

Evaluation of results measurement and how this can be improved

A stylized illustration of a balance scale, rendered in a light beige color. The scale is shown in a slightly tilted position, with the right pan hanging lower than the left, suggesting it is heavier. The background is a solid light beige color.



Norad

Norwegian Agency for
Development Cooperation
P.O.Box 8034 Dep, NO-0030 Oslo
Ruseløkkveien 26, Oslo, Norway

Phone: +47 22 24 20 30
Fax: +47 22 24 20 31

Design: Siste Skrik Kommunikasjon
ISBN: 978-82-7548-7894

Can We Demonstrate the Difference that Norwegian Aid Makes?

Evaluation of results measurement
and how this can be improved

April 2014

Itad in association with the
Chr. Michelsen Institute

Note on layout and language

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

Preface

Policy makers and the aid administration have had results measurement and reporting high on the agenda for many years. Several initiatives, procedures and tools aim at documenting the difference that Norwegian aid makes. Yet, while we have a good overview of the interventions supported by Norwegian aid, we have less information about how and to what degree the support has made a change. External reviewers such as the Office of the Auditor General of Norway and the OECD DAC have recently pointed to the need to improve the results measurement and reporting of Norwegian aid.

Weak results measurement does not necessarily mean that Norwegian development aid does not generate results, only that results have not been measured and reported on sufficiently. However, insufficient attention to and understanding of the results of aid, may prevent learning and can potentially result in poor aid effectiveness.

By assessing all stages in the aid management cycle - from planning to evaluation – this report looks into how the aid administration could improve its reporting on results. The evaluation concludes that there is room for improvement in many areas of aid management; from guidelines and training through to planning and approval of development interventions to reviews and evaluations. The report provides examples of how results can be documented and suggests ways to build on experiences from other agencies that are also working to improve their documentation of results.

However, while improvements in aid management are necessary, organizational changes are equally needed. The evaluation finds that results are not sufficiently requested from management, the staff does not have the time needed and they have few incentives to focus on results. Moreover, there are many other – and sometimes conflicting – priorities. Thus, although results are high on the agenda, results management appears not to be a priority in practice.

We hope this report may contribute to an informed debate on how to better document the results of Norwegian development cooperation. It is in everyone's interest – including the beneficiaries and Norwegian tax payers – to have better information on what Norway achieves by its aid.

The evaluation was commissioned and managed by the Evaluation Department of the Norwegian Agency for Development Cooperation (Norad) and carried out by Itad in cooperation with Chr. Michelsen Institute. The consultants are responsible for the content of the report, including the findings, conclusions and recommendations.

A handwritten signature in black ink, reading 'Tale Kvalvaag'.

Tale Kvalvaag

Director, Evaluation Department

Acknowledgements

This report has been prepared by a team drawn from Itad Ltd working in association with the Christian Michelsen Institute (CMI). The team consisted of Derek Poate (Team Leader) and Rob Lloyd (Deputy Leader) from Itad; Dr Espen Villanger from CMI and Dr Achim Engelhardt (Lotus M&E Group). Dr Merima Ali (CMI) contributed to the analysis while Ingrid Hoem Sjursen and Lars Gunnar Christiansen provided valuable research assistance. Jodie Ellis was project officer for Itad.

Derek led on aspects of the study design, the review of comparator agencies, and worked on the reviews of grants. Rob designed the overall approach, led on the analysis of grant management systems and procedures, and undertook reviews of grants. Espen led on the analysis of evaluation reports and consultants, working closely with Achim and Merima. They also reviewed a sample of grants.

Contents

Preface	iii
Acknowledgements	v
List of tables and boxes	viii
Acronyms and abbreviations	x
Executive summary	xv
SECTION A: Background and methodology	3
1. Introduction	5
2. Methodology and analytical framework	9
2.1 Description of the institutional set up that supports evaluability in the Ministry of Foreign Affairs (MFA) and Norad	9
2.2 Approach to the research	12
2.2.1 Assessment of grant management processes	13
2.2.2 Assessment of evaluation procedures and reports	14
2.2.3 Study of comparator agencies	15
2.3 Analytical framework	16
2.4 Limitations of the study	16
SECTION B: Evaluation findings	17
3. Results measurement and evaluability in the grant management system	19
3.1 Results measurement in the grant management process	19
3.2 Analysis of comparator organisations	22
3.2.1 Grant framework	22
3.2.2 Grant application and the planning stage	24
3.2.3 Quality assurance (QA)	24
3.2.4 Follow-up and monitoring	25
3.2.5 Completion	26
3.3 Policies, systems and procedures for grant management	28
3.4 Staff training and technical support	42
3.4.1 Findings on the training available to staff on results measurement	42
3.4.2 Findings on the technical support available to staff on results measurement	49
3.5 Implementation to ensure evaluability and results measurement	56
3.5.1 Findings on how well procedures on results measurement are being followed	56
3.5.2 Factors that influence attention to results in grant management	61

4. Results management and evaluability in EVAL commissioned reports	71
4.1 Planning, commission and quality assurance (QA) of evaluations	71
4.2 Competencies of evaluators	79
SECTION C: Conclusions and recommendations	83
5. Conclusions	85
6. Recommendations	91
6.1 Detailed recommendations on grant management systems	91
6.2 Recommendations for Norad's Evaluation Department	95
6.3 Options for implementation to address more structural issues	96
References	99
Annexes	105
Annex 1 Terms of Reference	107
Annex 2 List of persons consulted	117
Annex 3 Table summarising findings, conclusions and recommendations ...	121

The following annexes are available for download at www.norad.no/evaluation

Annex 4 Data and instruments
Annex 5 Comparator agency study
Annex 6 Methodology and analytical framework

List of tables and boxes

Tables

Table 1:	Distribution of number and value of grants in 2012	21
Table 2:	Results requirements across the Grant Scheme Rules	30
Table 3:	Minimum requirements for results measurement outlined in the Grant Management Manual	32
Table 4:	Information central to evaluability that is required as part of the planning of grants	34
Table 5:	Level of practical guidance offered by core documents on results measurement	36
Table 6:	Staff perceptions of the utility of policies and procedures in review grant applications in key areas of results measurement ...	37
Table 7:	Training courses available to staff that include a focus on results measurement	43
Table 8:	Key results concepts and themes covered in training courses	45
Table 9:	Survey responses to the question: 'How effective were training courses at building your skills in results measurement?'	46
Table 10:	UKS attendance records for results-focused training courses for 2012 and 2013 (to date)	47
Table 11:	Percentage of survey respondents that requested technical support on results measurement during the preparatory and follow-up phases of their most important current project/programme ...	50
Table 12:	Assessment of the content of the technical advice provided by Norad to embassies and the MFA on results measurement for the six of the 20 reviewed grants where advice was requested	51
Table 13:	Summary of results from six questions in the staff survey	64
Table 14:	Consultants' self-assessment of their expertise by evaluation approach	80
Table 15:	Consultants' self-assessment of their expertise by evaluation methods.....	81
Table 16:	Specification of consultants' expertise in Norad evaluation ToRs ...	82

Boxes

Box 1:	The difference between reviews, reports and evaluations in the Norad/MFA grant management system	11
Box 2:	Objectives of EVAL evaluation activities	12
Box 3:	Comments from staff on the Grant Scheme Rules	31
Box 4:	Comments on the minimum requirements on results measurement	38
Box 5:	Comments on the quality of evaluations	39
Box 6:	Comments on the barriers to training attendance	48
Box 7:	Comments on the need for practical guidance on results measurement	49
Box 8:	Comments on technical advice provided by AMOR on results measurement	52
Box 9:	Comments on the challenges faced by Norad staff in advising on results measurement	54
Box 10:	Examples of good practice in results measurement in grant management	58
Box 11:	Example of good practice for short reviews	61
Box 12:	Comments on the results orientation of the leadership	63
Box 13:	Comments on time available to work on results	66
Box 14:	Comments on the pressures to disburse funds	66
Box 15:	Comment on the incentives and sanctions around results measurement	69

Acronyms and abbreviations

AMOR	Avdeling for Metode og Resultater / Department for Quality Assurance (Norad)
AR	Annual Review (DfID)
ARPP	Annual Review of Portfolio Performance (World Bank)
DAC	Development Assistance Committee
Danida	Danish International Development Assistance
DFID	Department for International Development (UK)
DO	Development Objectives
EVAL	Norad's Evaluation Department
FGD	Focus Group Discussion
GMM	Grant Management Manual
GRPR	Global and Regional Programme Reviews
ICR	Implementation Completion Report (World Bank)
IEG	Independent Evaluation Group (World Bank)
ISR	Implementation Status Report (World Bank)
M&E	Monitoring and Evaluation
MFA	Ministry of Foreign Affairs
MTR	Mid-Term Review
NOK	Norwegian Krone
Norad	Norwegian Agency for Development Cooperation
OECD	Organisation for Economic Cooperation and Development
OPCS	Operations Policy and Country Services
PCR	Project Completion Review (DFID)
QA	Quality Assurance
QAG	Quality Assurance Group (World Bank)
QER	Quality Enhancement Reviews
RBM	Results-Based Management
RM	Results Management
SDI	Slum Dwellers International
SMART	Specific Measurable Assignable Realistic Time-related
ToR	Terms of Reference
UKS	Utenrikstjenestens kompetansesenter / Foreign Service Institute (MFA)
UNEG	United Nations Evaluation Group

Executive Summary



Executive Summary

Introduction

This report presents the findings, conclusions and recommendations from a study into the reasons why grants supported by the Norwegian Aid Administration are proving difficult to evaluate and therefore why it is hard to judge how successful Norway's aid programme is.

Analysis by Norad's Evaluation Department (EVAL) in 2011 found that none of the reports on grants that were evaluated could reach firm conclusions about the results being achieved. Reports showed well what money was being spent and what direct activities or services were being delivered. But critical questions about whether those services gave rise to real benefits for poor people and other target groups proved elusive.

This evaluation was framed around a series of questions, with hypotheses that could be tested: were the arrangements for planning results in grants adequately designed and specified; were staff adequately trained to manage for results in grant management; were policies and systems correctly implemented when grants were approved? The study also checked to see if problems were arising in the way EVAL designed and managed evaluations: did EVAL ensure evaluation designs placed an appropriate emphasis on measuring results; and were consultants recruited to evaluate sufficiently competent?

Methodology and limitations

Data collection was designed around three lines of enquiry. First, a comprehensive review of policies and guidance for new grants and training for staff developed during the period 2008-12. Then assessment of the quality of a sample of 20 grants implemented over the same period, and 20 end-of-grant reviews and evaluations. Four focus group discussions were held with staff from departments at the Ministry of Foreign Affairs (MFA) and Norad headquarters, followed by a survey of a sample of staff who work with grant management, drawn from the MFA, embassies and Norad. Some 126 staff responded in total.

A second area of analysis was to examine a sample of six evaluation reports commissioned by EVAL and completed since 2010, relating to the period reviewed for policies and guidance. A brief skills profile survey was administered to the consultants whose contact details were available. The third study area was a comparative desk review of grant management and evaluation policies, systems and procedures at Danida, the UK Department for International Development (DFID) and World Bank.

Sample sizes were set in the terms of reference (ToR). Although all grants and reports were selected at random, the small samples mean that the findings may not fully represent the diversity of actual experience and should be considered as case studies. All interviews were voluntary and survey responses were self-selecting. The possibility exists that respondents had a higher than usual interest in results management, but there are no *a priori* grounds for thinking these would be more for or against results procedures, hence any bias is not thought to be systematic.

Findings

Comparison with other development organisations

The comparison with DFID, the World Bank and Danida highlights some key differences. Norway does not prepare grants so there is less scope for interaction over results at the planning stage. Norway's requirements for objectives and indicators are less thorough and do not include a description of the links between activities, outputs and outcomes and the underlying assumptions (the intervention logic), a review of supporting evidence or an evaluation plan. The three comparators all have arrangements for quality assurance (QA) of individual grants or projects at or before appraisal. And reporting requirements by other agencies go beyond indicators to include a rating assessment of performance that can be used for comparative analysis.

Grant systems, policies and procedures,

There are 45 different Grant Scheme Rules and each has its own set of rules for results measurement. This has led to some confusing and inconsistent procedures, with little cross-referencing between relevant texts. Central guidance describes minimum content but not standards of that content for new applications. Applicants are not required to articulate their theory of change, indicate the evidence supporting their programme design and its overall credibility, or to describe their planned results measurement system.

Staff training and technical support

Although training courses on results are of good quality, there are gaps in content and they are not reaching enough staff to be effective. The use of technical support or independent appraisal of grant applications is mostly at the discretion of staff. Current arrangements with Norad's Department for Quality Assurance (AMOR) are valued for their quality, but the formality of procedures dissuades some staff and there is insufficient capacity for the potential demand.

Implementation of a results-focus in grants

Of the 20 grants reviewed, most did not have well-developed frameworks for results, reflecting the limited guidance and gaps in minimum standards. Staff think they possess the skills to review applications and monitor grant performance but argue that pressure of time and a low priority by senior management reduces their effectiveness. Reporting at the end of a grant is particularly weak and reflects confusing guidance and the absence of sound planning for evaluation when a project is designed.

Planning, commissioning and Quality Assurance of evaluations

The current arrangements for planning evaluations by EVAL tend not to generate high-quality reports. A variety of factors, such as large numbers of questions in the ToR, poor specification of evaluation objectives and a hands-off approach to managing consultants contribute to a divergence between the ToR and the deliverables in the final evaluation report. The overall assessment of the reports indicates that they are generally of a sufficient quality for providing information about outputs but not about causality.

Competencies of evaluators

The majority of the consultants have substantial experience with evaluation, and have formal training in the discipline. Many have a solid foundation in the application of core evaluation approaches and tools, but less so in the techniques for more advanced results analysis. It is not possible to determine the adequacy of consultants because ToRs do not describe the competencies that are needed for teams to deliver more rigorous impact or attribution evaluations.

Conclusions

We conclude that four of the five hypotheses in the ToR can be rejected by our findings:

- Current policies, systems and procedures are too fragmented, insufficiently comprehensive and do not provide the necessary guidance for staff.
- Training reaches too few staff, there are significant gaps in coverage and there is little supporting material. Formal advice and quality assurance is too limited to be effective at ensuring evaluability.
- Implementation of a results-focus fails to ensure evaluability, partly because there is little clarity about minimum standards, but also pressures of time on staff, low priority by senior managers and a lack of incentives to prioritise results.
- EVAL-commissioned evaluations are not designed and managed in a way that ensures they measure and report on results.

There was insufficient evidence to reach a conclusion about the hypothesis on the competency of consultants. Necessary competencies need to be expressed clearly in ToR and bidders assessed against those.

Recommendations

The findings in this evaluation lead to many potential recommendations. However, proposing too many changes might run counter to the Norwegian approach to development cooperation. For that reason, we present our recommendations in three parts. First, we present a candidate list of technical changes that would resolve the specific shortcomings or gaps in current grant policies, guidelines and operations. Secondly, we make recommendations

dealing with the work of Norad's evaluation department. Thirdly, we consider some options for implementation that try to match change with culture and working practices. Full details can be found in Chapter 6.

Detailed recommendations on grant management systems

- As part of the planning of grants, partners should be required to outline in greater detail how they plan to measure results. The details requested of partners should be expanded beyond indicators, baseline and objectives to include: a theory of change; a review of the evidence base that underpins the programme design (and an assessment of its quality); details of how data will be collected and analysed; whether reviews and/or evaluations will be commissioned and of what type; and the budget implications.
- More detailed and comprehensive guidance should be developed on how to put results into practice, specifically how to appraise results frameworks and support partners in developing effective measurement systems. This guidance should include 'how to' guides on: developing theories of change, assessing the quality of evidence that support a programme design, planning and managing grant-level evaluations and reviews, and how to appraise results management systems.
- Standard Quality Assurance checklists should be developed for staff to use when appraising results frameworks, progress reports and final reports. These checklists would, in effect, provide a clear specification of what the minimum requirements on results measurement detailed in the Grant Management Manual mean in practice, and how staff can make a judgement on what is considered good enough.
- Partners should be required to use the standard templates that have been developed, rather than using their own formats, to create greater consistency within the system and to ensure that appropriate level of detail is provided by a partner in their application, progress reports and final reports. These templates follow international good practice and can help strengthen national systems.
- A rating assessment should be added to the Progress and Final Report templates for partners to complete themselves. Reporting currently makes little use of indicators and does not require any judgement about performance. This is in contrast to the practice of the agencies reviewed for comparison. A rating is something that can be done by the programme officer even where the partner has poor indicators, and would help build awareness about evaluation.
- A more considered and strategic approach to the use of evaluations and reviews at the grant level should be developed. As part of the preparatory phase of a grant, greater consideration should be given to whether reviews and/or evaluations should be commissioned, and the budget implications. This should include a review of the existing base of evidence.

