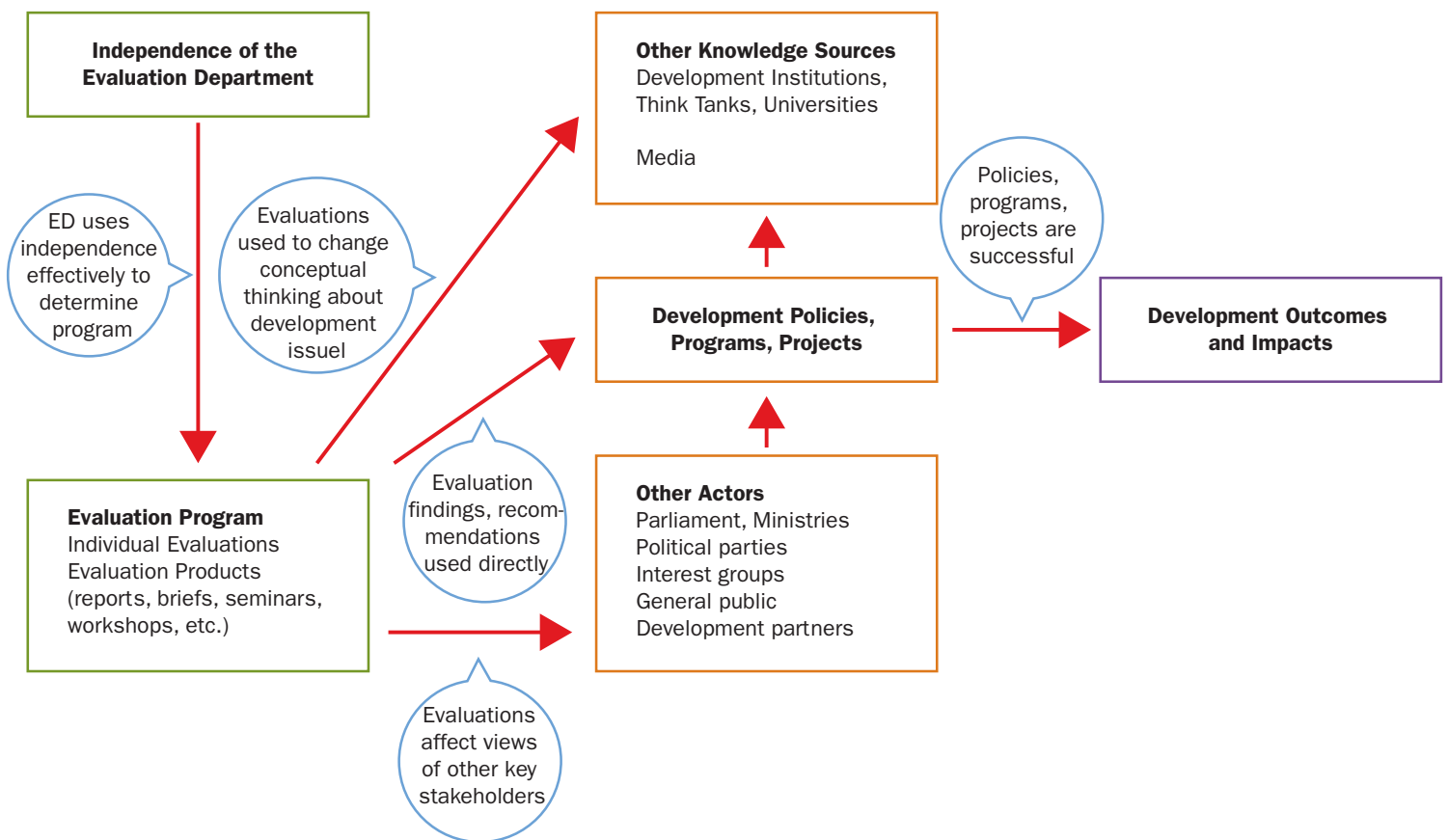
B. Recommendations

157. Based on the findings and conclusions from this study, we have the following recommendations. First, the Evaluation Department should strengthen the quality of its evaluation reporting by taking the following steps:

- Monitor the frequency of use of individual consultants to ensure that the range of those selected through competitive tendering is sufficiently broad to avoid the appearance—and perhaps reality—that it relies too heavily on a small group with pre-digested views; at the same time, continue efforts to broaden the range of competitors.

- Develop guidelines to ensure that report recommendations are clear and fact-based, that they can be implemented by the agencies to which they are addressed, and that it is feasible to measure whether and to what extent they have been carried out.

- Provide short, clear executive summaries for all reports.

- Develop short, non-technical briefs for all reports to make the results more widely accessible; the summaries prepared for Annual Reports can be used for this purpose.

- Provide targeted briefings to those units, particularly country offices, involved in the evaluations by arranging to brief them on the findings of relevant evaluations as part of regularly-scheduled unit staff meetings

- Emphasize dissemination—beyond the dissemination seminar—as a key part of the work of the Evaluation Department by adding time for dissemination activities as a standard evaluation activity and actively explore additional channels for communication, such as external networks, social media, and podcasts, to better target and disseminate its messages.

- Improve the "visuability" of reports, as discussed in Annex 3.

158. Second, the Department should improve its processes by:

- Formulating Terms of Reference that are clear, but also flexible enough to allow evaluators to pursue the most promising paths for answering the evaluation questions.

- Making efforts to keep stakeholders more engaged with the evaluation while the work is on-going, perhaps through periodic updates on progress.

159. Third, the Department should work with the leadership of Norad and MFA to address the structural ambiguities under which it operates currently. One option might be to provide a routine venue for the Director to deal directly with the Ministry's top management in order both to allow for an airing of organizational issues affecting the Department, but also to open a channel of communication that can increase the use of evaluation findings and recommendations at the highest levels. A more far-reaching possibility is to have the Department report directly to Parliament, as is the practice in the Netherlands. (A full review of such options was beyond the scope of this evaluation, however.)

160. Fourth, on the issue of follow-up of recommendations, we believe that this is the responsibility of MFA as the implementing organization. MFA could take several steps to ensure adherence to the requirements for responding to Evaluation Department reports, including:

- Making timely production of follow-up plans and actions an explicit element in managers' performance metrics.

- Conducting an annual review of compliance with requirements for responding to evaluation recommendations, with reference to specific reports, and disseminating the results to its management team and to Parliament.

- Promoting evaluation use by highlighting "success stories" of good use.

161. The Evaluation Department also should consider how it might promote accountability. One option, which would not detract from its independence, would be to produce and issue a report every two years on the status of recommendations with a focus on those which MFA has agreed to implement.

Annexes

ANNEX 1: Case study and other key informants

Erik Aakre Deputy Director, Section for Results Management Department for Quality Assurance, Norad

Knut Aarhus Senior Audit Adviser, Performance Audit Department II, Office of the Auditor General of Norway

Tone Bækkevold Allers Deputy Director, Section for Peace and Reconciliation, MFA

Ole Winckler Andersen Head, Evaluation Department, the Ministry of Foreign Affairs, Denmark

Marit Brandtzæg Director, Department for Quality Assurance, Norad

Eva Bratholm Director, Department for Communication, Norad

Beate Bull Senior Advisor, Evaluation Department, Norad

Vegard Bye Team Leader for the Evaluations of Norwegian Development Cooperation to Promote Human Rights, Scan Team

Thomas Dam Senior Adviser, Department for Quality Assurance, Norad

Arne Disch Team Leader for the current Evaluation of Norway's *Oil for Development Program,* "Facing the Resource Curse: Norway's Oil for Development Program" Report, Scan Team

Vibecke Dixon Senior Advisor, Evaluation Department, Norad

Lena Fargerlund Evaluation Department, SADEV, Sweden

Andreas Gaarder Deputy Director General, Department for Regional Affairs and Development, MFA

Marie Gaarder Director, Evaluation Department, Norad

Jon Hanssen-Bauer Former Special Envoy to Sri Lanka, MFA, Currently Special Advisor, the Middle East Section

Hilde Haraldstad Senior Advisor, Section for Eastern Asia, MFA

Ida Hellmark Advisor, Evaluation Department, Norad

Ted Kliest Policy and Operations Evaluation Department, Netherlands Ministry of Foreign Affairs

Villa Kulild Director General, Norad

Siv Lillestøl Senior Advisor, Evaluation Department, Norad

Geir Løkken Assistant Director General, Section for Human Rights and Democracy, MFA

Lene Lothe Senior Advisor, Global Health Section, Department for Global Health, Education and Research, Norad

Trine Rønning Mathisen Senior Advisor, Section for Management of Subsidiary Agencies & Development Funds, MFA

Tom McDonald Head of Secretariat, Independent Commission for Aid Impact, UK

Liv Marte Nordhaug Senior Advisor to the Director General, Norad

Petter Nore Director, Department for Economic Development, Energy, Gender and Governance, Norad. (Former head of section, Oil for Development)

Ingvar Theo Olsen Senior Advisor, Department for Global Health, Education and Research, Norad

Jan Håkon Olsson Deputy Head of Mission, Head of Development Cooperation, Royal Norwegian Embassy, Malawi

Beth Ann Plowman Team Leader, *Evaluability Study of Partnership Initiatives – Norwegian Support to Achieve Millennium Development Goals 4 & 5.* Report 9/2010 – Study

Inge Herman Rydland Senior Advisor, Section for International Development Policy, Department for Regional Affairs and Development, MFA

Unni Silkoset Senior Adviser, Global Health Section,Department for Global Health, Education and Research, Norad

Balbir Singh Senior Advisor, Evaluation Department, Norad

Siri Skåre Director, International Programs, Norwegian Centre for Human Rights, University of Oslo

Margot Igland Skarpeteig Team Coordinator, Rights and Social Equity Team, Department for Global Health, Education and Research, Norad

Lisbeth Skuland Deputy Director General, Section for Management of Subsidiary Agencies and Development Funds, MFA

Erik Solheim Former Minister of Environment and Development, former Special Peace Envoy to Sri Lanka, Currently Chair of OECD Development Assistance Committee

Gunnar Sørbø Team Leader, *Pawns of Peace: Norwegian Peace Efforts in Sri Lanka* Chr. *(1997-2009)* Michelsen Institute

Jon Teigland Former Senior Advisor, Evaluation Department, Norad

Hanne Tilrem Senior Advisor, Section for South-Asia and Afghanistan, MFA

Cliff Wang Senior Advisor, Department for Climate, Environment & Natural Resources, Norad

Nick York Head, Evaluation Department, DfID, UK

ADDITIONAL PERSONS CONTACTED FOR INTERVIEWS
Aslak Brun MFA, Evaluability Study
Cathrine Halsaa MFA, Human Rights
Anette Haug Leave of Absence from Evaluation Department
Sigvald Tomin Hauge MFA, General
Hege Hertzberg MFA, General
Aud Lise Norheim MFA, Human Rights
Inger Sanges MFA, Evaluability Study
Halvor Sætre MFA, Human Rights
Tone Tinnes MFA, Human Rights

DECLINED INTERVIEW (Generally due to position change)
Ervil Helga Fastrup MFA
Lasse Bjørn Johannessen MFA
Atle Leikvoll MFA
Ingrid Mollestad MFA
Thomas Stangeland MFA
Christian Syse MFA
Aud Mariat Wiig MFA

CASE STUDY INTERVIEWEES BY CASE AND GENERAL INTEREST

General

Knut Aarhus, Performance Audit Department

Marit Brandtzæg, Director, Quality Assurance, Norad

Eva Bratholm, Communications, Norad

Marie Gaarder, Evaluation Department

Ida Hellmark, Evaluation Department

Villa Kulild, Director General, Norad

Trine Rønning Mathisen, Section for Management of Subsidiary Agencies and Development Funds, MFA

Liv Marte Nordhaug, Senior Advisor to Director General, Norad

Lisbeth Skuland, Section for Management of Subsidiary Agencies and Development Funds, MFA

Evaluability Study of Partnership Initiatives

Erik Aakre, Quality Assurance, Norad

Thomas Dam, Quality Assurance, Norad

Lene Lothe, Health, Norad

Siv Lillestøl, Evaluation Department [Evaluation Manager]

Ingvar Theo Olsen, Global Health, Norad

Jan Håkon Olsson, Deputy Head of Mission, Head of Development Cooperation, Royal Norwegian Embassy, Malawi

Beth Ann Plowman, Team Leader [Consultant]

Unni Silkoset, Senior Adviser, Global Health Section, Department for Global Health, Education and Research, Norad

Hanne Tilrem, Senior Advisor, Section for South-Asia and Afghanistan, MFA

Cliff Wang, Climate, Environment & Natural Resources, Norad

Pawns of Peace – Evaluation of Norwegian Peace Efforts in Sri Lanka, 1997-2009

Beate Bull, Evaluation Department [Evaluation Manager]

Tone Bækkevold Allers, Section for Peace and Reconciliation, Department for UN, Peace and Humanitarian Affairs, MFA

Jon Hanssen-Bauer, former Special Envoy to Sri Lanka, MFA

Hilde Haraldstad, former Ambassador to Sri Lanka, currently Section for East Asia and Oceania, MFA

Inge Herman Rydland, Department for Regional Affairs and Development, MFA

Erik Solheim, former Special Envoy to Sri Lanka, MFA, former Minister of International Development

Gunnar Sørbø, Team Leader [Consultant]

Evaluation of the Norwegian Petroleum-Related Assistance

Arne Disch, Team Leader [Consultant-Scanteam]

Liv Marte Nordhaug, former Advisor for Norwegian Petroleum Directorate, East Timor case, Norad

Petter Nore, Director, Department for Economic Development, Gender, and

Governance, Norad, former head of section, Oil for Development
Balbir Singh, Evaluation Department. Evaluation Manager for Oil for Development study
Jon Teigland, Evaluation Department, Evaluation Manager

Evaluation of Norwegian Development Cooperation to Promote Human Rights
Vegard Bye, Team leader [Consultant, Scanteam]
Vibecke Dixon, Evaluation Department [Evaluation Manager]
Andreas Gaarder, Deputy Director General, Department for Regional Affairs and Development , MFA
Geir Løkken, Assistant Director General, Section for Human Rights and Democracy, MFA
Siri Skåre, Director, International Programs, Norwegian Centre for Human Rights, University of Oslo
Margot Igland Skarpeteig, Team Coordinator Rights and Social Equity Team, Department for Global Health, Education and Research, Norad

ANNEX 2: Semi-structured interview guide for case study interviews

Note: Actual interview questions will be customized for interviewees depending on their overall role.

1. What is the nature and extent of your role in using evaluations/promoting their use?
   a. In general
   b. Specific to the 4 evaluations

2. To you, what would constitute ideal use of an evaluation?
   a. In general, what would it look like?
   b. Specific to your position/unit

3. Do you see yourself as a stakeholder for the evaluation(s)? Primary or secondary?
   a. How were you involved?
   b. What were facilitators/barriers to your involvement?

4. Who do you see as the key stakeholders for the evaluation(s)?
   a. To your knowledge, were they appropriately involved in the evaluation(s)?
   b. Examples/evidence

5. How familiar are you with the key findings and recommendations of the evaluation(s)?

6. In your opinion, did the organization/your unit make good use of the report(s)? What about primary stakeholders?
   What factors affected their use?
   a. Barriers to use?
   b. Facilitators of use?
   c. Specific examples/evidence

7. How could use of evaluations be facilitated within and outside of the organization?
   a. Structural/organization issues
   b. Procedural issues
   c. Quality issues
   d. Communication issues (e.g. professional dissemination, media, social media)

8. Who else would you suggest I talk with about use of evaluation in the organization and outside of it?
    a. Who else would you suggest I talk with about one of more of the four evaluations?
    b. Where else might there be evidence of use of these evaluations?