- Standard checklists should also be developed for quality assuring grant level evaluations and reviews. These should cover both quality at entry (ToR, inception report, etc.) and exit (final evaluation/review report).
- Develop a more comprehensive training programme to support staff capacity in results measurement. The training programme should offer more in-depth and longer-term training for those that want to deepen their skills in results measurement and evaluability. It should include greater focus on: theories of change, reviewing evidence, and appraising results measurement systems.
- The requirements on technical assistances and quality assurance should be harmonised across all grant scheme rules.
- An online resource hub should be developed that provides staff with access to examples of good practice in results measurement and pools sector-specific resources.
- Develop the capacity of grant recipients to measure results such as by *e-learning*, but perhaps also through a 'partners' guide to managing for results.

Recommendations for Norad's Evaluation Department (EVAL)

- Tighten the design specifications for evaluations. Draft ToR with tighter specifications for the purpose, objective and scope of evaluations so it is clear when outcome or impact is to be evaluated in addition to outputs.
- Keep evaluation questions focused. Reduce the number of evaluation questions to be covered so that resources are clearly prioritised for key results.
- Require evaluators to clearly describe the programme logic of the intervention being evaluated as a basis for the design.
- Be more specific in ToR about the required consultants' skills. More consideration should be given to the specific skills and expertise required for either the team leader or core team members.
- Monitor the progress of evaluations more closely. Once an inception report has been agreed, EVAL should plan periodic check-ins with evaluation teams to ensure the process is on track, and delivering according to what has been agreed.
- Develop a clear process for deciding and managing impact evaluations. To be most effective, impact evaluations should be set up at the design stage of an aid project and will require a joint decision with the implementers or partners.

- When conducting impact evaluations that seek to quantify attribution, ensure the appropriate competencies exist among the EVAL staff managing the evaluation and the consultants.
- Ensure the specification of methodologies, data requirements and competencies of evaluators are in line with the requirements for outcome and impact evaluations.

Options for implementation to address more structural issues

Different combinations of reforms are possible to tackle what we regard as a core weakness in the system: the fact that current guidelines are not being followed. We put forward two different approaches that can be used to address this problem: either by concentrating results expertise or by broadening it. Each has their own set of recommendations. The two approaches are not necessarily mutually exclusive. Both (or elements of both) could be taken forward in tandem. Three cross-cutting recommendations provide the foundations for both approaches.

Cross-cutting recommendations

- Strengthen the support at senior manager level for results measurement. There needs to be more visible action among the leadership of the MFA and Norad to insist on better results management and evaluability (and provision of the resources that are required to support this).
- There should be a clear requirement for mandatory technical assistance and quality assurance of partners' results frameworks, reporting and evaluation plans, for all grants greater than a certain amount, and sample assessment of mid-size grants.^{1,2}
- Improve staff incentives for measuring results. A strong incentive would be to incorporate continuing professional development in results management as a positive career attribute to be recorded on personnel files and factored into career development.

The final part of our recommendations is to propose the two contrasting approaches to results and evaluability.

1 The precise limits would need to be agreed, but considering the distribution of project and programme grants in 2012, setting the mandatory requirement at greater than NOK 25 million would mean reviewing 83 grants and cover 64 percent of the committed value. Depending on the available staff and workload, a limit of NOK 15 million might be practical. Mid-range grants could then be NOK 5 million to NOK 15 million. These are the same limits that have been used under past arrangements.

2 We are aware that this is the approach taken in the old Development Cooperation Manual. This is a consistent and easier approach for staff to work with.

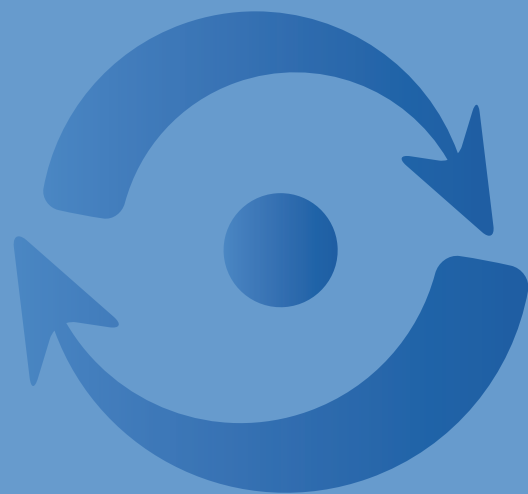
Approach 1: Concentrating expertise. This would involve re-designing the approach to quality assurance expanding the role of the Department for Quality Assurance (AMOR) so that all eligible grants receive quality assurance on their content before approval. This approach would bring a high degree of consistency, but has substantial implications for staffing in the short time span necessary to process grants. The recommendations that would need to be taken forward are:

- **Resource AMOR to provide more comprehensive support to all eligible grants on its results measurement frameworks and evaluation plans at approval and completion.** This could be contracted out or could be staffed as temporary teams of peer reviewers such as were used in the World Bank Quality Assurance Group (QAG).

Approach 2: Broadening expertise. This would involve mainstreaming skills to a large number of staff across Norad, the MFA and the embassies, and would bring advice closer in space and time to grant managers. This approach is designed more to work within current practices and build through progressive rather than radical change. It broadens involvement by developing staff capacity and working more closely with partners, which would help build their capacity as well. The recommendations that would need to be taken forward are:

- **Build a cadre of staff specialised in results measurement and evaluation.** We recommend, alongside an improved training programme, that a small number of staff from across the organisations are given intensive training in evaluation and results measurement. Embedded advisers would offer a more flexible and informal form of peer support to staff around how to measure results. To encourage participation, incentives should be created to bring career rewards.
- **Design a new approach to outreach to partners, with a combination of improved technical guidance, some direct capacity building and access to the self-study materials.** More focus should be put on developing the capacity of partners to improve the measurement and reporting of results.

Main Report



SECTION A:

Background and methodology

1. Introduction

The catalyst for this evaluation was the finding by Norad's Evaluation Department (EVAL) that 'none of the evaluations and studies commissioned by EVAL and finalised in 2011 could report sufficiently on results at the level of outcomes or impact,' (Terms of Reference, p.1). EVAL commissioned this evaluation to understand why this is the case and to better understand the factors and dynamics that are at play within the Norwegian Aid Administration that support partners in, and prevent them from, measuring results.

The purpose of the evaluation is to contribute to further learning and progress in the aid administration's follow-up of the government's demand for a strengthened focus on results by identifying reasons for the insufficient results documentation, and provide evidence-based recommendations for improvements in this area.

To explore this issue, the evaluation was set up to test a series of positive statements, or hypotheses.³ These were divided into two groups. The first set of hypotheses relate to the policies, systems and practices of grant management within the Norwegian Aid Administration. EVAL reports draw heavily on the results reported by individual grants. Therefore, if parts of the grant management system are not functioning effectively, and grant level reviews and evaluations are either lacking or of poor quality, this would be a contributing factor to EVAL reports not being able to report on results. The three hypotheses related to grant management are:

1. Internal policies, systems and procedures to ensure evaluability and results documentation in the grant management process provide appropriate and comprehensive guidance.
2. Staff receive appropriate training and technical advice/support to effectively ensure evaluability and results documentation as part of the grant management process.
3. The policies, systems and procedures that are in place (to ensure interventions are evaluable and robust results data are being collected) are being correctly and adequately implemented.

³ The hypotheses were reformulated in the inception report from negative to positive statements. They are also being analysed here in a different order than they were originally presented in the terms of reference

The second line of enquiry pursued in the evaluation is that the insufficient reporting of results in EVAL reports is related to the functioning of EVAL itself. The thinking here is that perhaps the way that EVAL plans and manages evaluations is a contributing factor to reports not being able to sufficiently show results, or that the skills of the consultants that are hired to conduct the evaluations might be a factor. The two hypotheses related to EVAL that were tested are:

4. The planning, commissioning and quality assurance of evaluations places an emphasis on measuring results.
5. Evaluators have adequate competencies to effectively measure results and find/use evidence.

To test these hypotheses we undertook eight lines of enquiry:

- Desk-based document assessment of guidelines, handbooks, rules etc. dealing with evaluability and results measurement.
- Assessment of courses, training and capacity building of staff.
- Review of grant management processes including quality assurance.
- Comparative review of approaches for evaluability and results management at the World Bank, DFID and Danida.
- Review of current practice through a sample of recent grants.
- Interviews and a survey of staff to explore current practice and opinions.
- An assessment of the results-focus in a sample of recent evaluations by EVAL.
- An assessment of evaluators' competencies in the sample of evaluations.

The lines of enquiry centre on the concepts of results and evaluability. Our understanding of these is as follows:

- **Results** are 'the output, outcome or impact (intended or unintended, positive and/or negative) of a development intervention.'⁴

4 This definition is taken from the OECD/DAC glossary of terms in evaluation and results-based management, 2002.

- **Evaluability** is the degree to which an intervention is possible to evaluate in a reliable and credible manner (i.e. availability and reliability of data).⁵ This includes whether there is 'clarity in the intent of an intervention to be evaluated' (e.g. existence of a theory of change or intervention logic).⁶ Importantly, evaluability is not associated with any particular approach or method of evaluation.

The report is divided into six chapters, grouped into three main sections:

Section A covers the background to the evaluation and includes the introduction in Chapter 1, and a description of our methodology in Chapter 2.

Section B contains the findings of the evaluation. It starts with Chapter 3, which addresses the three hypotheses concerning results measurement in grant management. It first describes the Norwegian grant management system (3.1); it then presents the results of a comparison with three comparator agencies (3.2). It then deals with policies and systems (3.3); followed by a review of training (3.4); then finally, it looks at the actual performance of results measurement in grants in recent years (3.5). Next we examine, in Chapter 4, the hypotheses concerned with EVAL's work. First, we examine how the department manages its commissioned evaluations (4.1), then we review evaluators' capacities (4.2).

Section C draws together conclusions from the evaluation in Chapter 5, and presents recommendations in Chapter 6.

A large body of evidence is provided in supporting annexes: Annex 1 contains the terms of reference (ToR); Annex 2 lists people interviewed under the study; Annex 3 presents a table summarising the findings, conclusions and recommendations from the evaluation; Annex 4 contains all the main references to data and data collection instruments; Annex 5 presents findings from a review of three comparator agencies to see how they perform with respect to results in grant management; Annex 6 provides details about the methodology and analytical framework. Annexes 1-3 are found at the back of this report. Annexes 4-6 are in a separate volume of the report which is available electronically at www.norad.no/evaluation.

⁵ This definition is taken from the terms of reference for the evaluation.

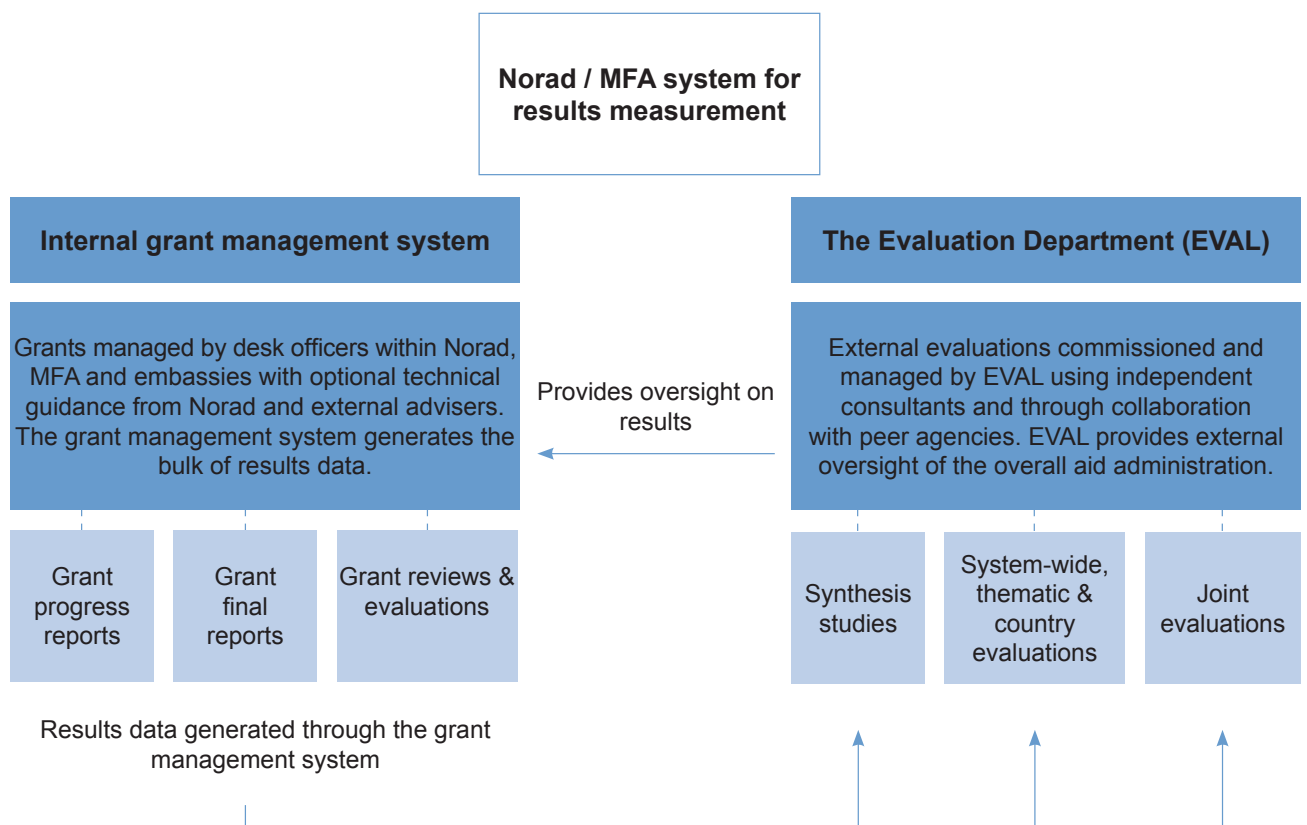
⁶ This is a key element of the United Nations Evaluation Group (UNEG) understanding of evaluability: 'Before undertaking a major evaluation requiring a significant investment of resources, it may be useful to conduct an evaluability exercise. This would consist of verifying if there is clarity in the intent of the subject to be evaluated, sufficient measurable indicators, assessable reliable information sources and no major factor hindering an impartial evaluation process.' (UNEG Norms for Evaluation in the UN System, April 2005).

2. Methodology and analytical framework

This chapter summarises the research strategy and methods, with references to supporting material in annexes. To start with, we describe the institutional arrangements that have a bearing on evaluability. This sets out the main organisational and institutional features that underpin the systems and processes which we analyse. Next, we present an overview of our research methods, and then explain the three main strands of work that were undertaken. Finally, we conclude with a note on the analytical framework.

2.1 Description of the institutional set up that supports evaluability in the Ministry of Foreign Affairs (MFA) and Norad

Figure 1: Overview of the results measurement system within Norad and the MFA



As outlined in Figure 1, the system for results measurement in Norad and the MFA has two components: the internal grant management system, which generates the bulk of results data, and EVAL, which provides independent oversight of results. The **internal grant management system** guides how Norad and the MFA manage the finances disbursed through the Norwegian aid budget. Grants are the basis of results data within the Norwegian aid administration. The responsibility for measuring and reporting on results sits with the partner and takes place through periodic progress reports and final reports at the end of a grant.

The relationship with partners is managed through a programme officer. Programme officers are responsible for ensuring results are reported effectively and in a timely manner and are responsible for ensuring overall quality. In undertaking this role they can draw on technical support from Norad or externally.

Within the Norwegian system individual units are responsible for their own quality assurance (QA), with formal QA provided through the AMOR Legal Section, while technical guidance and advice can be requested from Norad Sector Advisers, the AMOR Results Management Section and/or EVAL. Quality Assurance by the Legal Section takes place during the preparatory phase of a grant and involves checking the grant Agreement (the contract), including its compliance with the relevant Grant Scheme Rules and the Grant Management Manual.

In some cases this may also involve a review of a grant's results framework. In these cases the Results Management Section is consulted. Programme officers can request technical advice at any point in the grant management process. Whether it is requested is mainly at the discretion of the individual programme officer. In the few cases where it is mandatory, this is stipulated in the Grant Scheme Rules. The Grant Scheme Rules also state when formal Quality Assurance from the AMOR Legal Section is required.