ANNEX 3: Norad evaluations – a review of reports as products

Introduction

As agreed between Norad and the Rand team, this study of evaluation use was expanded modestly, at no additional cost, by adding a discussion of the conclusions and recommendations of Norad evaluations, based on the four reports already selected for case study analysis.

All four of these reports can, from one perspective or another, be considered atypical for "regular" development evaluations, and so the general applicability of the conclusions presented in this note may be open to some discussion. Also, the note looks at the reports as such, without reference to annexes or companion volumes, and it does not take into consideration the processes that may be followed in Norway once an evaluation report has been presented to Norad's Evaluation Division.

Also please note that this is not a review of the quality of the work done for the evaluations. Rather, for this review we have limited ourselves to asking the following questions:

- How well have the conclusions in the report been presented?
- Can the recommendations reasonably be traced back to these conclusions?
- How well have the recommendations been formulated for clear understanding and ease of action and follow-up?

In doing this limited analysis, we also added another perspective – what can be called the "visuability" of the reports – how easy is it to trace conclusions and recommendations back to the analyses and findings. In the following we start with the latter point before discussing conclusions and recommendations.

The Visuability of an Evaluation Report

In The Road to Results[40] Mora Imas and Rist make, inter alia, the following points on the writing of an evaluation report:

---

40   Linda G. Morra Imas and Ray C. Rist, *The Road to Results*. The World Bank 2009. Chapter 13: Presenting Results, pages 471-472.

- The purpose of a report is to communicate with readers.

- Organize the material in the body of the report into sections that address major themes or answer key evaluation questions.

- Place major points first in each section, with minor points later in the section. Open each paragraph by stating the point it addresses.

- Support conclusions and recommendations with evidence.

- The body of an evaluation report should contain the following components: introduction, description of the evaluation; findings; conclusions; and recommendations.

There is no single "gold standard" for how to write an evaluation report, but the above points do provide some guidance to a user-friendly report. Also drawing on the practical experiences of the Rand team, we would like to note that a busy reader of any report:

- Is unlikely to read it – or even skim it – from beginning to end.

- Rather, he/she may skim the Executive Summary (not discussed in this note) or go directly to the main report.

- In most cases, any look at the main report will start with the final chapter for conclusions and recommendations.

- Then if interested the reader may peruse quickly the earlier chapters of the report, either turning the pages for general interest (perhaps especially if looking at a print version), or seeking specific points of interest.

- Under either approach, a user-friendly report – in our language a report with high visuability - is one that permits the reader to find and absorb quickly the pertinent points.

The following paragraphs discuss the four reports regarding their visuability. The points emphasized by Morra Imes and Rist are central in this regard.

Pawns of Peace. Evaluation of Norwegian peace efforts in Sri Lanka, 1997-2009

This very interesting report is highly atypical for development evaluation – given the topic it is more akin to an academic political treatise. Given that the effort being evaluated has been completed there is also a less obvious scope for forward-looking recommendations. However:

- No paragraph numberings, no sidebars, some sections are pages long with no subtitles or paragraph headings. No indication as to which chapter you are reading (unless you are on the opening page of that chapter.) Some paragraphs are very long and should have been broken up. Very little use of cross-references (whether to the text or annexes).

- A number of the chapters have their own ending conclusions (chapters 2-3 and 7-11) marked by a sub-heading, but these are not shown in the table of contents.

- Final chapter: Heading is appropriate: Conclusions and Broader Lessons. Length is appropriate for a report this length (but could have benefited from a modest shortening). Only three subheadings for a chapter of 13 pages is sparse. No points underlined before "Broader lessons" (where these are shown numbered and bolded). The number of such broader lessons (seven) is appropriate, but there are no cross-references to where in the text these matters are discussed.

Evaluation of Norwegian Development Cooperation to Promote Human Rights

This report covers a wide-ranging and complex material. As a result the report is long (100 pages). Looking at it from a reader's perspective:

- No paragraph numberings (sections are numbered, but that is not the same thing although these numbers make it easier to find your way than if there are none), no sidebars, some words bolded in the text often to distinguish between the countries under discussion, some use of sub-headings, very few if any cross-references.

- Chapters 4, 5 and 6 have their own ending sections with findings and conclusions and these are appropriately shown in the table of contents

- Final chapter: Heading is appropriate: Findings, Conclusions and Recommendations. Length is fully appropriate (six pages), but also without cross-references. Some formulations are too cryptic especially in the absence of cross-references – see page 95 third paragraph final sentence: "In comparison to both these channels, the official Norwegian aid system has a lesson to learn." (In the absence of any explanation, who can understand what that lesson might be or where to find it?) It would also be preferable to do with all discussions and conclusions before presenting recommendations, to avoid ex-post explanations after a recommendation (as after Recommendation 1).

Evaluability Study of Partnerships Initiatives. Norwegian Support to Achieve MDGs 4&5

A study of evaluability cannot easily provide the same type of operational recommendations as the other studies in this sample. With that caveat in mind:

- No paragraph numberings. Sections are numbered, and while that is not the same thing most sections here are quite short. There are no sidebars (but a few bolded excerpts and some tables that help readers' understanding), a number of bolded section headings and a few bolded sub-headings, and some use of italics that in print are easy to overlook. Very few if any cross-reference, but less of a problem here than in the other reports (short and quite visible report).

- Structure: The structure of the report could seem unbalanced: A summary of more than five pages, a main report (before the final chapter) of 12 pages and a final chapter with "Main Recommendations" of seven pages

- The Heading of the final chapter is appropriate: Main Recommendations, but are there no conclusions worth highlighting – even for this special topic? The chapter is too long for the subject matter and the shortness of the report as a whole. There are five overall recommendations that better could be numbered (for ease of reference, and preferably in some order of priority) and headed in bold. Several of the recommendations are now followed by unmarked paragraphs that would seem in substance to represent additional recommendations – better to mark them clearly as such or alternatively move or drop them.

Evaluation of the Norwegian Petroleum-Related Assistance

This report probably comes the closest to a "standard" evaluation of an ongoing project or program. It is however a compilation of four case studies, mentions (page 1) that there has been at least one other major evaluation, but does not leave the impression (from a quick reading of the report) that this latest product integrates the findings of the earlier report.

Comments on the current report:
- No paragraph numberings, while sections are numbered, and mostly with significant use of sub-section numberings. There are no sidebars, while the frequent bolded section and sub-section headings tell about contents but do not give a flavor of the findings. Few if any cross-references.

- Structure: This report does not have a proper summing-up chapter. The final chapter deals with "Specific Issues" but without a condensation of key findings or of conclusions, while recommendations can only be found in the Executive Summary. In fact, the Table of Contents does not show the word recommendations at all. The apparent lack of attention to recommendations is in spite of several paragraphs with recommendations around in the text (such as pages 43 and 49), but you have to search the text to find them.

- For the recommendations you have to go to the Executive Summary. The report seems to have a lot of recommendations (18) although the Summary may perhaps not be separating clearly between recommendations, sub-recommendations, conclusions and comments. It would have been better for the current last chapter – Chapter 7 Specific Issues – to have been properly incorporated into the report, and including any important conclusions and recommendations with those coming from the rest of the report.