As part of the management of grants, reviews and evaluations can also be commissioned. They are complementary to ongoing results monitoring and allow a more in-depth analysis of performance (see Box 1 for an explanation of the difference between reviews and evaluations). Reviews and evaluations are an important component of the ongoing monitoring of grants, and can be commissioned by either the partner or the programme officer. Whether they are to be conducted is often included in the agreement with a partner.

Box 1: The difference between reviews, reports and evaluations in the Norad/MFA grant management system

Reviews⁷

Relatively light-touch exercises based on existing documentation and a limited amount of primary data collection that examine programme effectiveness and document lessons learned. They are described in the Management Manual, section A16, page 66. Reviews can be conducted during programme implementation to identify if a programme is on track (mid-term reviews) and/or at completion to assess whether a programme has reached its objectives (end reviews). Special reviews can also be conducted that look into specific challenges during implementation; for example, related to environmental or social issues such as gender and poverty.

Reviews can be led by external consultants, internally, or through a mix of internal and external stakeholders. The need for regular reviews is assessed in the appraisal phase and included in the agreement with a partner. Reviews, however, may also be commissioned *ad hoc* when the Grant manager, Partner or other donors find it important.

Reviews are rarely mandatory. They depend on the rules for the particular scheme under which the grant is being awarded and are at the discretion of the programme officer.

Progress and final reports

Progress reports are a summary of a project's activities and results over a specific period and must be submitted by the grant recipient at intervals specified in the agreement, as governed by grant scheme rules. A detailed progress report form template (S11/S61) may be used or followed as a checklist.

Final reports are mandatory for all grants and are to cover the whole project and support period, not just the Ministry/Norad's contributions. The report is the grant recipient's own presentation and assessment of the project results and thus a form of self-evaluation. A detailed final report form template (S21/S81) may be used or followed as a checklist. The progress and final report templates link to the template for a grant application and provide a coherent sequence of information about planned and actual performance.

Evaluations

Evaluations are independent, comprehensive assessments based on systematic data collection. They are always conducted by external consultants. Evaluations are carried out for learning and accountability purposes, and to ensure that programmes are relevant, effective and efficient.

Evaluations are supposed to satisfy internationally agreed quality standards such as the OECD/DAC's Quality Standards for Development Evaluation. Evaluations are normally agreed upon between partners, and will often be referred to in the agreement. They may, however, also be commissioned by the Grant Manager (i.e. the MFA, the Embassy or Norad), co-financing donors or a partner.⁸

⁷ This definition is taken from the 2005 Development Cooperation Manual. The 2013 version of the Grant Management Manual states simply 'A thorough assessment of project or programme, with focus on performance in relation to plans and goals.'

⁸ Ibid. The 2013 version of the Manual states 'Systematic and objective assessment of an ongoing or terminated project/programme or grant scheme conducted by an external entity.'

The second pillar of the Norwegian aid administration's approach to results measurement is EVAL. EVAL is a semi-independent entity, located within Norad, responsible for initiating and implementing evaluations of activities financed through the Norwegian aid budget. It reports directly to the assistant director-general of the MFA (see Box 2 for EVAL objectives).⁹ As Figure 1 indicates, the data for EVAL-commissioned evaluations comes from grant-level results monitoring. While evaluation teams will conduct primary and secondary data collection, the results data generated through progress reports, final reports and grant reviews and evaluations provides the foundation for any evaluation exercise.

Box 2: Objectives of EVAL evaluation activities

- Evaluate goal achievement and results relative to adopted plans;
- Evaluate whether the consumption of resources is commensurate with the results achieved;
- Systematise experiences in order to quality-assure and improve future activities through effective learning processes; and
- Provide information to funding authorities and the general public.¹⁰

2.2 Approach to the research

The overarching logic of this evaluation is deductive. A set of hypotheses were developed by EVAL as part of the approach analysis for the evaluation. This evaluation has tested those hypotheses through observation and analysis to confirm or reject them. The underlying theory on which the hypotheses are based (not articulated in the ToR) can be summarised as follows:

'The evaluability of a grant is determined by the extent to which the planned intervention is designed around a clear explanatory logic that specifies: the programme theory by which resources translate into outputs, which in turn stimulate changes in behaviour. The intervention logic should draw on evidence from either social or natural science theory, or supporting information from similar interventions in other places or times, and should take into account contextual factors and the potential risks to the intervention.'

Evaluability may be high, yet evaluations fail to determine outcomes. This might be the cause of poorly specified evaluation studies or inadequate practice and competencies of evaluators. This last point formed the basis of two supplementary hypotheses, also tested under the evaluation.

⁹ According to the department's mandate, it should report to the director-general, via Norad's director-general, while in practice it reports to the assistant director general, who has been delegated the responsibility of the Ministry's development cooperation.

¹⁰ EVAL mandate for evaluating the Norwegian Development Administration, http://www.norad.no/en/evaluation/_attachment/393700?_ts=138d70a555a&download=true

The methodology for the study has three main components. To preserve reader-friendliness, these are described in full in Annex 6 and in summary here:¹¹

- Assessment of grant management processes (2.2.1)
- Assessment of EVAL (2.2.2)
- A desk review of systems and procedures in three comparator agencies (2.2.3)

2.2.1 Assessment of grant management processes

The assessment was carried out in a sequence of six stages:

- **Step 1** – Map grant management systems to identify key steps across the multiple processes and guidance currently in use.
- **Step 2** – Assess the quality of grant management¹² processes and systems for ensuring evaluability using a checklist derived from current practice across a wide range of development organisations (see Annex 4 for the checklist template) followed by interviews with key informants.
- **Step 3** – Assess the quality of results-based management training for staff, looking at scope and content with a particular focus on to what extent it provided support in key areas of evaluability.
- **Step 4** – Assess the practices of evaluability and results measurement across a sample of 20 grant-funded interventions. Details are given in Annex 6, Box 1. The grants were reviewed using a composite checklist that combined policy compliance, evaluability and quality assurance. We also assessed 20 grant-level end/mid-term reviews and evaluations using a composite checklist that assessed the quality and credibility of their analysis and findings.
- **Step 5** – Survey a wider sample of staff to test emerging theories.
- **Step 6** – Validate the findings with key stakeholders through four focus group meetings in Oslo, separately for clusters of staff from different departments in the MFA and Norad.

Staff survey and focus groups selection bias

Participation in the focus groups was through an open invitation directed through heads of departments. It is not known the extent to which the invitation reached all levels of staff, nor the guidance that was given about participation. It was clear in some meetings that section heads were present rather than more junior staff.

¹¹ Further details can be found in the inception report for this evaluation, available upon request from EVAL.

¹² The term grant management is used for the Norwegian *tilskuddsforvaltning* and includes the entire grant management process: appraisal/approval/decision documents, results measurement, quality assurance, reviews and evaluations.

Those that attended were experienced staff members, mostly thought to be mid- or late-career professionals with experience of grant management, often at embassy as well as headquarters level. Participation in the staff survey was voluntary because the survey was transmitted directly to individuals by email. We have considered the question of selection bias in both the focus groups and survey and do not consider it to be a problem for the interpretation of our findings. More details are provided in Annex 6.

Throughout the report we use a range of quotes taken from the focus groups and interviews. To ensure anonymity we have not included the names of individuals or their institutional affiliation. We have also not shown whether the quote was sourced from a focus group or interview.

2.2.2 Assessment of evaluation procedures and reports

The approach for this part of the evaluation was to examine the quality of a sample of reports from EVAL's Evaluation Reports series and to check the processes followed against EVAL's documented standards.

A sample of **six evaluation reports with a results-focus out of the 37 reports completed in the past three years** was selected, and a detailed assessment made of how the processes and systems for commissioning, managing and quality assuring evaluations have operated in practice. The sample size followed guidance in the ToR, which indicated a sample of five evaluation reports.

Our approach applied **backward induction** to reveal the critical factors that have led to the quality of each of the final evaluation reports. Backward induction is a method where reasoning backwards in time from the final outcome (or output) to the previous step all the way to initiation will reveal the sequence of actions that led to the result of interest.¹³ This approach was supplemented with an assessment of whether EVAL is sufficiently results-focused in its planning of evaluations.¹⁴

Avoiding conflicts of interest

The nature of this evaluation meant that at times, particularly in relation to our testing of Hypotheses 4 and 5, the evaluation team was required to assess the work and competencies of other consultancy companies, some of which both Itad and CMI compete against for work. Given the perceived conflict of interest that this posed, it was important that we managed this process sensitively. To do so we took a number of steps. First, we defined clear criteria upon which our judgements of evaluation quality and competency were based. These were drawn from current best practice. The criteria we used can be found in Annex 4. Second, we shared all of the original assessments with EVAL with our

13 See Villanger (2005), <http://www.sciencedirect.com/science/article/pii/S0014292105000164> for a formal application of backward induction in foreign aid relations. For a useful introduction to backward induction as a tool in wider applications, see http://en.wikipedia.org/wiki/Backward_induction.

14 Even if all evaluations are of a high quality, one could still have a situation where none of them address outcomes or impacts.

justifications for scores, so that if someone wished, they could review and verify our work. Third, in an effort to reduce bias within the scoring, the assessments of evaluation reports were reviewed by multiple team members.

ITAD and CMI have also been involved in a number of evaluations for Norad in the past three years. In our original technical proposal we set out an approach to avoid any possible conflicts of interest and identified four contracts we had worked on to be excluded from our analysis. When we came to sample evaluation reports we realised that further exclusions were necessary, as explained in Annex 6.¹⁵ A total of nine reports were excluded from the population of 37 to avoid conflicts of interest. All in all, the exclusion of reports, the openness and transparency in all assessments and the reliance on documented best practice create a verifiable process. Any perceived issues of conflict of interest can be reviewed by outside parties.

2.2.3 Study of comparator agencies

By a combination of guidance in the ToR and discussions during the Inception Phase, three agencies were selected for an analysis of comparator systems and procedures. These were Danida, DFID and the World Bank.¹⁶ In addition, some information was reviewed from an assessment of results-based management in the Finnish Ministry for Foreign Affairs.

Data for the comparisons were collected through a combination of **documentary evidence** (such as policies, processes and structures that guide results management in the grant-making or lending process and the functioning of the evaluation units) and **telephone interviews** to clarify points with key informants. The findings reflect a review of what the organisations say they do (policies and systems) rather than their actual practices, which would be beyond the scope and resources of this study.

We recognise that there are differences between the three agencies and Norway's aid administration. The comparisons we have drawn focus on approaches to ensure clear specification of objectives and results, and arrangements to achieve evaluability of grants and projects. These are common aspects of good public sector management and transcend differences in modes of operation or political guidance.

15 We have not excluded reports where Itad and CMI competed for the assignment. Among the selected six reports, Itad and CMI were competing for five of them. We believe that the measures taken as explained above are sufficient for avoiding conflict of interest.

16 DFID and the World Bank were specified in the terms of reference. We suggested that one other donor within the Nordic countries would provide a closer contrast from a country with comparable institutions and aid management. Sweden was rejected because of substantial changes in aid policy in recent years, and with the decision to close the Swedish Agency for Development Evaluation (SADEV) it may not have provided a stable recent experience from which to draw lessons. The approach of Finland for project cycle management closely follows EC guidelines and was reviewed within the past two years in a study of results management. Our analysis drew on that study without the need for further investigation. The proposed comparator was therefore Denmark.

2.3 Analytical framework

In designing this study, we gave careful thought to the challenge of drawing rigorous conclusions from a small number of enquiries based around mainly qualitative information. We have followed a realist approach, which postulates that outcomes are a result of the interaction between an intervention mechanism and the context in which it is applied.

In the case of this evaluation, the intervention is the steps and procedures adopted by the MFA and Norad to ensure the quality of grants being approved for financial support. The institutional structures are represented by the systems, guidelines and procedures by which staff are advised to interact with grant applications and grantees to ensure an appropriate specification of results. Contextual and cultural factors are the checks and balances, such as arrangements for quality assurance, and the *de facto* prioritisation given to results management through leadership, incentives and institutional culture.

2.4 Limitations of the study

Data collection was designed around three lines of enquiry. First, a comprehensive review of policies and guidance for new grants, training for staff and assessment of the quality of a sample of 20 grants out of 12,000 implemented between 2008 and 2012 and 20 end of grant reviews and evaluations. In order to work with a broad population of grants and to allow for a full implementation cycle from design to completion we elected to use this time period.

We recognise that practice followed for the planning and management of grants was evolving through this period and that can be seen in the attention to results and risk management in the updated Grant Management Manual in 2013, especially through the revision of V04 'Guide to management of results and risks'. Approaches were also harmonised across different grant schemes at the same time. In view of the changing guidance, for policy compliance we used the Development Cooperation Manual, which had been in circulation since 2005. For evaluability and quality assurance we drew on international good practice.

Sample sizes were set in the ToR. Although all grants and reports were selected at random, the small sample sizes mean that the findings may not fully represent the diversity of actual experience. All interviews were voluntary and survey responses were self-selecting. The possibility exists that respondents had a higher than usual interest in results management, but there are no *a priori* grounds for thinking these would be more for or against results procedures, hence any bias is not thought to be systematic.

The analysis in the comparative desk review only looked at the specification of policies, not how they are actually implemented in practice and their influence on the quality of development results.

SECTION B:

Evaluation Findings

The following section presents the findings of the evaluation. It is structured in two parts. First, in Chapter 3 we explore the procedures and practices of results measurement within the grant management system to understand the extent to which results are being measured, and explore what is working well and any possible blockages and barriers.

The second part of our findings, Chapter 4, is focused on Norad's Evaluation Department and explores whether the insufficient reporting on results in their reports is a product of dynamics internal to the way the department works.

The key findings from each chapter are summarised at the end of each section.

3. Results measurement and evaluability in the grant management system

The following chapter is focused on the functioning of the grant management systems. We first describe how results are measured through the grant management process (3.1); we then outline the findings from our assessment of the grant management process, or equivalent, within three comparator agencies (the World Bank, DFID and Danida) (3.2); next, we present findings from our review of the appropriateness and comprehensiveness of the policies, systems and procedures for ensuring results measurement (3.3); following that are detailed findings on the appropriateness of the technical advice and training that is available for staff (3.4). Lastly, we discuss the extent to which policies and systems for ensuring results reporting are being put into practice, and present an analysis of the factors which both support and prevent this from happening (3.5).

The overall aim of this part of the assignment is to come up with recommendations that can strengthen how results measurement and evaluability are embedded within the grant management system and support learning within Norad and MFA.

3.1 Results measurement in the grant management process

The grant management process is divided into three broad phases: preparatory (which includes entering into agreement), follow-up and completion. In each phase, there are activities that are mandatory for all grants and activities that are determined by the scheme rules that a grant sits under. These are detailed in Figure 2, using solid lines and dotted lines, respectively.