- Also, when reading the report it is clear that the authors have found a number of country-specific issues that perhaps should be better reflected in the Summary as well as in a proper final chapter and likely among the recommendations as well.[41]

This report more than the others could also raise some more general issues regarding the "intellectual framework" for some Norad evaluations, in two respects:

- The evaluations – both TORs and the reports – focus on "doing things right". However, from time to time it is also important to consider whether Norwegian aid (in this case) is "doing the right things" – the strategic and policy choices. As one example, the discussion of Bangladesh does not discuss the significant main issues that in this country have been affecting energy generally and the hydrocarbon sector in particular for a long time. (Based on some limited on-and-off connectivity with the country and sector we remember two questions in particular: (i) Energy pricing and the use of scarce resources to subsidize in this respect the better-offs in the country; and (ii) The reportedly mixed to low quality of government entities operating domestic gas fields and pipelines at least some time ago when the Norwegian cooperation with Bangladesh was ongoing.) It is not clear from the report whether Norway has ever considered how its assistance should best help addressing such priority sectoral issues.)

- **The relationship of Norway's efforts to those of other donors in a sector** – were there discussions of priorities, of organization, of linkages, of divisions of labor? If not – should there have been?

---

41   Page 10: Too little focus (Mozambique) on a number of issues. Page 14: Angola – no systematic institutional analysis nor of training needs – (page 15) impossible to assess impact of a program over 20 years. Page 15: Angola – Budgets seriously under-spent – indicating lack of effectiveness, while overspending on administration indicates lack of efficiency. Page 16: Bangladesh: Gap in BPI activities with focus on geophysical issues (as per page 17 the resident adviser was a geophysicist) while having left out a number of other important areas. Also staff in BPI and HCU had left at the end of the technical assistance. Page 19: Norad's lack of dialogue with Petrobangla. (The report indicates without apparently saying so outright that Norad was determined to follow the Norwegian model for sectoral organization.)

Conclusions: Visuability

On the basis of these four reports, the Rand team believes that the Norad evaluation reports can be improved from the perspective of readers' convenience:

- Keep the reports to a reasonable length, even if that means making use of annexes that could be attached to the main volumes.

- Also maintain some proportions in length between Summary, Main Text and the Final Chapter.

- The reports would benefit from systematic use of tools to distill for the readers the matters and findings being discussed. Two alternative tools would be (a) side-bars to distill the matters being discussed; and (b) using bolded first sentences to each paragraph to give the main findings/issues/conclusions of these paragraphs – as is done to some extent in the Evaluability report.

- More use of explicit cross-references would also make it easier for readers to dig down on the issues of importance to them – and it might be good practice always to do so in the final chapter with references to the discussions in the previous chapters.

The Rand team also considers it to be good, reader-friendly practice always to have a final chapter that should not provide new material, but that should summarize the most important findings and conclusions and on this basis provide the recommendations from the evaluation. This does not seem to be standard practice for Norad's reports. Such a final chapter may of course overlap significantly with a good executive summary – this is in our view an overlap worth having for readers' convenience, but if the summary is to serve also as the summing-up chapter, this should be stated clearly for readers' orientation.

The Conclusions and Recommendations of the Reports

The book *The Roads to Results* also (page 473) comments that the last part of a report should be its conclusions and recommendations, which readers often read first. It stresses in this regard the following:

- Evaluators have often difficulty distinguishing findings from conclusions.

- Findings describe what was found in the evaluation, and should be supported by evidence.

- Conclusions are based on professional assessment of the findings. They should be made about each evaluation sub-objective as well as the overall objective of the program or policy.

- No new information should be presented in the conclusions section

- Recommendations advocate action, indicating what the report wants the client or other stakeholders to do.

- Reports should not include "laundry lists" of recommendations, which should be limited in number.

- The recommendations should not be overly prescriptive, but at the same time cannot be so general that they have no teeth. Within these borderlines, recommendations should be clear and specific enough so that all understand what needs to be done to satisfy them, which organization or unit needs to take action, and when it should be done.

Drawing on these points, the following paragraphs discuss the four reports regarding their findings and recommendations.

Pawns of Peace. Evaluation of Norwegian peace efforts in Sri Lanka, 1997-2009

- The conclusions seem to flow well from the earlier chapters.

- All of the broader lessons are of a general (academic) nature, of interest to anyone anywhere who might be contemplating some sort of peace-making effort in a country under conflict.

- A reader looking for the more specific recommendations for a country like Norway (what should the country do differently if it should at some future stage contemplate a similar effort somewhere) will find these more under the earlier pages (133-135 on Assessment of Norway's role).

- Perhaps for a report of this "academic" kind there could have been two sets of recommendations – both brought out clearly – first on the broader lessons and second on what Norway should consider if in such a situation again.

- A question could also be whether there would be any more concrete actions that should be considered regardless of any future events: would there be a Norwegian manual requiring modifications, anything in the budgeting processes that should be revised, any training program for officials at different levels?

Evaluation of Norwegian Development Cooperation to Promote Human Rights

- The conclusions seem to flow well from the previous chapters, although this is difficult to ascertain with any precision in the absence of cross-references and/or other "Visuability" tools. Certainly, a reader interested in reading more about any of the conclusions will have to work!

- The number of recommendations (five) is appropriate, as is the use of bullets under some of the recommendations, if reading these as examples or illustrations of the broader point in the respective recommendations, and not as separate recommendations.

- The recommendations are also quite specific and presumably therefore actionable, although the report does not address the recommendations to any particular entity within the Norwegian government system. Rather, they are written – as seems normal for Norad evaluations - in the passive voice.[42]

Evaluability Study of Partnerships Initiatives: Norwegian Support to Achieve MDGs 4&5

- As mentioned above, this evaluation report has a final chapter with Main Recommendations, but no conclusions. For that you need to go to the Summary, which is inconvenient.

- Also, the Summary deals much more with approach and methodology than about findings and conclusions, except for the specific discussions of the four specific country programs.[43]

- The final chapter is too long for the shortness of the report as a whole.

- There are five overall recommendations that better could be numbered (for ease of reference, and preferably in some order of priority) and headed in bold.

- Several of the recommendations are now followed by unmarked paragraphs that would seem in substance to represent additional recommendations – better to mark them clearly as such or alternatively move or drop them.

---

42   We recognize that for outside consultants it may at times be difficult to direct recommendations with any precision.

43   By the way, for an outside reader it would have been very useful with some more discussion of the four programs – their structure and modalities, any linkages to work/activities by other donors, and the linkages to the countries' own health issues, strategies, administration and financing.

Evaluation of the Norwegian Petroleum-Related Assistance

- As mentioned above this report does not have a proper final chapter. What is now the final chapter deals with "Specific Issues" but without a condensation of key findings or of conclusions, while recommendations can only be found in the Executive Summary.

- The discussion of "Specific Issues" would seem to belong in an earlier chapter, but it does raise a number of points that could well belong as conclusions in a final chapter.

- The presentation of conclusions in the Summary could probably have been extended somewhat.

- For the recommendations you have to go to the Executive Summary. The report seems to have a lot of recommendations (18) although the Summary may perhaps not be separating clearly between recommendations, sub-recommendations, conclusions and comments.

- Also, when reading the report it is clear that the authors have found a number of country-specific issues that perhaps should be better reflected in the Summary as well as in a proper final chapter and likely among the recommendations as well.

- The recommendations are in places rather exhortatory, not addressed to anyone in particular, and their implementation would not be easy to track. With modest effort they could probably have been converted into about half a dozen real recommendations, with illustrative bullets in some cases, since most of these recommendations are of the "good housekeeping" variety and could therefore be merged.

- But the evaluation would also have benefited from some clearly actionable items – items which could be followed up (without significant massaging by Norad's Evaluation Division).

Summing up: Conclusions and Suggestions

The sections dealing with conclusions and recommendations have been treated differently in the four reports. Also, the recommendations presented by the consultants have tended to be of a quite general nature, held in passive voice/tonality, not directed at anyone in particular, and for these reasons not necessarily easy to convert into action-and future oriented tasks.[44] We suggest that Norad could consider the following:

---

44   We recognize that Norad has a regular post-evaluation process to deal inter alia with such matters. We also recognize – as is also mentioned in several of the four reports – that the absence of clear, monitorable objectives, preferably with baselines, may make more difficult the production of such desired, firm recommendations.