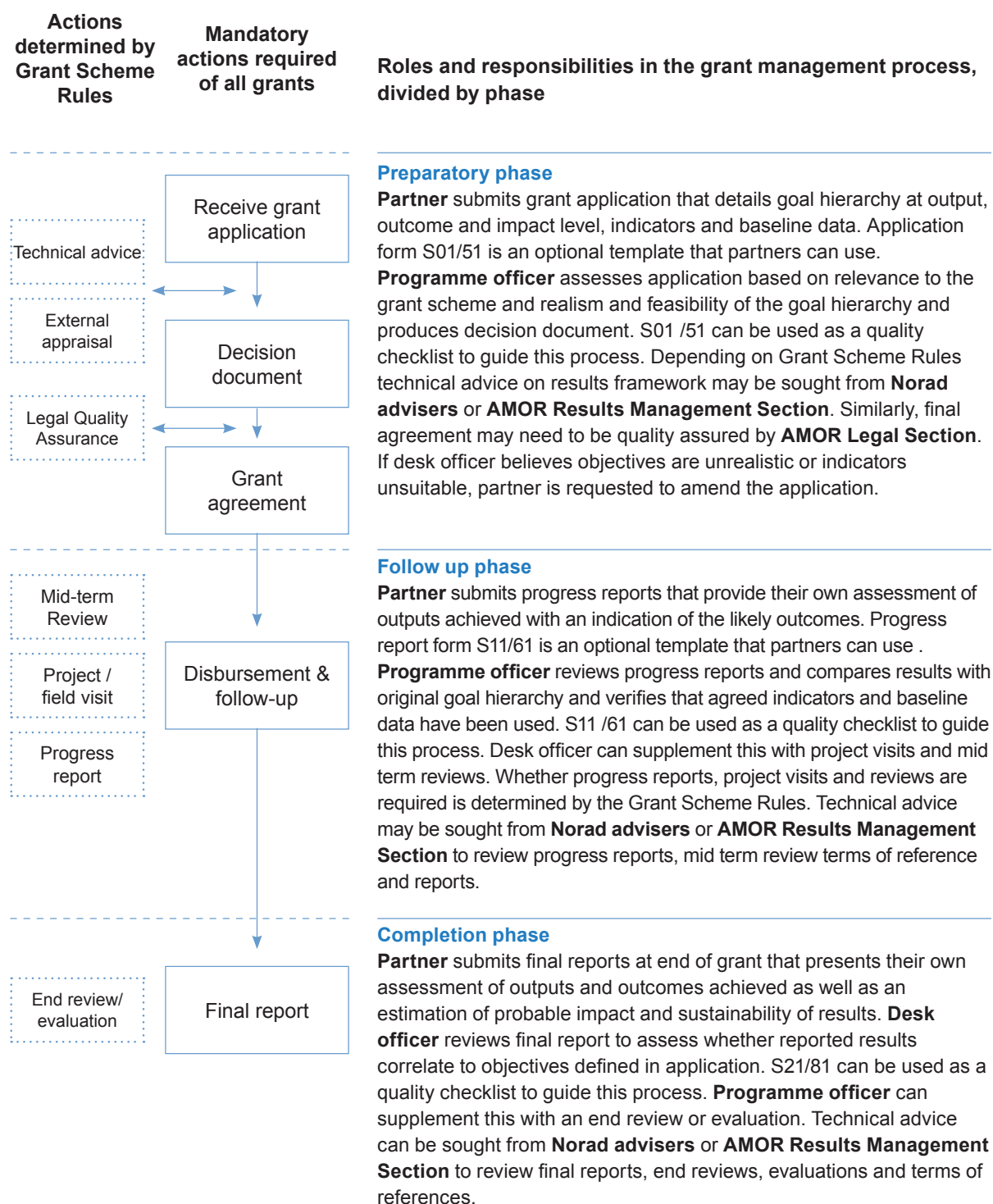
Across the three phases, **partners** are responsible for articulating how they will measure results in their application to Norad or the MFA and providing self-assessments of their progress through progress reports and final reports. Relevant templates are provided for each of these stages: applications (S01/51), progress reports (S11/61) and final reports (S21/81).¹⁷ Their use by partners is however, optional.

Programme officers in the MFA, Norad or an embassy are responsible for ensuring that the results framework and subsequent reports are of a sufficient quality to ensure results are documented in a robust and credible way. In doing this, programme officers play both the role of quality assurer and also of capacity builder. As discussed previously, at each phase of the grant management process programme officers can request, and in some cases are

¹⁷ Templates S01, 11 and 21 are the Norwegian language versions of the documents, S51, 61 and 81 are the English language versions.

mandated to receive, technical advice from Norad on the results framework, progress and final reports. Optional technical advice can be sought from the **AMOR Results Management Section** or other **Norad advisers**.

Figure 2: Outline of the grant management process, including key actions, roles and responsibilities, divided by phase¹⁸



¹⁸ Details of the grant management process have been taken from the 2013 Grant Management Manual.

Before presenting our analysis of grant procedures it is also worth summarising the overall grant portfolio that staff within the MFA, Norad and embassies manage. This helps explain the relative size and importance of different strands of grants. Table 1 summarises grants committed in 2012.

The Type I, Project and Programme Grants, which are the subject of this study, are the most numerous and have the greatest share of value of all grants. Small grants (Type III) are also numerous but of low total value. The remaining types (II and IV to VI) concern other methods of disbursement and are not the concern of programme officers. Within the Project and Programme grants, it is important to point out that the distribution of grant size is skewed. Four out of five are less than NOK 10 million in value; the largest 19 percent of grants accounts for 82 percent of all grant disbursements. Thus, the effort of programme staff on managing for results needs to be considered against the number and size of grants they are dealing with. It is not possible for staff to apply the same level of results-focus across the entire system and it seems reasonable that some prioritisation is given to larger grants.

Table 1: Distribution of number and value of grants in 2012

Number and value of grants committed by grant size (million NOK) in 2012							
Grant type		Size					Total
		1<	1-5	5-10	10-25	25<	
I - Project and programme	Number	422	383	160	146	81	1,192
	Value	178	903	1,081	2,192	7,684	12,037
III - Small-scale grant	Number	816	89	18	6	0	929
	Value	178	184	111	100	0	573
All other grants	Number	77	58	19	29	39	222
	Value	14	156	128	429	10,358	11,084
Total	Number	1315	530	197	181	120	2,343
	Value	370	1,243	1,320	2,720	18,041	23,695

Note: Grant types II and IV to VI concern other methods of disbursement and are not included in the table.

Major findings about institutional arrangements and the portfolio of project and programme grants

- The responsibility for ensuring plans and arrangements for a grant meet requirements for intervention logic, results and evaluability rests with grant managers at the MFA, Norad and the embassies. EVAL relies on information about grants generated from grant reviews and reporting for their evaluations.
- Neither progress reports nor grant reviews are mandatory but depend on requirements under each grant scheme's rules.
- Programme officers face a skewed workload, with 80 percent of all the grants they manage accounting for only 18 percent of the value of the total portfolio.

3.2 Analysis of comparator organisations

In order to set the Norwegian approach in a wider context, we looked at the grant management and QA procedures of the World Bank, DFID and Danida. Details of this analysis are in Annex 5. We compared the processes of grant approval, implementation and evaluation of these three organisations with those of Norway, as outlined in the introduction in Chapter 1 and analysed in detail later in this chapter. Figure 3 summarises the *de jure* actions as defined in manuals, handbooks and guidelines.

The stages do not show in full detail every aspect of each agencies' systems, because these are complex with numerous detailed variations according to type and size of loan or grant. Nor have we been able to investigate how well those organisations actually follow their stated procedures. However, the stages shown have been chosen to highlight those fundamental steps in the process where comparisons about a focus on results and evaluability can be drawn between the approach adopted by the MFA and Norad, and those of the comparator agencies.

The shaded boxes indicate process stages where a comparator has an approach that has some characteristics that contrast with the Norwegian system. Although there is, unsurprisingly, much common ground in the processing stages among all four organisations, key findings have emerged about the way evaluability is handled which we discuss under five headings:

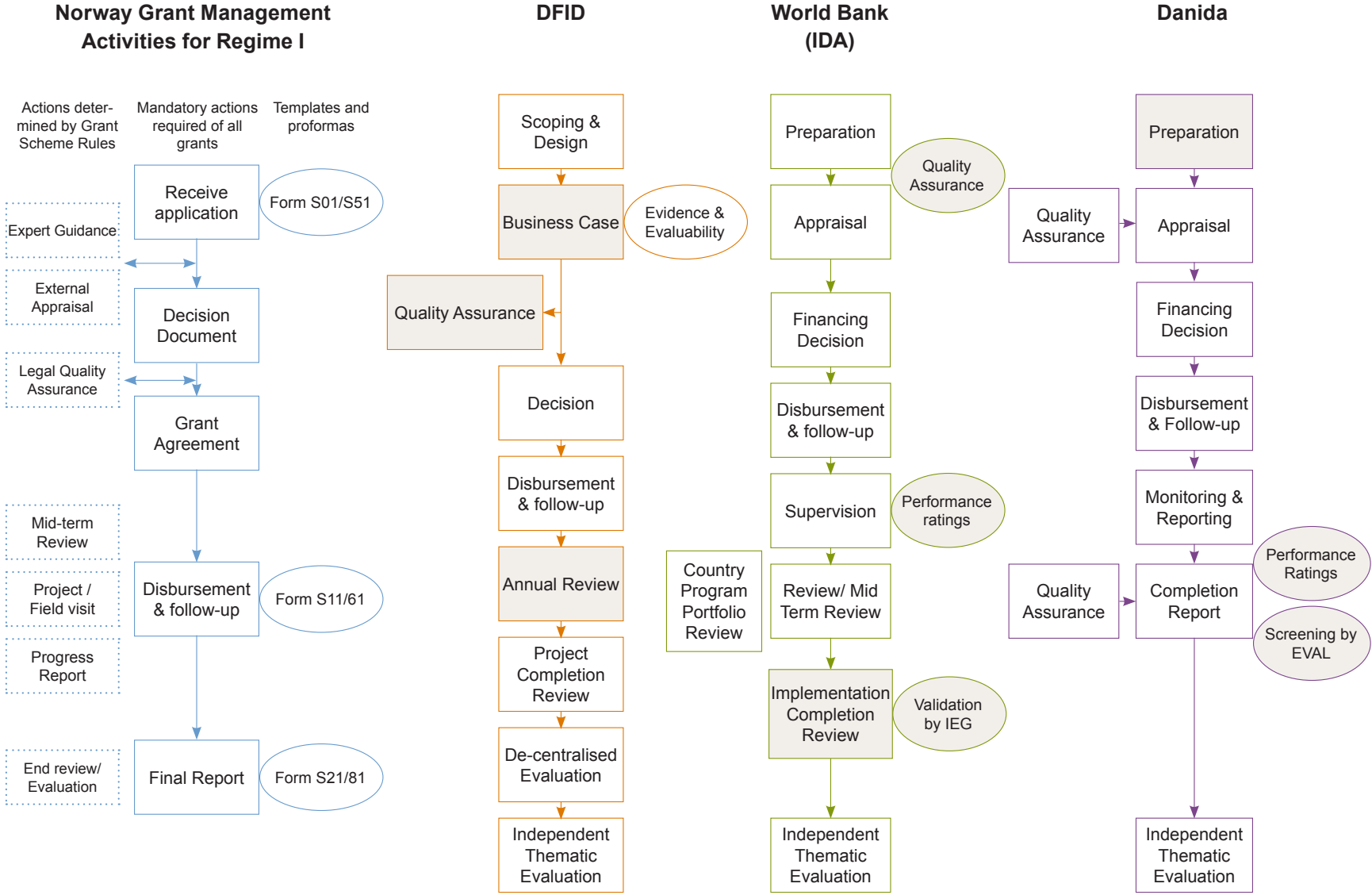
- The granting framework (3.2.1)
- Grant application and the planning stage (3.2.2)
- Quality assurance (3.2.3)
- Follow-up and monitoring (3.2.4)
- Completion (3.2.5)

3.2.1 Grant framework

Neither Norad nor MFA staff are responsible for the design of projects. A fundamental difference between the Norwegian aid administration and other agencies is that staff of the MFA and Norad do not prepare projects. In some cases, Norway receives applications from potential grant recipients and negotiates the objectives, work plan and financing of the project. In others, such as for grants contributed from the budget lines of UN departments, support is initiated based on Norwegian policies matching the identified need for intervention and results. Then the quality, competence and capacity of potential partners are assessed and a dialogue concerning implementation is established. Among the comparator agencies, responsibility for grant/preparation rests more substantially with staff of the agency themselves.

Basic specifications about Norway's grants are contained in Grant Scheme Rules, of which there are 35 for the MFA and 10 for Norad. These define exactly which actions are mandatory and which are not. Since 2010, Norway has had a standard template proforma (form S01 in Norwegian, S51 in English) for an applicant, which includes details about the objectives of a grant and the indicators to monitor performance. The form can be completed online or used as a checklist against a grant applicant's own documentation. The use of the template is optional. Norway's large numbers of grants of varying sizes means that guidance about design is not always closely followed and, as noted earlier, might be better concentrated on larger or more prominent or riskier grants.

Figure 3: Comparison of grant/project processing procedures among four development organisations



Note: The shaded boxes indicate process stages where a comparator has an approach that has some characteristics that contrast with the Norwegian system.

3.2.2 Grant application and the planning stage

All three comparator agencies specify requirements about results and evaluability in more detail than Norway. If the provisions of Norway's grant application template form S51 are followed correctly, then a grant applicant will set out a hierarchy of objectives together with planned indicators for follow-up. But guidance is very brief, with no detailed explanatory text or examples. All the comparators have some variations in approach to project processing according to size or purpose of grant or loan.

DFID, in its current project document the 'Business Case', requires project managers to provide evidence to support the intervention being proposed; a theory of change and logframe to describe the intervention and its indicators; and a discussion about the need for subsequent evaluation and an evaluation plan for follow-up and completion. The strength of evidence available to support the planned intervention becomes a key criterion for deciding on subsequent evaluation. An intervention that is well proven in other settings would need less investment in evaluation than an innovative, unproven design.

The **World Bank** does not have such a strong emphasis on results as **DFID** but does require an abbreviated version of the logframe and arrangements for monitoring and evaluation (M&E) in the project document.

Danida's guidance includes a specific requirement for intervention logic, both for projects designed by **Danida** and for projects planned by partners.

In all cases, including Norway, the approach to evaluability is managed by organisational units responsible for processing projects or grants, not by the organisation's evaluation department.

3.2.3 Quality assurance (QA)

All three comparator agencies require more detailed and consistent Quality Assurance of project design and results frameworks prior to financing than Norway does. **DFID** has a multi-stranded approach to QA, varying the intensity according to size of the investment. All projects are required to have a quality review, which is conducted by a specialist unit for all large projects and self-administered against a checklist for smaller projects.

The QA checklists include detailed information about results and evaluability, including questions about: theory of change which explores and provides evidence for the critical assumptions and linkages from input to impact; analysis of behavioural change; an evaluation plan; a logframe with a clear results chain; and the quality of indicators, baselines, milestones and data sources.

Arrangements for QA are in a process of change at the **World Bank**. Quality at Entry reviews used to be carried out on samples of projects by an independent task team drawn mostly from Bank staff, but the approach was abandoned in 2008 in favour of arrangements managed separately in each region of operations. Since then, analysis of the Bank's portfolio has revealed a decline in

ratings of project outcomes, which has been linked to a number of problems including quality of design and readiness for implementation.

New QA arrangements had not firmed up at the time of this review, but are expected to involve a review of all proposed projects and include optional Quality Enhancement Reviews during preparation. At both DFID and the World Bank, Quality Assurance is a key feature that is used to assess how well projects meet design standards.

Danida has provision for QA through its Technical Advisory Service that appraises major programme support proposals before submission to the granting authorities. A separate Quality Assurance Department follows up during implementation and provides feedback on indicators, results and Danida's management. This unit also assesses the Completion Report.

Norad's quality assurance department, AMOR, has a role to help develop staff capacity and conducts Grant Management Reviews, but these examine compliance with processes rather than the quality of grant project design, and takes place only in Norad and at the Embassies, not the MFA.

3.2.4 Follow-up and monitoring

Both DFID and the World Bank have structured monitoring reports that use rating systems to assess performance. Norway's approach to monitoring grants is more *ad hoc*. DFID has a formal annual review (AR) and project completion review (PCR). The AR scores projects against actual achievement of expected results alongside an assessment of the outcome. All ARs and PCRs are published, and the format has been designed with a view to providing relevant information clearly to the general public. The process is mandatory for all projects approved since January 2011, and follows on from a former system in which large projects of longer duration were rated for performance in delivery of outcomes.

The **World Bank** has a long-established six-monthly progress report, now called the implementation status and results (ISR) report. This includes ratings for progress towards development objectives (DO) and implementation progress. The DO rating has to be supported by reference to indicators from the project logframe.

Day-to-day monitoring for **Danida** is done by the implementing partner. The project Steering Committee is responsible for overseeing that activities lead to the expected outputs and outcomes. The partner needs to be capable of providing sufficient information and able to use SMART indicators and established baselines.

Norway's approach to follow-up and monitoring is more *ad hoc*. Whether progress reports are required and their frequency is determined by the Grant Scheme Rules. Templates for progress reporting exist, but their use is optional. There is no requirement to score performance either by the partner or the programme officer managing the grant.

3.2.5 Completion

Unlike Norway, all three comparators use mandatory standardised templates for completion reporting and require a rating of result achievement. DFID's PCR uses the same scoring as the AR but includes a rating for the achievement of outcome as well as outputs. All projects must have a PCR.

The **World Bank** requires an implementation completion report (ICR) for all projects with ratings of performance. All ICRs are then subject to validation by the Independent Evaluation Group to cross-check the performance ratings. A sample of projects is then subjected to ex post evaluation. **Danida** requires a PCR based on the implementing partner's final report, with an assessment and rating of the results of the investments at output, outcome – and if possible – impact level.

In the **Norwegian** system, all grants have to produce a final report. A proforma S21/S81 provides a structure for the information that is required, which includes some provision for reporting against indicators for outcome and outputs. Use of the form is not mandatory and it can be followed as a checklist. There is no requirement to rate performance. There is provision for Norway's grants to conduct an end review or evaluation, but this is not mandatory and depends on the grant scheme and the risk assessment in the decision document.