- Always insist on a final chapter with Conclusions and Recommendations.

- Recommendations should be clearly linked to the main conclusions as presented in the final chapter.

- There should normally be not more than six recommendations, grouping as needed smaller or related aspects as bullets under the relevant recommendations as examples for consideration.

- Encourage consultants to show firmness and directness in formulating their key recommendations, if possible with indications of priority and with possible timetables.

- Be explicit in linking recommendations to the findings of the evaluations.

ANNEX 4: Norad Coding Scheme (Example)

Code 1: Interviewee

(Response Options)

    a. MFA
    b. Norad
    c. Embassy
    d. Evaluation Department

Code 2: Purpose of Interview

(Response Options)

    a. General Background
    b. Specific Report

Code 3: Report

(Response Options)

    a. Sri Lanka
    b. Evaluability Assessment
    c. Petroleum-Related Assistance
    d. Human Rights

Code 4: Role of Interviewee

(Response Options)

    a. Current (Indicate what role, e.g. senior health advisor)
    b. During study of earlier relevance (Indicate what role, e.g. highly involved in structuring the evaluation)

Code 5: Structural Independence of Evaluation Department

(Response Options)

    a. Positives
       i. Work hard to maintain independence
    b. Negatives
       i. Appearance is that <u>Eval Department interferes</u> with the findings and conclusions of the consultants. "They should accept the report as written

and then simply get the Management response to it."
  ii. Eval Depart seen as NORAD
  iii. Must have unit that does follow-up
  iv. Act of getting comments on drafts makes them appear not independent
  v. Lots of process/bureaucracy

Code 6: Quality of Evaluation Department Staff
(Response Options)

  a. Positive
    i. Generally good
  b. Negative
    i. Need to focus more on development issues
    ii. Less use of DAC criteria as organizers
    iii. No capacity to do the work themselves as TLS even with help from consultants
    iv. Need to think outside the box
    v. Have rote approach
    iv. Need new methods

Code 7: Procedural Issues
(Response Options)

  a. Contractual Process
    i. Positives
      1. Lot of input into TOR: highly involved in structuring the evaluation
      2. Study requested by MFA
      3. Study Interest at highest level-PM
    ii. Negatives
      1. Cannot get who want to do the work-process decrees
      2. Don't even know who bids
      3. Eval Depart too technical in approach
  b. Program Process
    i. Positives
    ii. Negatives
      1. Diffuse and complex responsibility for program
      2. Steering Committee with unclear mandate
      3. Steering Committee is highly political
      4. Steering Committee being restructured
  c. Other
    i Positives
      1. Giving extensive comments
  d Type of Uses For Report
    i Instrumental
    ii. Decision-making
    iii. Conceptual
    iv. Symbolic
    v. Communication

     vi. Valuing

     vii. valuation learning

Code 8: Quality of Team Contracted to Do the Work (Case specific)

(Response Options)

  a. Positives
     i. Efficient [in setting up seminar and follow-up note]
     ii. Took note of extensive comments on the draft-improved
     ii. Lead author seen as objective
  b. Negatives
     i. Too theoretical
     ii. Need direct practical experience
     iii. Too much academics
     iv. Potential conflict of interest (e.g. writing a book on the same process)
     v. Team does not take guidance (e.g. focuses on history of Sri Lanka's development aid rather than keep to political process)
     vi Lack of MFA buy-in of results
     vii. Conclusions before study data collection
     viii. Not viewed as neutral or objective

Code 9: Report Quality Issues

(Response Options)

  a. Positives
     i. Easy to read
     ii. Lots of learning
     iii. Showed political dilemmas
     iv. New: evaluation of a political process
     v. Typically OK reports
  b Negatives
     i Recommendations not always relevant
     ii. Team could not get Visas to go in-country-reduced credibility
     iii. Sri Lankan members had to withdraw and stay anonymous
     iv. Should be able to keep whole report confidential
     v. Many evaluations going on at the same time
     vi. Often have greater funding
     vii. Others more comprehensive and sophisticated
     viii. Lots of players
     ix. Latest approaches focus on regional institutions
     x. Narrow focus means less use
     xi. Same message-lack of clear objectives, data

Code 10: Use

(Response Options)

  a. Positives
     i. Lots of use of report

    ii. TL (Eval Depart.) gave a follow-up internal briefing

    iii. Olso Forum-external session

    iv. Working on MFA guidelines

    v. Contributed to others' guidelines (UN mediation guidelines)

    vi. Innovative panel format for seminar

    vii. Use to review strategic direction of units

    viii. Whole process of engagement—discussion—was good

    ix. Study confirmed need for solid data

    x. Relevant

    xi. No conflict about findings

    xii. Was a Pay-for-Performance pilot; supported already developed M&E plans for the pilot

    xiii. His unit used to advocate own role; demonstrate

b. Negatives

    i. Confirmed their thinking; nothing new

    ii. No surprises

    iii. Not specific enough guidance to change anything

    iv. Gap between publication and seminar. Yields suspicion and negative press

    v. Hard to see more use

    vi. Narrow focus makes more difficult to read and less use

c. Suggestions (Here are examples from Evaluability Assessment)

    i. Weakness that org cannot attract more attention to evals

    ii. Need evals with wider scope

    iii. Think outside the box

    iv. Need to think upfront about marketing

    v. Need targeted briefings

    vi. In-country briefings

    vii. On-line discussions

    viii. Reports must be issue oriented

    ix. Focus on key issues

    x. Get away from DAC criteria as report organizers

    xi. Do less reports but go deeper

    xii. Method can create interest e.g. recreate baseline

Code 11: Best Quotes

Code 12: Comments

Code 13: Other Codes Not Captured From Above

ANNEX 5: Terms of Reference (TOR)

Study of evaluation use in the Norwegian development cooperation system.

"All evaluations have a cost but not necessarily a value. Their value does not depend on their cost but on their use" (Osvaldo Feinstein, 2002). [45]

1.0 Introduction

Billions of Norwegian Kroner are spent on development programmes. In 2011 alone spending amounted to 27, 7 billion Norwegian Kroner. To improve the efficiency of development spending, we need a better understanding of what works and how. A primary tool for acquiring understanding about effectiveness, relevance and impact of development cooperation programmes is evaluations, studies and other related research. While it is internationally recognized that there is a need to ensure that more development programmes are being evaluated, the value of these evaluations ultimately depends on their use. The challenge is to use these studies for better learning and decision-making.

This study will examine the use of evaluations, mechanisms for knowledge sharing and learning in the Norwegian development cooperation system.

2.0 Background for the study

Use of evaluations

The use of evaluations can take different forms, but use is assumed to always form part of the objectives of an evaluation. In the OECD/DAC principles for Effective Aid, 1992, it is stated that "to have an impact on decision-making evaluation findings must be perceived as relevant and useful (…)."[46] The OECD/ DAC principles moreover state that the main purposes of evaluation are:

- to improve future aid policy, programmes and projects through feedback of lessons learned, and
- to provide a basis for accountability, including the provision of information to be public.

Most evaluation functions in bilateral and multilateral development cooperation agencies, including in Norway, base their evaluation activities on these principles and include institutional arrangements to secure follow-up and feed-back from the evaluations to programme and policy departments. However,

---

45   Feinstein, Osvaldo (2002), p. 433: "Use of Evaluations and the Evaluation of their Use" in Evaluation, Vol 8 (4), pp- 433-439
46   P. 8,"Evaluating Development Cooperation – summary of key norms and standards", OECD/DAC network on development evaluation, OECD.

there is little knowledge and research about the degree to which these evaluations are actually used, when and how they are used, and which factors influence their use.