Major findings about the comparative approaches by DFID, the World Bank and Danida

The findings from these comparators need to be considered against the context of their own organisational structures, governance and aid modalities.

The **World Bank** is an international financing institution that lends money on a large scale against project or policy-based objectives and has obligations to report on results to its board.

DFID is a bilateral donor like Norway, but with a much larger grant portfolio and a strong institutional separation from the UK Foreign and Commonwealth Office. DFID, too, has to meet searching enquiries from the UK Parliament about the effectiveness and efficiency of aid spending.

Although institutionally **Danida** is probably closest to Norad and Norway's MFA, there are still operational differences, especially with regard to project planning and preparation.

The relevance of findings for Norway is arguably more in the nature of the conceptual approach, and types of systems, than in the detail of rules, procedures and policies. Both DFID and the World Bank impose strict procedures on their borrowers or grant recipients and these can have significant transaction costs. Norway has long followed a partner-led approach and nothing in this review should be interpreted as arguing against that way of working. So the question is, what findings and lessons are relevant for Norway?

Grant application and planning

- Norway's requirements on results measurement at the planning stage give less guidance than those of other agencies and, as such, are less suited to supporting evaluability. Comparator agencies require a description of the intervention logic, a clear articulation of the monitoring systems, and an evaluation plan.
- DFID is the only comparator agency that requires, at the planning stage, a review of the evidence base (and its quality) to support an intervention. This innovative approach provides a logical foundation for deciding the type and detail of evaluation to be planned, and helps ensure a more strategic use of evaluations.

Quality assurance (QA)

- Each of the three comparator agencies requires more thorough QA of individual grants at or before appraisal than Norway. This increases the attention given to results before a grant is approved by DFID, the World Bank and Danida.

Follow-up and monitoring

- Norway requires basic results reporting, but the frequency and content of progress reports and reviews vary by type of grant and are at the discretion of the programme officer. The World Bank and DFID, in particular, take a more consistent approach and require significantly more detailed progress reports that are analysed and reported to top management.

Completion

- Unlike Norway, all three comparators use mandatory standardised templates for completion reporting and require a rating of result achievement.

Overview and lessons

The main implication arising from these comparisons is the need for a coherent system, in which questions about objectives, evaluability and evaluation planning are built in to all stages of the grant or project cycle. To achieve that, staff need adequate technical guidance and systems of checks and balances – periodic reviews and QA, and to know that measurement of results is a high priority for senior management in the MFA and Norad and will influence decision-making. That could include refusing to approve a grant with inadequate arrangements for measuring results.

Experience from both the World Bank and DFID suggests that results measurement is a difficult challenge and systems need to be periodically reviewed and their effectiveness questioned. A results orientation is a management strategy more than a set of technical tools, which is why strong and decisive leadership about the need to plan for and demonstrate results is so important.

There is some evidence that excessive layers of review and Quality Assurance can have negative effects, with staff more concerned to avoid risk to their personal reputation than to take risks with project design. Striking the right balance between flexible guidance and rigid prescription is desirable, especially so that programme officers can work *with* partners rather than just impose demands on them.

3.3 Policies, systems and procedures for grant management

The following chapter presents our findings in relation to Hypothesis 1. It explores the extent to which **grant management policies, systems and procedures provide appropriate and comprehensive guidance to staff to ensure results are measured and grants are evaluable** (Hypothesis 1). This is the first of three chapters that look at the workings of the grant management system across the MFA, Norad and the embassies.

The evidence for this chapter draws on five main sources: 1) a review of all core grant documents; 2) the results of an online survey of staff; 3) four focus group discussions with staff from the MFA and Norad; 4) the results of our comparative analysis of grant management processes of other development agencies; and 5) an assessment of the quality of 20 grant-level mid-term and end reviews and evaluations.

A number of grant management rules, policies and guidelines have been identified that relate to results measurement. The overarching policy framework for how results are measured across the Norwegian government, including the MFA, Norad and embassies, is set by a number of key Norwegian laws and regulations:

- The Administration Act (*Forvaltningsloven*);
- The Budgetary Regulations (*Bevilgningsreglementet*); and
- Regulations on Financial Management in Central Government. Provisions on Financial Management in Central Government (*Reglement for økonomistyring i staten. Bestemmelser om økonomistyring i staten*).

The Regulations on Financial Management, in particular, provide the basis for all Norwegian grant management. These regulations state that grant managers must ensure that grant reports should contain information on results and whether Norwegian money is spent effectively.

The main MFA- and Norad-specific documents are:

- **Grant Management Manual: Management of Grants by the Ministry of Foreign Affairs and Norad.** This outlines a common approach to grant management across the MFA and Norad. It details the requirements for each of the four grant scheme regimes¹⁹ and provides details of key activities at each phase of the grant management cycle (preparation, follow-up and completion). Guide V04 'Results and risk management: key concepts and methodology' provides specific details on what needs to be considered when assessing results and risks. The Grant Management Manual was updated in 2013 partly to improve the focus on results measurement during planning and implementation.
- **Results Measurement in Norwegian Development Cooperation: A Practitioner Guide.** This provides practical guidance on measuring results. It clarifies terminology; explains the role of indicators and results chains; introduces the concept of an M&E plan; and provides a list of 'issues to consider' at each of the phases of the grant management cycle. It is an introductory guide to results measurement.
- **Grant Scheme Rules.** There are 45 different Grant Scheme Rules and each one provides background information on the fund and the criteria that will be used to make funding decisions. Depending on the scheme, details may also be provided on the results framework for the scheme, whether grants need to be quality assured and the evaluation provisions that are in place.

Alongside these three documents are two others that relate to results measurement. The **Logical Framework Approach** is a practitioner guide to using logframes to plan, monitor and evaluate projects. Although this guide is no longer in circulation, it is still publicised on the Norad website and therefore has been included in our study. The **Development Cooperation Manual** has a focus on results measurement. However, with the introduction of the Grant Management Manual this has been superseded. Interviews with staff indicated that some still use this document.

At the state level, there are two additional guides to note: **Measuring Results: Results Management in the Central Government**²⁰ and **Evaluating Central Government Grants**.²¹ Neither are referenced in core Norad or MFA documents, which raises questions about how widely they are actually used. The MFA Grant Management Unit, however, has informed us that they were used to inform the development of the Grant Scheme Rules and are referenced in certain training material.

¹⁹ The four grant scheme regimes are: project and programme support; general grants; small-scale grants; and direct disbursement of grant funds.

²⁰ Resultatmåling, Mål- og resultatstyring i staten.

²¹ Evaluering av statlige tilskuddsordninger.

The number of Grant Scheme Rules and the variation in requirements around goal achievement, Quality Assurance and evaluation present a confusing set of procedures for staff to follow. The Grant Scheme Rules are a core reference point on results measurement with Norad. There are 45 in total, 35 for MFA, 10 for Norad. Although the Grant Management Manual outlines the basic requirements on results measurement, such as the existence of indicators and baselines, the details of whether results frameworks need to be quality assured, whether evaluations will be conducted and so on, are contained within the Grant Scheme Rules.

Our review of the 45 Grant Scheme Rules found significant variation. Of the 10 rules for Norad, 4 have a general statement on the need to document results,²² 1 has no requirement at all, and 6 describe the results chain for the Grant Scheme and require grants to have objectives structured in the same way. Of the 35 MFA scheme rules, 9 are with no goal statement, 16 have a specific statement, and 10 have a general statement (see summary in Chapter 10 of Annex 4).

Similarly, for QA, across the schemes there is a mix of requirements that ranges from technical assistance, such as expert appraisal being optional, not required, mandatory under specific conditions (e.g. depending on the size and complexity of the grant) or, in a very limited number of cases, mandatory. Again, the provisions associated with whether grants are evaluated are similarly variable. In some cases, the Grant Scheme Rules outline a clear plan for evaluation, but in others there is only a general statement of intent (see Table 2).

Table 2: Results requirements across the Grant Scheme Rules

Results issues	Criteria	MFA	Norad
Quality assurance (QA)	No requirement	26	10 ²³
	Optional ²⁴	4	0
	Mandatory under certain conditions (only legal)	5	0
Technical assistance	No requirement	22	4
	Optional	9	5
	Mandatory under certain conditions	4	1
Evaluation of scheme	No specific evaluation provisions	30	9
	Clear evaluation plan in place	5	1
Criteria for goal achievement	No approach described	9	1
	General statement on approach to measuring goal achievement	10	4
	Specific approach to measuring goal achievement	16	6

22 Two frequent statements used across Grant Scheme Rules are: 1) 'The unit responsible for the scheme is in charge of assessing performance to the extent possible, by putting together results information from grantees regularly. The relevance of the project to the objective of the grant scheme must be documented. Clear objectives and targets, and indicators for achieving these must be drawn up at project level. The total results achieved for the target group should give an indication of the social effect.'; 2) The Norwegian funds make up a small proportion of the total funding provided for this objective, and the social effect can also be influenced to a considerable extent by other factors. This makes it difficult to measure results at the grant scheme level. The aim is therefore to examine general progress within the area, and how the Norwegian funding may have contributed to this.

23 In four cases, Norad relies on the QA systems of the partner. These have been classified as 'no requirement'.

24 In those grant schemes where QA is optional, the guidance is that QA should be sought if the project is large and/or complex.

The explanation offered for this diversity of requirement on results measurement relates to the Provisions on Financial Management in Central Government – one of key regulations setting the overarching results measurement across the Norwegian government (see above). This requires the management of grants to be adapted to each individual grant by an assessment of risk and essentiality. This is to secure cost-efficient grant management. This principle is built into the Grant Scheme Rules and, as a consequence, few have clear cut requirements on when appraisals, reviews and/or evaluations should be carried out.

The problem with this approach, however, is that it means all grants are treated on a case-by-case basis. This presents a confusing mix of requirements for the individual programme officer to manage. The staff survey indicated that 54 percent of respondents agreed that they are often unclear around which rules and guidelines on results should be followed when managing grants. Evidence from the focus groups, interviews and the qualitative responses to the survey supported this finding. Box 3 presents an illustration of some of the views of Norad and MFA staff involved in grant management.

Box 3: Comments from staff on the Grant Scheme Rules

“Grant Scheme Rules just focus on the objectives and process. They are not helpful [to help measure results]. And there is a confusing mix of them.”

“Grant Scheme Rules and other internal grant rules need to be clarified and made simpler. They seem to be overlapping and are confusing.”

“The challenge with Grant Scheme Rules is that they are all slightly different. In the case of humanitarian emergencies for example, it’s difficult when you shift from emergency relief to long-term development as you move between grant scheme rules that have different requirements.”

The minimum requirements on results measurement that are outlined in the Grant Management Manual are not adequate to ensure evaluability. In particular, the level of information requested of partners during the planning of grants is not sufficient. An important part of fostering consistency in how results are measured is to define minimum standards. The Grant Management Manual outlines a number of requirements that all grants must meet. These are defined for each of the three phases of the grant management cycle (see Table 3).

Table 3: Minimum requirements for results measurement outlined in the Grant Management Manual

Phase of the grant management cycle	Minimum requirements
Preparatory	- Grant applications should be assessed based on their relevance to the grant scheme and the realism and feasibility of the goal hierarchy. - The goal hierarchy needs to include impact, outcome and output statements. It should also include indicators for outcomes and outputs and baseline data. - If the programme officer believes that the objectives are unrealistic or indicators are unsuitable, the applicant must be requested to amend the application.²⁵
Follow-up	- Progress reports should include an assessment of the preliminary results that have been achieved. The focus should be on outputs with an indication of the likely outcomes. - In reviewing a progress report, a programme officer should compare the results with the original goal hierarchy and verify that the agreed indicators and baseline values have been used.
Completion	Final reports should present an assessment of outputs and outcomes as well as an estimation of probable impact.

The stage of the process where we see the most room for improvement is at the preparatory phase. During the application process, partners are asked to provide a basic level of detail on their proposed intervention. This includes: a description of the project, including the need it is addressing and the rationale; the implementation plan; details of the budget and the goal hierarchy; indicators and baseline data. These requirements are detailed in the Grant Management Manual. While these are all important to ensuring results can be measured, they are not sufficient for ensuring the evaluability of grants. Based on the comparison with peer agencies and the approaches of others, we have identified a number of additional issues central to evaluability:

- Articulation of a **theory of change** – While there is no single definition of what theories of change are, a recent DFID study²⁶ into their use in international development highlights a number of key characteristics: a detailed description of the **context** for the initiative, including social, political and environmental conditions; the current state of the problem the intervention is seeking to influence and other actors able to influence change; an articulation of the **long-term change** that the initiative seeks to support and who it ultimately benefits; the process/sequence of change anticipated to

²⁵ Grant Management Manual, p.87

²⁶ Vogel, I. (2012) Review of the use of theory of change in international development: Review report, April, DFID, UK. http://r4d.dfid.gov.uk/pdf/outputs/mis_spc/DFID_ToC_Review_VogelV7.pdf

lead to the desired long-term outcome; and **assumptions** about how these changes might happen, as a check on whether the activities and outputs are appropriate for influencing change in the desired direction in this context.

While grounded in similar thinking to logic models and goal hierarchy, theories of change push thinking further and require a much deeper analysis and articulation of how an intervention aims to bring about change, the assumptions that are being made and the influence of the wider context. Goal hierarchy, as defined in the Grant Management Manual for example, only requires a simple articulation of expected cause-effect relationships over time.²⁷

In recent years, the theory of change has been adopted by some development agencies as a more flexible approach to describe an intervention. Theories of change are particularly useful when working in complex areas of work such as governance or civil society strengthening.

- **Reviewing the evidence to support an intervention design** – An evidence review is the systematic analysis of the evidence that supports an intervention's design. It provides an indication of whether an intervention is well grounded in established evidence, or if it is innovative in nature. A recent addition to this practice is not only to review the evidence, but also to review its strength; this requires making a judgement around the credibility of the evidence. Both practices are important to evaluability because they indicate whether the intervention is building on an established body of credible evidence of what works, or is using an innovative design that has yet to be tested. This judgement should inform the approach to M&E. If the evidence base for the intervention is strong there may not be a need for an evaluation or review, or if one is commissioned it might focus on process or partnership issues. When the evidence base is weak or absent, and the intervention is more innovative, the M&E budget may need to be higher. DFID, as part of its Business Case process, requires any applications for funds (internal or external) to include a review of the evidence that supports the proposed design of an initiative and an assessment of its strength using a 'How to Note' on assessing the strength of evidence' (Annex 5).
- Putting in place a clear **results measurement system** – Evaluation is difficult in the absence of an effective monitoring system. Routine monitoring of data is the foundation on which evaluations are based. This requires a well thought-through results measurement system that considers: what data are being collected; how they are triangulated; who is collecting the data, how and when; the resources that are available to do this; when and how are evaluations to be commissioned, etc.

27 Ibid, p. 83

Table 4: Information central to evaluability that is required as part of the planning of grants

Issues key to ensuring evaluability	Grant Management Manual
Goal hierarchy	Required
Articulation of theory of change, including assumptions	Not required
Available evidence that underpins programme design	Not required
Quality of available evidence that underpins programme design	Not required
Indicators and baseline data	Required
Systems of results measurement including: data collection schedule, budgets, roles and responsibilities, commissioning of evaluations/ reviews	Not required

Our assessment of the Grant Management Manual against these criteria reveals a number of gaps (See Table 4). First, while the application process requires applicants to outline their activities, outputs, outcomes and impact in the form of a **goal hierarchy**, they do not have to articulate a **theory of change**. There is no requirement to unpack the underlying logic of their proposed project explaining how activities link to outputs, to outcomes and eventual impact, the assumptions that underpin these relationships or the influence of other factors and actors outside of the intervention to the desired change.

Second, while applicants are asked to outline **indicators** and **baseline data**, they are not required to describe the overall **systems that are in place for results measurement**. At no stage in the preparatory phase are grant applicants required to provide details on issues such as: when data are to be collected; the tools that will be used to collect data; how data will be triangulated; the resources that will be required; roles and responsibilities for data collection and analysis; or whether evaluations will be commissioned and for what purpose. Similarly, in the case of large strategic institutional grants there are no requirements for organisations to detail their organisation-wide approach to measuring and reporting results.

These are crucial details for programme officers to feel confident that a partner has given adequate thought to how results will be measured and has the systems to be able to generate robust and credible results data. The DFID Business Case, for example, requires projects managers and those applying for funds to detail how the project will be monitored and the proposed evaluation plan. It also requires details of the project logframe and indicators (Annex 5). Interestingly, the Grant Management Manual states that programme officers as part of their assessment of the ‘realism of the planned results’ should assess the systems for results achievements.²⁸ However, there is no further explanation of what this actually means in practice.

²⁸ Grant Management Manual, p.86.

Third, while applicants are required to describe the rationale for their proposed project, they are not asked to present the **available evidence base that underpins their intervention design or assess its strength.**

While there is reference to results measurement across a number of policies and guidelines, together the documents fail to provide a coherent body of guidance material that supports staff in practically appraising results frameworks and supporting partners in measuring results. As part of our review of grant management policies and procedures we assessed the documents that have a clear focus on measuring results to see how practical their guidance was. We made this assessment across five key areas of results measurement. The assessment used a four-point scoring scale:

- High – the document indicates that the issue is important and provides detailed guidance on how to put it into practice, including how to appraise it.²⁹
- Medium – indication that the issue is important, with basic guidance on how to put it into practice.
- Low – indication that the issue is important, but no guidance on how to put it into practice.
- None – no mention of the issue.

As Table 5 demonstrates, what we found is that while documents highlight the importance of having indicators, baselines and so forth, there is limited depth to the support that is provided in translating these issues into practice. What is particularly surprising is that there are no mandatory standard templates or checklists for Norad advisers or programme officers to use in quality assuring results frameworks and/or results reporting. This is a contrast with the practices in comparator agencies, where quality assessment procedures make use of standardised checklists.

²⁹ It is important that Norad/MFA guidelines not only give details of putting results measurement into practice, but also offer support on how issues such as results frameworks and progress reports can be appraised and a judgement made on their quality.

Table 5: Level of practical guidance offered by core documents on results measurement

Issues key to ensuring evaluability	Grant Management Manual	Results Management in Development Cooperation	The Logical Framework Approach	Evaluating Central Government Grants ³⁰	Measuring Results: Results Management in the Central Government ³¹
Goal hierarchy	Low	Medium	High	High	Medium
Theory of change and assumptions	Low	Low	Low	None	Low
Availability and quality of evidence that underpins programme design	None	Low	None	Low	Low
Quality of indicators and baseline data collection plans	Low	Medium	Medium	Low	Medium
Systems of results measurement including: data collection schedule, budgets, roles and responsibilities	Low	Low	Low	Medium	None
Management and design of mid-term reviews, final reviews and evaluations	Low	Low	Low	Low	Medium

In contrast to our assessment, the survey results indicate that staff generally find the guidance material on results measurement to be useful (Table 6 summarises staff views). It is important to note, however, that ‘utility’ does not equate to ‘adequacy.’ While staff may find the guidance useful, our argument is that the guidance material fails to cover all of the necessary issues.

While it may be useful, we do not think it is comprehensive enough to support the effective appraisal of partners’ results frameworks or evaluability. In addition to this, the survey results also suggest scope for improvement, particularly around reviewing baseline data, indicators and data collection plans, and M&E budgets.

³⁰ Evaluering av statlige tilskuddsordninger.

³¹ Resultatmåling, Mål- og resultatstyring i staten.

Table 6: Staff perceptions of the utility of policies and procedures in review grant applications in key areas of results measurement

Issues key to ensuring evaluability	% (number of respondents)				n=
	Very useful	Useful	Somewhat useful	Not useful	
Assessing baseline data*	20% (19)	35% (33)	39% (37)	5% (5)	94
Reviewing how measurable and achievable project objectives are at different levels (outputs – outcomes – impact)*	19% (18)	45% (42)	29% (27)	6% (6)	93
Reviewing the logic of the project design	17% (16)	52% (48)	27% (25)	4% (4)	93
Assessing the quality of indicators and data collection plans*	14% (13)	42% (39)	37% (34)	8% (7)	93
Assessing the appropriateness of the budget for monitoring and evaluation	10% (9)	40% (36)	36% (33)	14% (13)	91
Reviewing management and design of mid-term reviews, final reviews and evaluations*	16% (15)	50% (46)	29% (27)	4% (4)	92

Note: Rows marked * do not sum to 100 % due to rounding of numbers.

The minimum requirements on results measurement, and their practical application, are not consistently understood by staff. While the Grant Management Manual outlines a number of requirements on results measurements (see Table 3 above) the understanding of these among staff is mixed. On the one hand, the staff survey indicated that 83 percent felt they had a clear understanding of the minimum requirements that a partner's results framework should have.³² On the other, the interviews and focus groups revealed that staff felt very unsure of what the Norad/MFA approach to results measurement was, and felt that individual staff members were given too much individual discretion to decide what is good enough when appraising results frameworks. Box 4 provides an illustration of some of the views we heard.

This contradiction in the evidence could be explained in two ways. First, the minimum requirements on results are not being effectively communicated. The most recent version of the Grant Management Manual was only released in 2013, so perhaps staff have still not familiarised themselves with its contents. For example, we heard from staff that some still prefer to refer to older versions of the manual because it is what they are most familiar with. The confusion around minimum requirements could therefore simply be a reflection of the manual being in transition.

³² We recognise this is just one part of the minimum requirements; however, we have used it as a proxy for wider awareness of the minimum requirements.

Another explanation is that, while staff may be aware of what is the minimum that needs to be included in a results framework and reporting, they are not clear on what quality looks like. This is analogous to the finding above about the absence of more practical guidelines: while the minimum standards define what needs to be present in a results framework or in reporting, they do not help staff in understanding what is good enough in terms of quality. For example, staff may be aware that indicators and baselines need to be present, but they are unsure about what constitutes a good indicator or baseline, and what should be considered as substandard.

Box 4: Comments on the minimum requirements on results measurement

“The guidance we have available to us is very vague. It’s optional and it exists in too many places. Our approach to grant management is very vague.”

“We don’t have clear criteria for deciding what’s appropriate on results. It’s very much left up to the individual to decide.”

“We don’t know what the bare minimum is on how results are measured.”

“We don’t have a clear understanding of what results measurement is in Norad; everyone has a different understanding.”

“Too much discretion is given to the individual member of staff, the programme officer, on what is acceptable results measurement.”

“Guidance on results is there but there are no clear criteria on what matters, what is good enough or what is not acceptable.”

“It is not clear in the system what the red lines [are]. Staff commenting on results frameworks should make a judgement on if it is good enough. We need to indicate if something needs small changes or if it needs fundamental changes and money should not be disbursed before the issues are addressed.”

There is limited guidance available to staff on how to design, manage and/or quality assure grant level reviews and evaluations. Evaluation and reviews are described in the Grant Management Manual as a key tool for assessing results during both the follow-up and completion phases of grants. They provide an opportunity to go beyond the self-reporting of grantees and provide an independent judgement on the effectiveness of an intervention.

As was discussed above, while Grant Scheme Rules make reference to evaluations and reviews, the decision to commission one mostly lies with the programme officer managing the grant. This same person has responsibility for quality assurance. Despite the importance of reviews and evaluations the guidance available to staff on how to ensure high quality designs and implementation is poor. This is particularly problematic given the finding, discussed later in the report in Chapter 3.5.1 that, based on our review of 20 grant-level reviews and evaluations, few met basic quality standards.

The Grant Management Manual contains no guidance other than to say reviews and evaluations are important. Results Management in Development Cooperation describes the OECD-DAC evaluation criteria and provides a list of indicative questions for all criteria. The most comprehensive and practical

document is Evaluating Central Government Grants (which deals with grant schemes rather than projects or programmes). It provides a good introduction to the principles of evaluation and includes checklists for staff to use. However, this does not seem to be a widely circulated document within Norad or the MFA. It is not cross-referenced in any of the core policies or guides and was never raised in any of our interviews or focus group discussions with staff. There is possibly scope to make better use of this document.³³

Other comparator agencies have a much more comprehensive approach to providing guidance in evaluation and reviews. DFID for example has developed an Evaluation Handbook for staff that explains different evaluation designs, and practical considerations at each stage of the evaluation process (planning, implementing and completion). It also has a standard template that has to be used to assess quality at entry and at exit of all evaluations and reviews. In support of procedures, DFID also produces brief 'How to Notes' that give more detailed information about a range of topics that staff can use to develop their own understanding and are referred to in checklists.

In contrast to our relatively critical assessment of existing guidance, staff have an overall positive view. The staff survey indicated that 66 percent (n=92) of respondents found the existing guidance on reviews/evaluations useful. While it is important to take note of this finding, it is also necessary to recognise the limits of what it tells us. While staff may appreciate what exists, our review has indicated a clear absence of material on how to design, manage and quality assure grant level reviews or evaluations.

Its absence is of particular concern, given that there is no overall oversight of grant level evaluations and reviews quality in the Norwegian system. Quality is managed and overseen by the individual programme officer. Moreover, the interviews and focus groups indicated that staff are not always happy with the quality of the evaluations and reviews, which suggests the guidance is perhaps not helping programme officers manage evaluation or review processes effectively. See Box 5 for the views of two staff on this issue. This finding is supported by our own review of 20 end reviews, which indicates high variability in quality (see Chapter 3.5)

Box 5: Comments on the quality of evaluations

"We could use evaluation so much better. The quality [of evaluations] isn't great. Evaluators are often fearful of making a clear judgement [on effectiveness]."

"The reviews I have commissioned have not been very good. They have been too afraid to make a judgement. It is always recommended that a project be continued. If it is good it should continue. If it is bad it should continue so that it can improve."

33 The Development Cooperation Manual explains the difference between reviews and evaluations and provides a ToR template. For more detailed information on evaluation it refers the reader to the Sida (Swedish International Development Cooperation Agency) evaluation guidelines. However, as mentioned previously, this document is being phased out.

The flexibility provided to partners in how they present results frameworks and report on results creates inconsistency in the quality of how results are measured. The Grant Management Manual restates policy that due to Norway's 'partner orientation' there are no specific requirements on how applications are submitted. There are no mandatory application templates. All that is required is a letter requesting assistance accompanied by a programme document that can be used as a basis for assessing the application.³⁴

While interviews indicate that this approach to requesting funds and reporting on results is valued by partners, it poses a challenge to Norad and the MFA in maintaining consistent quality in how results are measured and reported. This point was raised in a number of the focus groups and emphasised in the recently published OECD-DAC Peer Review report (2013).³⁵ As one focus group participant remarked:

"Partners can sometimes get away with not having good results frameworks because we do not have any specific results-based management requirements [within Norad/MFA]."

The absence of mandatory application and reporting templates is also likely to contribute to the confusion around what the minimum requirements are on results measurements. For example, while proforma templates exist for applications, progress reports and final reports (templates S01/S51, S11/S61, and S21/S81, respectively) their use is optional. Programme officers decide whether and how they are used. Applicants can submit in their own internal format and the programme officers use the template as a checklist. This flexibility undermines the cross-referencing between templates and guidance that can serve to reinforce messages, and in turn clouds the clarity of communication on what are minimum standards.

An additional consequence is that the Grant Management Manual fails to draw a distinction between a final report (which is mandatory) being an example of self-evaluation by a grant recipient, and a review, which is an independent study commissioned by a programme officer. Self-evaluation has a potential capacity-building role among grant recipients, but this characteristic is neither explored nor developed in the guidance material.

Major findings about the policies, systems and procedures for grant management

Minimum requirement on results measurement

- The minimum requirements on results measurement outlined in the Grant Management Manual are not adequate to ensure evaluability. In particular, the level of information required of partners during the planning of grants on

³⁴ Grant Management Manual, p.22.

³⁵ The report stated: "Despite Norad's quality assurance and advisory roles, it is not mandatory for programme units and embassies to consult Norad about results frameworks agreed with their partners, or to use a common template provided by the agency, and the resulting variance in quality has been highlighted by Norad as a challenge," p.80.

how results will be measured and evaluation used is not sufficient. Key gaps include:

- Partners are not required to unpack in detail the underlying logic of their proposed project, the assumptions that underpin the cause-effect relationships, and the role of factors and actors outside of the intervention in influencing change.
 - Partners are not required to describe the overall systems that are in place for results measurement, including their planned use of evaluations.
 - Partners are not required to present the available evidence base that underpins their intervention design or assess its strength.
- The minimum requirements on results measurement outlined in the Grant Management Manual are not consistently understood by staff.
 - The number of Grant Scheme Rules and the variation in requirements around goal achievement, Quality Assurance and evaluation present an inconsistent set of procedures for staff to follow. This creates unnecessary confusion around results measurement.

Standardised templates

- Norway does not require the use of mandatory reporting templates by partners. This produces high levels of variability in how partners present results frameworks and report on results. This creates inconsistency in quality and undermines compliance with the minimum requirements on results measurement.

Guidance and checklists

- While results measurement is referenced in numerous policies and guidelines, there is not a coherent body of guidance material that supports staff in the practical task of appraising results frameworks, supporting partners in measuring results and ensuring the evaluability of grants.
- There is an absence of guidelines and checklists that support staff in making judgements on whether partners' results information or results frameworks are of 'good enough' quality.
- There is limited guidance available to staff on how to design and manage grant-level reviews or evaluations. This is likely to be a contributing factor to the low quality of reviews and evaluations discussed later in this report.

3.4 Staff training and technical support

The following chapter presents findings on whether **staff have access to appropriate training and technical advice to effectively ensure evaluability and results documentation as part of the grant management process** (Hypothesis 2). Under technical advice, we consider the systems that are in place for quality assurance and expert review. While these could have been covered under the previous hypothesis given they are in a sense part of the procedures for providing guidance to staff on results measurement, we feel they were more accurately described as mechanisms for technical support and therefore considered under this hypothesis.

The evidence for this chapter draws on a number of sources: 1) a review of training material and post-course evaluations; 2) attendance records of training; 3) the results of the online staff survey; 4) findings from the focus group discussions; and 5) the comparative analysis of grant management processes of other development agencies.

The chapter is divided in two parts: first the findings from our review of training are discussed; this is then followed by the findings related to technical support.

3.4.1 Findings on the training available to staff on results measurement

A wide number of training courses are available to Norad, MFA and embassy staff that cover results measurement. These are provided through the Foreign Service Institute (UKS)³⁶ and are part of its annual training programme. The extent to which results measurement is a primary focus of these courses varies. There are three courses where measuring results in grant management is a primary focus of the course, and three that are broader in scope, but which include sessions on results measurement (see Table 7). In addition to the UKS structured training programme, AMOR also delivers *ad hoc* training. This is delivered in response to specific requests or where there is a specific need identified (e.g. through a grant management review).

36 Utenrikstjenestens Kompetansesenter (UKS).

Table 7: Training courses available to staff that include a focus on results measurement

Title of training course	Course lead	Duration of course	Mandatory/ optional course	Frequency of course	Results-focus
Results and risk management ³⁷	AMOR	3 hours	Optional	4 times per year	In-depth
Results and risk management (specialisation) ³⁸	AMOR	2 day	Optional	2 times per year	In-depth
Reviews and evaluations in grant management ³⁹	EVAL	1 day	Optional	Once per year	In-depth
Objectives, performance and risk management for managers in the foreign service ⁴⁰	Grant Management Unit	3 hours	Optional	2 times per year	Basic
Grant management and PTA – basic course ⁴¹	Grant Management Unit	3 days	Mandatory	7 times per year	Basic
Development policy and grant management for locally employed embassy staff ⁴²	Grant Management Unit	5 days	Optional	Once per year	Basic

The content of the training courses on results measurement provides staff with a basic level of knowledge on how to appraise results frameworks. However, there are notable gaps in content around developing theories of change, reviewing the quality of evidence and assessing partners' results measurement systems. As part of our assessment of the training available to staff we assessed the material from each of the training courses to see whether it covered the concepts and themes central to results measurement and provided guidance on how to translate these concepts into practice in the management of grants (see Table 8, below). We found that the depth and breadth of content varied considerably from course to course, but on the whole, the current training programme provided participants with a good basic knowledge of results measurement and how to appraise results frameworks. That said, we identified three notable gaps in course content:

37 Mål, resultat- og risikostyring i tilskuddsforvaltning.

38 Mål, resultat- og risikostyring i tilskuddsforvaltning, fordypning.

39 Gjennomganger og evalueringer i tilskuddsforvaltningen.

40 Mål, resultat- og risikostyring for ledere i utenriktjenesten.

41 Tilskuddsforvaltning og PTA – grunnkurs.

42 Utviklingspolitikk og tilskuddsforvaltning for lokalt ansatte.

First, only the training course on Reviews and Evaluations in Grant Management engages with the issue how to review the available evidence to support a programme design, and importantly, assess its quality and reliability. As discussed above, this is a key step in supporting evaluability as it reveals where a programme design is not grounded in an existing evidence of what works and therefore may require a greater investment in M&E. Given its importance, we would argue that it should be covered in more of the courses.