The Norwegian development cooperation system
The Norwegian development cooperation system is composed of the following organizational entities. Each has the following different responsibilities:

1. The Norwegian Ministry of Foreign Affairs is responsible for policy-making, -decisions and the management of 62% of the aid budget (including funding to multilateral organizations),
2. Embassies and delegations are responsible for managing 18% of the aid budget at country level, and
3. The Norwegian Agency for Development Cooperation (Norad) is responsible for management of 13% of the aid budget (most of which is allocated to the Norwegian NGOs), as well as for providing sector- and thematic advice and quality assurance services. In addition, Norad is responsible for evaluating all activities funded over the Norwegian aid budget.
4. Norfund, the Peace Corps and the Office of the National Auditor manage 7 % of the aid budget.

In this study we will cover the following organizational entities: Norad, the Norwegian Ministry of Foreign Affairs and embassies. The term 'Norwegian development cooperation system' will hereafter refer to these entities only

Norwegian development aid in 2011
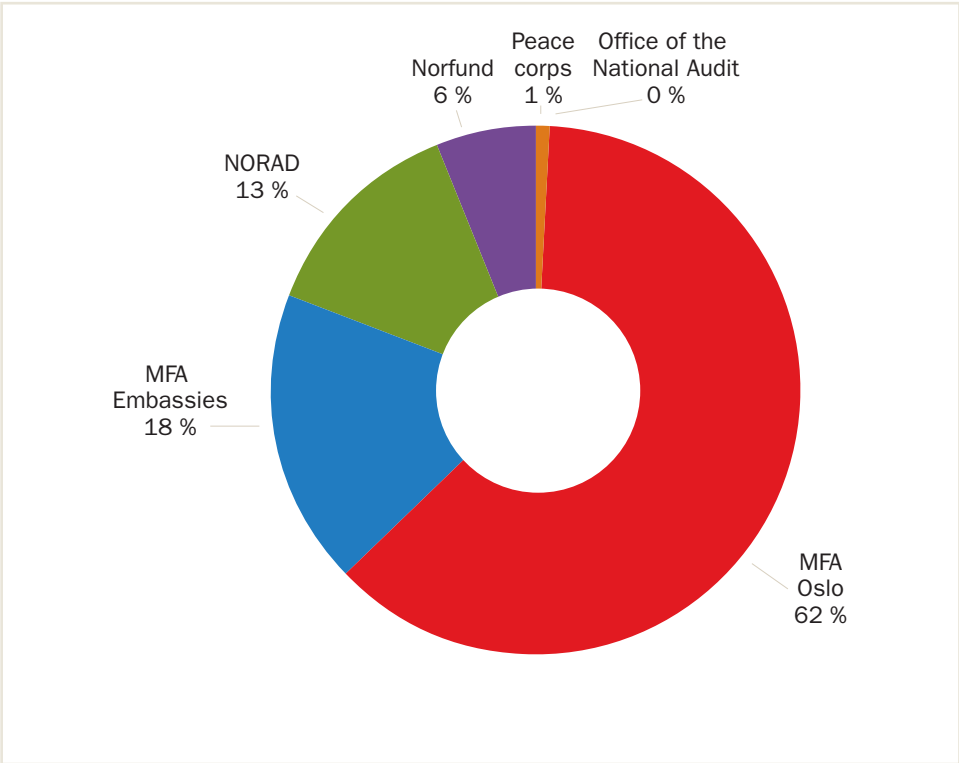
The Evaluation Department in the Norwegian Development Cooperation System

The Evaluation Department located in Norad is equipped with a separate mandate[47] to implement evaluations of all activities financed over the Norwegian aid budget. It is responsible for independently initiating evaluations, as well as communicating the results to the decision-makers and the general public. Its mandate also tasks it to contribute to good learning processes. With a budget that constitutes less than 0,1% of the aid budget, the evaluations carried out by the Evaluation Department are mostly thematic and/or institutional evaluations and rely on the quality of the evaluations built into individual projects and programmes[48] to be able to draw conclusions about results and outcomes. These 'thematic' or 'overview' evaluations will by their very nature often have more than one responsible entity or counterpart in the Norwegian development cooperation system. For example, when evaluating Norwegian support to the education sector through multilateral channels, this typically involves the departments/sections in charge of the multilateral aid within the MFA, the UN delegation in NY, as well as the various embassies in the selected case-countries. Similarly, a recent evaluation of Norwegian support to promote the rights of persons with disabilities would typically require that the implementation of actions and concrete measures percolates throughout the whole Norwegian development cooperation system.

Since 2006, the Evaluation Department in Norad has conducted around 60 evaluations, either alone or jointly with other donors. In addition, the Evaluation Department has contributed to evaluations in UNDP and the World Bank through partnership agreements with the two organizations. The evaluation mandate requires that the Evaluation Department reports, disseminates and provides follow-up recommendations for the evaluations that it is responsible for. The mandate also requires that the Evaluation Department organize hearings or consultations with the responsible section(s) or unit(s) in the MFA, embassies and /or Norad during the evaluation process, allowing them to comment on the terms of reference and the draft report.

System for follow-up of evaluations

The institutionalized/formal system for follow-up of evaluations in the Norwegian development cooperation includes a memo addressed to the Director-General of the Norwegian Ministry of Foreign Affairs, synthesizing the evaluation findings and conclusions, stakeholder views, and recommending actions for follow-up of individual evaluations. The relevant section(s) or unit(s) in the MFA or Norad is tasked with preparing a plan within six weeks of receiving the recommendations from the Evaluation Department to account for how they intend to follow them up, and by which timelines. Within a year, the responsible section /unit must report on progress made on these actions.

---

47   http://www.norad.no/no/evaluering/h%C3%A5ndb%C3%B8ker-og-referansedokumenter/_
     attachment/161205?=true&_ts=126e597dde7
48   Ballpark estimates suggest that between 1% and 10% of the budget for an individual programme or project
     should be set aside for evaluation purposes (with smaller projects requiring a relatively higher amount than
     larger).

According to the Evaluation Department's own records for the period since 2006, not all the evaluations produced by the Evaluation Department have the required MFA follow-up plans and/or progress reports. There might be several reasons for this, amongst other that there are multiple responsible units for the evaluations as mentioned above, frequent turnover of staff, or the perceived relevance or/and the quality of the evaluations are not found to be satisfactory by the responsible section, unit or entity. However, due to the formalistic character of these plans and reports, several questions are left un-answered such as how and why the evaluations in question are actually used or not, and when in the evaluation process use took place, etc.

Another question is whether there are systems in place to ensure broader or cumulative learning from the evaluations (as opposed to the individual evaluation) presented to the Norwegian development cooperation system. Some of the evaluations have recurrent findings and recommendations. Is there a system for uptake of learning across evaluations? For example, the annual reports of the Evaluation Department summarize recurrent findings, conclusions and lessons across the evaluations on a yearly basis. Norad's Results Report also summarizes lessons across individual evaluations and reviews. What happens with these reports and other learning products?

Types of evaluation use

In the literature on evaluation use, use is characterized as something that can take place in many different forms before and during the implementation of an evaluation, or long after, and not necessarily only directly after the findings are presented and lessons are communicated. In some cases decision makers use the findings to change or modify the programme. This is called instrumental use.[49] But in many cases use is less easy to capture or to document, such as when it occurs as cognitive, behavioral, programme, or organizational changes resulting not from the evaluation findings but from engaging in the evaluation activities or -process, and learning to think evaluatively. This is often referred to as process use.[50] [51] Another type of use that is not easily traced occurs when evaluation findings help the program staff or key stakeholders to understand the program in a new way, or just acquire broader learning. This is referred to as conceptual use.[52] Sometimes, knowledge about a planned evaluation is sufficient to start thinking differently about a programme, and new perspectives are gained. These types of use (conceptual and process use) are not always easily captured in formalized management response systems; There are rarely boxes to tick off regarding this type of use; there might be little awareness about this in the organisation, and/or there are attribution time lags and use does not occur until a while after the evaluation has been completed and feedback

---

49 Forss, Rebien and Carlsson (2002), also refer to five different types of process use: learning to learn; developing networks; creating shared understanding; strengthening the project; and boosting morale in "Process Use of Evaluations – Types of Use that Precede Lessons Learned and Feedback", in *Evaluation*: 8: 1: 29-45.
50
51 Patton, 2008, p. 109.
52 Fleisher & Christie, 2009.

provided. Conceptual and process use might also to a larger extent follow individuals rather than systems.