Second, none of the training courses cover how to appraise a partner's results management system. Again this is key to evaluability. Has the partner budgeted appropriately? Have they got a clear delineation of roles and responsibilities internally for data collection and analysis? Have they thought through how they plan to use reviews and/or evaluations to support learning, etc.?

Third, only the Review and Evaluation in Grant Management course covers theory of change as an approach to unpacking programme design and the assumptions that underpin it. Again, this is an important component of evaluability that staff managing grants should be skilled in.

A possible limitation in the evidence base for this finding is that we were unable to attend any of the training courses directly, so were unable to observe the training being delivered. We are fully aware that good content does not automatically translate into good delivery and good delivery can fill gaps in formal content. That said, we still feel confident in this finding. As is detailed below, the online survey revealed relatively high levels of satisfaction with the training. Using participant satisfaction as a proxy for quality of delivery, we think it is reasonable to argue that the training on results measurement is being delivered to a good quality standard.

The short duration of the training courses raises questions about how effective they are in building staff capacity. The duration of results training available to staff varies. The advanced courses such as Results and Risk Management (Specialisation) and Review and Evaluation in Grant Management last for two days and one day, respectively,⁴³ while the introductory course on results lasts for only three hours. The trainings that have a broader focus may only spend a couple of hours – if that – looking at results measurement.

⁴³ In fact, only one day is spent on results measurement as the other is focused on risk management. Moreover, in the case of Review and Evaluation in Grant Management, while the course is billed as a one-day course, a review of the agenda indicated that it is closer to half a day (starting at 9am and finishing at 2.30pm).

Table 8: Key results concepts and themes covered in training courses

Title of training course	Goal hierarchy	Theory of change	Reviewing /assessing the quality of evidence	Indicators and baselines	Appraising results measurement systems*	Planning for reviews/ evaluations
Results and Risk Management	Yes	No	No	Yes	No	No
Results and Risk Management (Specialisation)	Yes	No	No	Yes	No	No
Review and Evaluation in Grant Management	Yes	Yes	Yes	Yes	No	Yes
Objectives, Performance and Risk Management for Managers in the Foreign Service	No	No	No	No	No	No
Grant Management and PTA – basic course	Yes	No	No	Yes	No	No
Development Policy and Grant Management for Locally Employed Embassy Staff	Yes	No	No	Yes	No	No

*This includes how to appraise data collection and analysis plans, M&E budgets and staff resourcing, etc.

In the case of the advanced courses, it is difficult to see how the allotted time allows participants to go into detail around any of the issues covered, or to provide adequate time for them to share and reflect on their experience and debate issues with peers (all key components of adult learning). Similarly, the three-hour introductory course can only allow for little more than a presentation of results rather than training on results. Based on the evidence, it is fair to argue that the advance courses provide a good introduction to results measurement, while the introductory courses are more of an ‘awareness’ about results measurement.

Training is, of course, not the only way for staff to develop their skills. Staff can develop their skills on the job or gain experience from working in other organisations. However, training is an important means of building a consistent skill base among staff. And while not all staff need to have the same level of competencies, a range of courses of sufficient length should be offered to cover different needs. For example, given the support that Norad advisers provide on results measurement, the absence of a significantly more detailed and advanced course on appraising results frameworks in the current training programme is an obvious gap. DFID, for example, offers training and a 'How to Note' on the Business Case and a five-day intermediate-level course on evaluation (Annex 5).

Staff satisfaction with the results measurement training is generally high.

Post-course assessments for each of the three main results courses are between 4.4 and 5 out of a maximum of 6 for 2012 (figures from 2013 are not available). Similarly, the staff survey indicated that for each course more than 70 percent of attendees felt that it was effective at building their skills in results measurement (see Table 9). Moreover, 62 percent thought that the training programme as a whole was relevant to the realities of managing grants.

The only criticism of the trainings was that the courses are often too academic and do not link sufficiently to the practical realities on the ground. This criticism is likely to be linked with the length of the courses. Trainers have to provide participants with a sound technical foundation in results measurement, but with limited time available, they do not also address the implementation of these technical concepts in practice.

Table 9: Survey responses to the question: 'How effective were training courses at building your skills in results measurement?'

Title of training course	% (number of respondents)				n=
	Very effective	Effective	Slightly effective	Ineffective	
Results and Risk Management	33% (15)	51% (23)	13% (6)	2% (1)	45
Results and Risk Management (Specialisation)	50% (18)	36% (13)	11% (4)	3% (1)	36
Review and Evaluation in Grant Management	36% (8)	45% (10)	14% (3)	5% (1)	22

Note: 'Results and Risk Management' does not sum up to 100% due to rounding of numbers

Attendance levels at the training courses that are focused on results are low. The available evidence suggests a low number of staff attending the three results-focused training courses. The staff survey indicated that: 56 percent of respondents have attended Results and Risk Management (n=81); 46 percent of respondents have attended the more advanced course on Results and Risk Management (Specialisation)' (n=78); and only 28 percent of respondents have attended Reviews and Evaluations in Grant Management (n=79). This same pattern in attendance is repeated in UKS attendance records (see Table 10).

Table 10: UKS attendance records for results-focused training courses for 2012 and 2013 (to date)

Title of training course	Participant numbers ⁴⁴	
	2012	2013 ⁴⁵
Results and Risk Management	54	28
Results and Risk Management (specialisation)	20	13
Reviews and Evaluations in Grant Management	10	18

One possible explanation for the low attendance levels could be that staff do not prioritise training on results. This point was raised consistently in the focus groups (see Box 6). This explanation, however, runs counter to survey evidence that indicates that 74 percent of respondents believe that building their skills in how to plan for and measure results is a high priority for them. However, in the case of the MFA, the results were far less conclusive, with close to 50 percent of respondents seeing it as a priority and 50 percent of respondents saying they did not. This may be a result of the different career incentives and organisational culture within MFA (see Section 3.3). Another possible barrier to attendance could be that managers - the directors/deputy directors who define the training needs of employees – are not prioritising the issue. Again, this emerged in the focus group discussion as a possible explanation.

There is no easy solution to this issue. On the one hand, the results management courses could be made mandatory for anyone involved in grant management. However a number of staff involved in training emphasised that even then people still do not attend.⁴⁶ Ultimately, attendance is closely linked with incentives and culture. If training on results measurement is prioritised by managers, and staff see building their skills in this area as beneficial to their career, attendance is likely to increase. Our findings on the issues of incentives and culture are discussed in the next chapter.

⁴⁴ Attendance figures were requested from UKS from 2008-11. However, we were told that a different software program was used to record participant information before 2012 and this data could not be accessed. Also the training courses had other names than today therefore it was not possible to compare figures.

⁴⁵ These are attendance figures up to October 2013.

⁴⁶ The MFA grant management unit, for example, has found that attendance levels for Grant Management and PTA – Basic Course are lower than what is expected, despite this being a mandatory course for anyone involved in grant management.

Box 6: Comments on the barriers to training attendance

Training is not prioritised by staff

“Training is available to staff, but people don’t use it.”

“Building skills in this area [results measurement] is not high on the list of priorities for MFA staff.”

“It doesn’t help your career to go on a course of Results-Based Management.”

“Building your skills in grant management is not going to help your career in any way. It’s not valued.”

Training is not valued by managers

“Grant management training is not prioritised by our leaders compared with other more pressing and visible issues.”

“Heads of departments grumble when people have to take time off to go to training.”

“The top of the Ministry do not expect people to have the types of skills necessary for managing grants and measuring results.”

While training is an important means of building staff capacity on results, there is an absence of practical guidance on how to put results measurement into practice. Effective capacity building requires a number of complementary approaches. An important finding to emerge from the evidence is the need to supplement training with other types of support for staff on measuring results in practice. This point was consistently raised both in the online survey and during the focus group discussions (see Box 7). The types of support that are staff are requesting are three-fold:

First, **tools that help bridge from the theory of results measurement to applied practice.** There is currently a dearth of practical guidance on results. As is argued in Chapter 3.1, there are a range of documents stating the importance of results measurement, but very few provide support on how to put it into practice.

Second, **information and support on how to approach results measurement in specific sectors.** The challenges of results measurement vary from sector to sector. For example, specific methodologies have been developed for measuring results in the field of governance and these may be different to those used in health. Similarly, standard indicators and associated data collection methodologies exist in different sectors that represent the accepted way of measuring certain outcomes. Developing a set of sector-specific guides, along with banks of indicators and assessment methodologies, could provide a useful resource to staff as they appraise results frameworks and work with partners in specific sectors to strengthen their approaches to results measurements.

Third, **examples of what good practice looks like in results measurement**. Norad staff in their role as technical advisers, sometimes struggle to not only comment on results frameworks, but also provide support in improvement; a bank of good practice examples for such issues as theories of change, M&E plans, data collection tools for specific indicators, etc. would provide a useful resource for this and support them in making judgements around quality.

Box 7: Comments on the need for practical guidance on results measurement

“There is training on results, but there is a gap between the theoretic[al] approach and the reality. There isn’t a link between training and our day-to-day reality. More practical tools that bridge this gap would be very useful.”

“We have good planning and financial tools; we now need better tools to work on results and evaluation.”

“If I work in climate and environment and want to see what good indicators look like [for this specific sector], I don’t currently have anywhere to go”.

“Results frameworks will always differ between types of organisations and types of projects, and individual assessments will need to be made based on this. A few good examples of great frameworks would be helpful.”

“Pooling results information in one spot would be useful. We sometime ask AMOR for examples of best practice, but they don’t have anything they can share with us.”

3.4.2 Findings on the technical support available to staff on results measurement

The technical support available to staff on results measurement can take one of two forms: advice and guidance from either a Norad adviser or a results specialist from AMOR’s Results Management Section. Grant Scheme Rules dictate whether technical advice and/or quality assurance is necessary (although none require advice specifically on results).

Grant Management Reviews take place annually at the request of the MFA. They are part of the MFA’s internal control system and not part of the regular grant-/project-processing procedure and a Quality Assurance procedure as such. The aim is to ensure that grant management is in accordance with existing rules and regulations, not only for results measurement, but for grant management in general. The reviews have a dual purpose of control and learning.

There is a mixed picture on the use of technical advice from Norad on results measurement. The evidence presents a mixed picture on whether technical advice on results frameworks and/or reporting is being requested from Norad. Our review of 20 grants for example, found that technical assistance was sought in six cases. This is despite 14 of the grants being more than NOK 15 million, which under the old grant management manual was the threshold for mandatory technical advice for grants under the bilateral aid budget (03.10). Whether technical advice needs to be sought is now dictated by the Grant Scheme Rules. While this low number could be a result of grant managers simply not documenting the advice they received, we think this is unlikely given the clear paper trail that often exists when technical advice is sought.

The staff survey, on the other hand, suggests the use of technical support is more widespread: 75 percent of respondents indicated that they have recently requested support from AMOR on results frameworks during the preparatory phase of the grant management process; while 81 percent of respondents have requested support from Norad; and 65 percent of respondents from an external expert. This uptake is lower in the follow-up phase (see Table 11).

Table 11: Percentage of survey respondents that requested technical support on results measurement during the preparatory and follow-up phases of their most important current project/programme

Types of technical advice available to staff	% that requested support (no. of responses)	n=
Preparatory phase		
Specialists in Norad	82% (72)	88
AMOR	75% (56)	75
External experts	65% (43)	66
Follow-up phase		
Specialists in Norad	59% (47)	80
AMOR	48% (33)	69
External experts	49% (36)	74

The quality of technical support provided by Norad to the MFA and embassies on results measurement is mixed. As discussed above, our review of 20 grants revealed that only six had requested any form of technical support from a Norad adviser, the AMOR Results Management Section or AMOR Legal Section. In the majority of cases, results measurement was not the main issue being addressed. In most cases, comments and advice were around legal issues, traceability of funds, etc. Where results issues were addressed, however, we found that the advice frequently lacked depth and there was often not a clear judgement on quality. Table 12 presents the results of our assessment. The assessment was based on the following four-point scale:

- High –detailed assessment of quality, with a clear judgement
- Medium – good assessment of quality, but no clear judgement
- Low – covers the issue, but with limited assessment of quality
- None – no mention of the issue

Table 12: Assessment of the content of the technical advice provided by Norad to embassies and the MFA on results measurement for the six of the 20 reviewed grants where advice was requested⁴⁷

Grant number and name		Goal hierarchy	Indicator quality	Results measurement system	Evaluation/ review plans
GLO-07/387-2	Integrated Programme for Xingu	Low	Medium	Low	None
LKA-08/075	IFC Enterprise Development	High	Low	Low	None
GLO-09/853	Disaster Risk Reduction and Preparedness	None	Low	None	Low
QZA-09/220	Pilot Program for Climate Resilience	None	Low	Low	Low
MEU-11/0077	EBRD Southern and Eastern Mediterranean	None	None	None	None
08-1084031	UNICEF/Basic Education. Gender	Low	High	Low	None

Table 12 indicates the issue that was covered most consistently and with the most depth across the grants was the quality of indicators (five out of six). The issue that was covered least frequently and with the least depth was the assessment of the evaluation/review plan.

The support that AMOR's Results Management Section provides in reviewing partners' results frameworks and results reporting, while helpful, can be overly formalised. As a result, it can sometimes fall short of providing the practical guidance that programme officers need to support partners in developing more robust results frameworks. The current process for receiving technical advice from AMOR's Results Management Section is for a formal written request to be made outlining the issue. AMOR then provides a written response. There is a clear rationale for this approach. AMOR is expert in results measurement, therefore its services are in

⁴⁷ Across the six grants, technical advice came from a number sources: Norad advisers, the AMOR Results Section and AMOR Legal Section.

demand. In order to manage its workload, processes have to be in place for logging work and dividing it up within the team.

While the capacity of the AMOR Results Management Section has increased in recent years, it does not have enough staff to handle *ad hoc* requests. In addition, there is also the more delicate issue of managing the section's professional reputation, as informal comments on a results framework from AMOR could lead to claims that AMOR has signed off on a grant's results framework, despite not having had the opportunity to review the final product and ensure the comments have been taken on board.

The challenge with this model of technical support is that it can hinder interaction or discussion that allows programme officers to talk through ideas with AMOR and for staff at AMOR to gain sufficient detail of to provide technical input that is contextualised to a particular grant. In the absence of consistent personal interaction, we found that AMOR's advice can be viewed as rather technical and theoretical.

To be clear, we are not suggesting that there is never interaction between AMOR and programme officers; our focus group discussions highlighted cases where this has taken place. Our point is that this interaction and discussion seems to be the exception and the more frequent model is for more formalised interaction. The evidence suggests that this can mean the technical advice provided by AMOR can sometimes fall short of offering programme officers the level of practical guidance they need to effectively support partners in developing robust results frameworks. Box 8 provides an illustration of the types of views we heard about the interaction with AMOR.

Box 8: Comments on technical advice provided by AMOR on results measurement

"AMOR can sometimes set standards too high. There is a gap between what is considered good practice and what's possible with CSOs of limited capacity."

"The comments that AMOR provide are very technical. They are generalists. We need more support around how you can take general technical guidance and make it relevant in a particular context. We don't just want to know what a good indicator is; we want to know what a good indicator is for this particular sector programme, in this particular country."

"AMOR advice tends to be detached from realities. There is a great gap between theory and practice."

"The interaction with AMOR is very formalised. We need more of a team-based approach where we can have informal discussions with colleagues about results. The guidelines say to do an 'integrated assessment', yet AMOR feels like an external agency. At the moment getting informal support from AMOR relies on personal connections. I know that AMOR needs to manage its workload and keep track of how many requests it is getting, but the current system of submitting formal requests for review is very time consuming. We need more interaction with them."

The current model, whereby individual programme officers either quality assure results frameworks themselves or seek technical advice from Norad, leads to inconsistencies in the robustness with which proposed results information is being accepted and reported. The Norwegian system's approach to quality assurance is built on a model of delegation: individual programme officers have primary responsibility for quality assuring partner's results frameworks, but can seek technical advice when they think it is necessary.

While Grant Scheme Rules are supposed to provide direction, in no case did we identify a scheme that mandated external quality assurance of results frameworks;⁴⁸ what form quality assurance takes is at the discretion of the individual programme officer. Similarly when technical advice is sought, again, authority is delegated to individual Norad advisers who, in the absence of any specific quality assurance templates or clear guidelines, approach the quality assurance process in whatever way they see fit.⁴⁹ This model of quality assurance relies heavily on there being a high skill level in results measurement across the system and good levels of standardisation in approach. Available evidence suggests practice in both these areas is very inconsistent.