Some authors claim that the discussion about evaluation use favors intended and planned use, instead of unplanned use. This, they claim, restricts our understanding of use.[53] Another type of function of an evaluation (system) is not necessarily related to utilization, but instead to provide legitimacy for an organization or programme.[54]

The experience in the Evaluation Department is that in some cases use occur early in the process, whilst in other cases, the end product is taken as a basis for internal learning and decision making. Information about when and how in the process use occur, will help the Evaluation Department prioritize its resources to maximize the impact of its evaluations.

Factors assumed to impact evaluation use
There are different factors assumed to influence use of evaluations. We will here emphasize several such factors: the institutional environment (incentives and capacity for use), the relevance of the evaluations (timing, involvement of stakeholders, credibility) and the quality of their dissemination (evaluation product, communication channels and mechanisms). [55] Low levels of these factors can act as barriers to use. In terms of the institutional environment, this is expected to differ some from Norad to the MFA and the embassies. It is expected that MFA and embassies have much less resources to spend on evaluation use and follow up, than Norad. Yet, these are the two entities that manage most of the resources and take most decisions (80%). Other factors determining use are assumed to be evaluation type, level of control over the evaluation object and their follow up.

3.0 Purpose and objectives
The purpose of this study is to learn about the impact of its evaluations in the Norwegian development cooperation system, how the current follow-up system works and what are the incentives, capacities and mechanisms in place to ensure use of the evaluations in decision making and for broader learning.

Objectives
– To identify to what extent, how and when in the evaluation process, evaluations produced by the Evaluation Department are used
– To assess which factors have been most decisive in promoting/hampering use of evaluations in the Norwegian development cooperation system
– To assess whether the existing system for follow-up of evaluations meets the current needs of the Norwegian development cooperation system
– To assess whether the existing system for dissemination (evaluation product, communication channels and mechanisms) of evaluations meets the current needs of the Norwegian development cooperation system
– To identify possible mechanisms for cumulative learning
– To assess the main factors influencing decision-making

---

53   Boswell, 2009, Dahler-Larsen, 2007 in Hanberger, A. (2011) "The real functions of evaluations and response systems" in Evaluation,17 (4), 327-349.
54   Hanberger, A. (2011), p. 328.
55   Feinstein, O. "Use of Evaluations and the Evaluation of their Use" in *Evaluation* 2002;8; 433

– Based on the previous six objectives, provide recommendations on how the Norwegian aid administration can improve systems for learning and evaluation use, including explicit recommendations for the Evaluation Department and the evaluation function.

4.0 Scope

The study shall cover evaluations supplied by the Evaluation Department in the time period from 2006 to the present and the following organizational entities: Norad, the Norwegian Ministry of Foreign Affairs and embassies.

The study shall put considerable emphasis on analysing the factors at play that may be influencing use of evaluations.

Moreover, the study shall distinguish between different types of use; for example (and not exclusively): instrumental or direct use; indirect/conceptual or cognitive use; actual use (the way in which evaluations are actually used) versus apparent use (what seems to be the use). There is also a differentiation between potential use and actual use that could be taken into consideration if relevant. The former focuses on possible uses of evaluations that are more intensive and beneficial. Other forms of use that can be considered are planned use versus un-intended use and the provision of legitimacy. The study will distinguish between changes in knowledge, attitudes and behavior, when analyzing policy or program influence.

The consultant/researcher must bear in mind time lags in the attribution such as a possible 'gestation' period for the occurrence of use. It might seem that there is no evidence of use and therefore no use, but this might be a result of process leading from the production of the evaluation to its use takes time.

Reviews and evaluations conducted by embassies, partners, sections and departments in the Norwegian Ministry of Foreign Affairs and Norad will not be included in this study, unless they occur as a basis for policy decisions, when tracing backwards (according to objective six above). For evaluations of development partners, such as the multilaterals or the Norwegian NGOs, the study will look specifically at how the responsible entities in MFA and Norad brought the recommendations into discussions with the organizations, or into contracts or Board documents.

A part of this study is to trace which type of knowledge or evidence is in demand when policy or strategy decisions are made in the Norwegian development cooperation system. This will be done by tracing backwards from select decisions. In doing this we seek to establish not only to what extent the evaluations are used, but also which source of knowledge, learning, experience and information is in demand, as illustrated by figure 2.

Figure 2: Demand and supply of evaluations in the Norwegian development cooperation
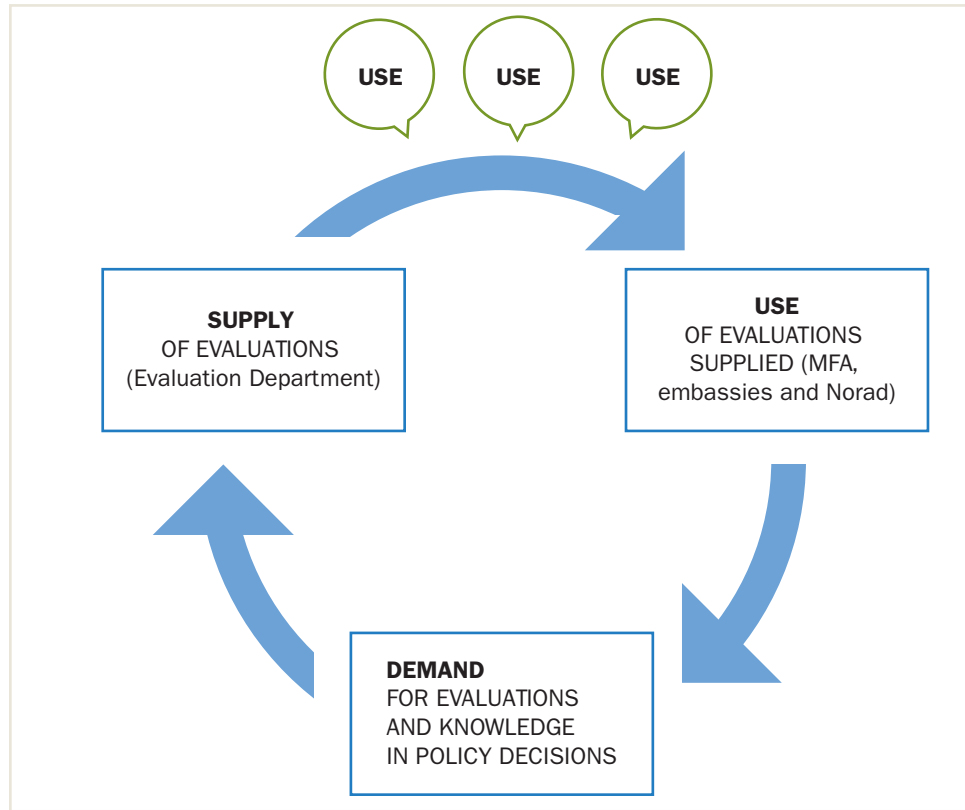
5.0 Methods/Approach

The study will be implemented in the following way to answer the evaluation objectives.

1. A desk review of evaluation use, and which institutional arrangements in development organizations facilitates use of evaluations. This will serve as a reference point for the analysis of the Norwegian development cooperation system. The desk review should include (mainly based on secondary sources such as existing evaluations, reviews and studies, and telephone interviews) a comparison of experiences, different evaluation functions, systems for follow-up and mechanisms in place for ensuring learning in relevant organizations (ex: Sida/SADEV, Dfid/ICAI, Netherlands, Danida). The review can also be complemented by face to face interviews, visits/ surveys, if resources team member locations permit.

2. Mapping: As part of the study a mapping of all evaluations and their follow up in the Norwegian aid administration since 2006 shall be done based on information available through the formal reporting system, interviews, and document analysis. The purpose of this part is to provide information on the degree of follow up for all the evaluations according to the formal response system. This is to serve as a benchmark for the next part; the in-depth case studies.

3. In-depth case studies of evaluation use and follow-up: Selection of a set of evaluations that shall be carefully studied in terms of how they were conducted, commissioned, degree of involvement, relevance, timeliness, quality, dissemination, and follow up, both informal and formal to find out about which factors influenced use in each case, and what type of use emanated from the evaluations. The evaluation cases will be selected based on the assumption that they will shed light on the questions above. A method for selection of cases that best answer the purpose of the study shall be provided by the consultant in the tender. Key informant interviews shall be conducted, as well as document analysis. The purpose is to provide more qualitative information about reasons for use, non-use or types of use, applying the study's analytical framework. Semi-structured or open ended interviews shall be used. Document analyses shall be undertaken of both the evaluations, annual reports from the Evaluation Department, their dissemination, their follow up and related /other relevant documents, to trace use.

4. In-depth case studies of policy decisions: a selection of a set of policy decisions will be made. The study will investigate how these policy decisions came about, which knowledge sources were used, what role did experience, evidence or other types of knowledge play. Were evaluations used? Criteria for selection of policy decisions: they must be preceded by deliberation processes and to be further devised in the tender. The analysis will start with the policy decisions and then seek to trace back what went into the policy decision in terms of knowledge, evidence, analysis and deliberations.

In addition, a survey among individuals in the Norwegian development cooperation system will be conducted by the Evaluation Department to contribute to identifying capacities, incentives and bottlenecks preventing or facilitating use of evaluations and other types of knowledge/evidence/experience in decision making in the Norwegian development cooperation system. The evaluation team/tenderer shall consider this survey when synthesizing the findings, conclusions and recommendations of the study.

Methodological criteria:
a. The proposals shall follow relevant OECD/DAC evaluation guidelines, including the Evaluation Quality Standards and the Criteria. In addition, for any impact evaluation, the 3ie principles of impact evaluation will apply (http://www.3ieimpact.org/doc/principles%20for%20impact%20evaluation.pdf).

b. All proposals are expected to explicitly address the issue of attribution, and how it will be assessed.

c. The less untested significant assumptions the evaluation proposal consists of, the higher it will be scored on quality. More generally, the quality of the proposal – and ultimately the study - will be assessed on the extent to which it identifies credible program theories and underlying assumptions. Furthermore, the extent to which the underlying assumptions are grounded on real

evidence and directly tested by the evaluation, will be important.

d. Triangulation of methodologies is key to increase reliability of the evaluation, and proposals will be assessed on the extent to which this is done.

6.0 Organisation and Requirements

The evaluation will be managed by the Evaluation Department, Norad (EVAL). An independent team of researchers or consultants will be assigned the evaluation according to prevailing regulations on public procurement in Norway. The team leader shall report to EVAL on the team's progress, including any problems that may jeopardize the assignment.

The MFA as a main stakeholder in the evaluation will be asked to comment on the following evaluation products: inception report, draft report and final report. However, all decisions concerning changes to the ToR, the inception report, draft report and final report are subject to approval by EVAL.

The team should consult widely with stakeholders pertinent to the assignment.

The evaluation team shall take note of comments received from stakeholders. Where there are significantly diverging views between the evaluation team and stakeholders, this should be reflected in the report.

The consultants are responsible for obtaining the necessary permits for collecting data. Data collected shall be made available upon request.

7.0 Budget and Deliverables

The project is budgeted with a maximum input of 23 person weeks. The deliverables in the consultancy consist of following outputs:

• **Inception report**
• **Draft Final Report** for feedback from the stakeholders and EVAL.
• **Final Evaluation Report** maximum 60 pages - prepared in accordance with EVAL's guidelines given in Annex 3 Guidelines for Report of this document.Seminar for dissemination of the final report in Oslo will be organised by EVAL.

Bibliography

Atkinson, Rowland, and Flint, John. 2004. "Snowball Sampling." In Michael Lewis-Beck, et al. (eds.), *The SAGE Encyclopedia of Social Science Research Methods.* Thousand Oaks, CA: Sage Publications.

Fleischer, Dreolin N. and Christie, Christina A. 2009. Evaluation Use: Results from a Survey of American Evaluation Association Members. *American Journal of Evaluation* 30, pp. 158-75.

Hendricks, Michael, and Handley, Elizabeth A. 1990. "Improving Recommendations from Evaluation Studies." *Evaluation and Program Planning* 13, pp. 109-17.

Henry, Gary T. and Mark, Melvin. 2003. "Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions." *American Journal of Evaluation* 24, pp. 293-314.

Jones, Harry and Mendizabal, Enrique. 2010. *Strengthening Learning from Research and Evaluation: Going with the Grain*. London: Overseas Development Institute, 2010.

Morra, Linda G. and Friedlander, Amy C. *Case Study Evaluations*. Washington: World Bank, 1999.

Norad. 2012. "A Summary of Results from the Norad Survey: How Do You Engage with Evaluation Knowledge in Your Day to Day Work." (draft).

Norad. 2012. *Evaluation of Norwegian Development Cooperation with Afghanistan, 2001-2011*. Report 3.12.

Norad. 2010. *Evaluation of the Norwegian Center for Democracy Support, 2002-2009*. Report 1.10

Norad. 2010. *Real-Time Evaluation of Norway's International Climate and Forest Initiative*. Report 12-18.10

Norad. 2012. *Real-Time Evaluation of Norway's International Climate and Forest Initiative Lessons: Learned from Support to Civil Society Organizations.* Report 5.12

Norad. 2011. *Evaluation of Results of Development Cooperation through Norwegian NGOs in East Africa*. Report 1.11

Norad. 2010. *Results Report. Capacity Development: Building Societies Capable of Sustaining Themselves.*

Norwegian Ministry of Finance, 2010, "Financial Management Regulations."

Norwegian Ministry of Foreign Affairs. 2006. "Instruction for Evaluation Activities in Norwegian Aid Administration."

Organisation for Economic Cooperation and Development. 2010. *Evaluation in Development Agencies. Paris: OECD Publishing.*

Patton, Michael Q. 2008. *Utilization-Focused Evaluation* 4th ed. Thousand Oaks, CA: Sage.

Preskill, Hallie and Caracelli, Valerie. 1997. "Current and Developing Conceptions of Use: Evaluation Use TIG Survey Results," *American Journal of Evaluation,* 18, pp. 209-225.

Rist, Ray C. and Kusek, Jody Zall. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System.* Washington, DC: World Bank.

Robert Wood Johnson Foundation. 2010. *An Ecological Understanding of Evaluation Use: A Case Study of the Active for Life Evaluation.* Princeton, NJ: RWJF.

Shadish, William R., Cook, Thomas D., and Leviton, Laura C. 1991. *Foundations of Program Evaluation: Theories of Practice.* Thousand Oaks, CA: Sage.

Smith, M. F. 1989. *Evaluability Study: A Practical Approach.* Boston: Kluwer Academic Press.

US Agency for International Development. 2010. "Conducting Mixed Methods Evaluations." Washington: USAID.

Weiss, Carol H. 1995. "Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families." In James P. Connell, et al. (eds), *New Approaches to Community Initiatives: Concepts, Methods, and Contexts* (Vol. 1), pp. 65-92. Washington: The Aspen Institute.

Winckler, Ole, Bull, Beate, and Kennedy-Chouane, Megan eds. 2013. "Evaluation Methodologies for Aid in Conflict." Routledge, (forthcoming).

EVALUATION REPORTS