The 'Provisions on Financial Management in Central Government' require the management of grants to be adapted to each individual grant by an assessment of risk and essentiality. This is to secure cost-efficient grant management. This principle is built into the Grant Scheme Rules and Grant Management Manual, and as a consequence few Grant Scheme Rules have clear cut requirements to when appraisals should be carried out or when reviews and evaluations, etc. are mandatory.

This approach to technical assistance is at odds with many peer agencies, and where Norwegian practice differs the most from comparator organisations. All others have an arrangement where a proportion of grants undergo a formal and systematic QA assessment. Danida, DFID and the World Bank all have had QA systems that are applied to review the design of projects before financing is approved. Systems at the World Bank are currently being revised. All three organisations vary the extent of coverage of QA either according to the size of the project (with all larger projects being reviewed) or through a sampling scheme (Annex 5).

While Norad staff are responsible for providing technical advice on results frameworks and reporting, many are sector specialists and are given no additional support to build their skills in results measurement. In the current Quality Assurance system, Norad staff are responsible for providing technical support on results frameworks. Sometimes this support comes from AMOR, but as is indicated above, in many cases it is provided by other Norad advisers, many of whom are sector specialists. Our survey indicated that 60 percent of Norad staff who responded feel confident in their skills to appraise result frameworks (n=43); however, the interviews and focus groups suggested a more complex picture (see Box 9 for some of the views that were expressed).

48 The five Grant Scheme Rules that do include mandatory provisions for QA require legal QA, not QA of the results framework. While we have seen instances of the legal QA covering results issues, the legal focus dominates these QA processes.

49 Apparently there is a process under way in the MFA to develop a template.

While staff may have the basic competency in results measurement, in the absence of any specialised training or guidance, many felt they did not have the level of expertise and skills needed to provide technical support to others. Staff may feel comfortable to identify when something needs improvement, but do not necessarily feel equipped with the expertise to suggest how it could be improved. This requires a different level of capacity. Based on the available evidence it is not clear whether there is the depth of skills with Norad (other than within AMOR) to enable the organisation to effectively provide technical advice on results measurement and reporting.

Box 9: Comments on the challenges faced by Norad staff in advising on results measurement

“Programme officers have an impossible task. We are technical experts and also supposed to comment on result frameworks.”

“What’s our role in grant management? We tell grantees to show results, but are not really able to help them. We tell them it needs to improve, but we are not sure how to do it.”

“We need to work with partners to understand and articulate objectives. Work with them to craft a results framework/negotiate a results framework. This is a difficult set of skills to master. It requires confidence in your technical abilities.”

Grant management reviews assess compliance with rules on results measurement rather than the quality of how results are being measured and reported. A Quality Assurance process that does take place on a regular basis is the grant management reviews. These reviews were frequently mentioned by staff during interviews. While staff commented on the pressures the reviews created, the feedback that is provided is seen as important to learning and improvement. A rating assessment is made of the work of a unit or embassy as a whole and this was broadly welcomed by staff and seen as a useful intervention.

The challenge with grant management reviews is that they are process focused and checking compliance with procedure. In this respect, their assessment of results is limited to, for example, whether staff have interacted to encourage the collection of baseline data and indicators, rather than the quality of what is produced. Similarly, they look for evidence that a grant manager has commented on a results framework, and not if the results framework and reporting is of a high quality and how it needs to improve. This is not to argue that checking that procedures have been followed is not important. But the evidence points to weaknesses in the quality of how results are measured and reported.

The issue of grant reviews only focusing on process has been a topic of discussion within the department for some time. Revisions were made to the process in June 2013, but their focus continues to be on compliance with rules and regulations.

Major findings about staff training and technical support

Staff training

- The training courses on results measurement are generally of a good quality, with high staff satisfaction. However, there are gaps in content in a number of areas that are central to effective results measurement and evaluability:
 - There is an absence of training on how to appraise a partner's results measurement systems and plans.
 - There is insufficient attention given to reviewing the quality of evidence as part of appraising programme design. This is touched upon in the Reviews and Evaluations in Grant Management course, but only to a limited extent.
 - There is insufficient attention given to how to develop and appraise theories of change. This is touched upon in the Reviews and Evaluations in Grant Management course, but only to a limited extent.
- The short duration of the training courses (between a couple of hours and a day) also raises questions about their effectiveness in building appropriate levels of staff capacity in results measurement.
- Attendance levels for training courses focused on results measurement are also low, and there is a perception among staff that building skills in results measurement neither supports career progression nor is valued by managers.
- Outside of the training there is also little other practical guidance available to staff on how to put results measurement into practice.

Technical support

- There is an inconsistent use of technical advice by programme officers to support the appraisal of results frameworks and reporting. A review of 20 grants indicated that in the majority of cases technical advice was not used.
- The inconsistent use of technical support results from the fact that the decision whether to request technical input rests with the individual programme officer; only in a limited number of cases is it mandated by Grant Scheme Rules.
- This approach to technical support is at odds with the practices in other peer agencies such as DFID and the World Bank. It is where Norwegian practice on results measurement differs the most from comparator organisations.
- A review of 20 grants indicated that when technical advice is provided by Norad to the MFA and embassies its quality is mixed. The advice frequently lacks depth and fails to provide a clear judgement on whether the results framework or report is of 'good enough' quality.

- The technical support provided by AMOR's Results Management Section, while helpful, is sometimes viewed as too formalised and lacking the practical guidance programme officers need to effectively support partners in developing robust results frameworks.
- The Norad staff responsible for providing technical advice on results frameworks and reporting are often sector specialists and are not given additional support to build their skills in results measurement.

3.5 Implementation to ensure evaluability and results measurement

This section moves on from the examination of training and technical advice and draws together evidence to examine how well **the policies, systems and procedures that are in place to ensure interventions are evaluable and results data is being collected are being correctly and adequately implemented** (Hypothesis 3).

The evidence for this chapter draws on a number of sources: 1) our review of 20 grants and follow-up interviews; 2) the results of the online survey of staff; 3) findings from the focus group discussions; and 4) the findings from other independent reports into results measurement in the Norwegian system.

The section is structured in two parts: the first presents our findings on the extent to which policies and procedures on results measurement are being put into practice; the second then explores the factors that influence this process.

3.5.1 Findings on how well procedures on results measurement are being followed

Many grants are found to have poorly developed results frameworks in terms of quality of the objectives and indicators. Across the sample of 20 grants that were reviewed, we found that while a good number provided a clear articulation of the problem that was to be addressed and the specific objectives that were sought (15 grants were scored satisfactory or above on this), there were consistently gaps in detailing the programme logic or theory of change. The most frequent assessment of programme logic and indicators was 'poor' (10 of the 20), only three projects were found to have 'good' indicators and none 'excellent'; only two projects were rated as 'excellent' or 'good' for follow-up reporting.

In only half of the instances had grants clearly outlined how activities, outputs, outcomes and impact were linked and the assumptions underpinning the intervention (this was conducted to a satisfactory level in nine instances). The designs were essentially grounded on a problem analysis that fitted with national and Norwegian policies, but without an explanation about how the activities that were to be funded or the outputs they deliver would stimulate progress towards the project purpose. Coupled with this, the quality of indicators was very mixed (with only half the reviewed projects being assessed as having satisfactory or good indicators) as was the justification of the programme design (only nine

were rated as satisfactory or above). Other recent reports have highlighted similar issues.⁵⁰

Progress reporting and final reports deal mainly with implementation of activities and financial status rather than with results. Our review of 20 grants identified very mixed quality in progress reports, with most being either poor (eight) or satisfactory (nine). Notably there was a consistent absence of outcome-level reporting. While we recognise that the grant management manuals (both former and current) only require the review of progress against outputs, it does state that progress reports should, if possible, also cover 'the project's assumed effect on the target group' (p.62.) Moreover, many reported activities rather than outputs.

A similar pattern was found in the final reports; of those reviewed (eight) only two offered robust reporting at outcome level. This is more problematic as there are specific requirements in the grant management manual to assess a grant's performance at both output and outcome levels. A significant finding from the survey is that the majority of staff who responded (77 percent in the case of progress reports and 73 percent for final reports) are satisfied with the quality of reporting that they receive. This suggests that staff are not placing a high priority on the reporting of outputs and outcomes, despite emphasis on that in the grant management materials.

Programme officers engage with grant applicants, but mostly about general issues of design, delivery and finance rather than results. Across the sample of 20 grants that were reviewed, we found consistent evidence of programme officers engaging with grantees during the preparatory phase and discussing areas for improvement in the grant design and delivery (14 cases). Rarely did these discussions focus on the grant results framework (only six cases). Mostly, they were about issues of financial management and clarifying the focus of the initiative: four raised questions about indicators; three about baseline information; one asked for an improved intervention logic; and one for a results matrix; two said that the risks needed to be explained more; and one called for a review to be planned. None discussed the evidence base for the intervention or the quality of evidence.

A similar pattern was identified during the follow-up phases of grants (evidence of interaction in 10 cases, with eight looking at results). In some cases, serious issues were identified in the results framework, which were not picked up that subsequently made results reporting much more difficult later during implementation. For example, in one instance the design document for an NOK 73 million grant

⁵⁰ Norad (2011), Annual Report 2011: Evaluation of Norwegian Development Cooperation, Norad, Oslo; OAG (Office of the Auditor General) (2011), The Office of the Auditor General's Investigation into Results Orientation in Norwegian Development Cooperation (Norwegian translation), Document 3:4 (2010-2011), 13 January 2011, Oslo; Gjennomgang av resultatfokus og risikohensyn i Norads faglige rådgivning, AMOR/22.01.09 Some examples from recent evaluations commissioned by EVAL: a) The report from the Evaluation of the Strategy for Norway's Culture and Sports Cooperation with Countries in the South points to both a lack of criteria for assuring quality and assessing performance and the absence of data for evaluating success or failure; b) The Evaluation of Results of Development Cooperation through Norwegian NGOs in East Africa reported that 'most projects lacked the data and information required to be able to measure changes in indicators for key results accurately.'; c) The Evaluation of Norwegian Development Cooperation to promote Human Rights could report on the increased number of university-based human rights centres in Indonesia as a result of Norwegian support. Still, the report had to conclude that 'there is no clear indication of what impact this has had on actual human rights compliance,'.

only outlined outputs and the overall objective of the project. No details are provided on outcomes. There is no evidence that this issue was picked up by the programme officer. It is unsurprising that in later reports there is no reporting on outcomes, as there were no outcomes specified in the programme design.

Box 10 provides three examples of grants where results measurement has been well addressed. In two cases they reflect extensive interaction with the grant manager.

Box 10: Examples of good practice in results measurement in grant management

IFC enterprise development facility

LKA-08/075, LKA-3124

This NOK 15 million grant was to the South Asian Enterprise Development Facility, a multi-donor advisory initiative for small and medium-sized enterprise development in the Maldives and Sri Lanka. The grant was to fund Phase II of the project, which aimed to continue and consolidate activities from Phase one. It was managed through the Norwegian embassy in Sri Lanka.

The application for this grant included an explicit logic model and an MTR from Phase I was used to inform and evidence the design of Phase II. The embassy in Sri Lanka used Norad technical support to review the design documents, including the results framework. This led to improvements in the indicators, baseline and greater clarity around key assumptions underpinning the programme. A mid-term review was conducted during the follow-up phase to assess progress, and plans are currently in place for an end evaluation. The embassy plans to seek input on from Norad on the ToR.

Emergency response for internally displaced people (IDPs) in Côte d'Ivoire

CIV-11/0005, CIV-1019

The one-year NOK 10 million grant was provided to the Norwegian Refugee Council for the emergency response in distribution, shelter and protection monitoring in two areas of Côte d'Ivoire in 2011.

The Norwegian Refugee Council followed the MFA's project proposal template and completed a detailed logframe that identified measurable results. The logframe helped to clearly articulate the intervention, to provide a logical intervention structure and to set realistic targets. Indicators, many of them SMART were provided to measure progress at outcome and output level. Outcome indicators included for example the percent of shelters for the most vulnerable households at sufficient standard or percent of targeted IDPs with non-food items to ensure hygiene and cooking activities during displacement. The final report reported against the logframe and the targets and along with the narrative provided a compelling and credible analysis of results achieved.

African Rift Valley

GLO-0630, GLO-08/449

This project supported with NOK 17.9 million of funding originated as a continuation of a grant to support a geographical focus of WWF-Norway's work and was incorporated into support for WWF-Norway's strategy for 2009-11. The purpose was stated as conservation and sustainable use of biodiversity improved and sustainable low carbon development promoted in conjunction with improving people's livelihoods based on more equitable access to, control over and sustainable use of natural resources.

Extensive interaction took place between Norad and the grantee, and was a major influence in the highly structured and detailed results framework for the 2009-11 strategy; good quality progress reports were produced following a results matrix format and a very detailed final evaluation report, which includes an updated logframe reporting results at output and outcome levels and a note about the baseline situation.

Most final reports are not consistently archived and therefore reduce the scope for future evaluation and accountability. In conducting our grant review. We found 12 instances where final reports were not saved in the archives. It is unclear if this was a result of the reports not being delivered, or not being uploaded to the archive. Either way, this points to a weakness in the Norad and MFA systems. The end reports are an important mechanism for internal accountability. The 2009 grant management manual specifically states that final reporting 'may include descriptions of activities during the previous period, descriptions of a project's end products, deviations in relation to plans, *goal achievement, effects on the target groups* and others, sustainability and summary of main findings,'⁵¹ (emphasis added).

The current Grant Management Manual, through the final report proforma template S81 states that 'it is required to report on the project's effect on the target group (outcome) and on the products/services delivered. The project's probable impact on society should also be indicated.' There is no doubt the intention is that grant reporting should include an assessment of progress towards outcomes. Five of the eight grants for which reports were available were assessed as being satisfactory or better, with one rated excellent because it had very good reporting on outcome and output indicators, making it easy to see the level of progress made at the end of the intervention.

Few of the assessed grants make use of mid-term or end reviews or evaluations to learn from experiences. Grant-level reviews and evaluations are independent external studies that assess results. They can be commissioned either during a grant or at the end. Of the 20 grants reviewed, only four commissioned reviews or evaluations as part of the grant management process. During most of the period covered by our review, grant management guidance was through the Development Cooperation Manual and reviews were only mandatory above NOK 50 million.⁵²

When grant-level reviews and evaluations are commissioned, their quality is generally poor. In order to fulfil our ToR and ensure a minimum number of grant-level mid-term and end reviews and evaluations were assessed we sampled a further 20 from a database held by Norad. The 20 were selected randomly from the published lists⁵³ of 'Reviews from Organisations' and 'Norad Collected Reviews' in proportion to the total number in each category. Our sample consisted of 16 drawn from the former list, which despite their titles are all reviews commissioned by the grantees and conducted by an external consultant. In addition, we selected four reports from the latter list, and these are reviews commissioned by embassies or other sections of the aid administration.⁵⁴

51 Grant Management Manual Web 2 p.61

52 Although widely used by staff, the manual was only *mandatory* for projects under the regional allocation, so relevant for only a few of the projects in the sample.

53 Selecting only end reviews, final reports and final evaluations from the lists <http://www.norad.no/en/tools-and-publications/publications/norad-collected-reviews?index=80&serie=collectedreviews> and <http://www.norad.no/en/tools-and-publications/publications/reviews-from-organisations>

54 The end reviews are coded QA1, QA9, QA12, QA14. QA1 and QA14 are called end reviews, while QA9 is called final report and final evaluation, and QA12 is called final report. Two were commissioned by embassies (QA9 and QA 14), one jointly between the embassy and the MFA (QA1) and one by Norad (QA12). The reviews commissioned by the grantees are coded QA2 to QA8, QA10, QA 11, QA13 and QA15-QA20.

Our assessment of the reviews and evaluations highlighted a number of weaknesses:

- None of the reviews or evaluations measured results at the outcome level.
- Over half (11 out of 20) of the reports provided insufficient analysis of results at the output level.
- Nine out of 20 do not even identify the outputs that stems from the intervention (QA 2, 3, 9, 12, 13, 15, 16, 17 and 18).
- Only two out of 20 (QA 5 and 20) establish and test any theory of causal links (theory of change or other logic) between the intervention and its expected results.
- Almost half (nine) of the reviews do not present a clear analysis where the conclusions are based on the evidence and data, and assumptions underlying the analysis are not discussed (QA 2, 3, 4, 9, 10, 13, 15, 17 and 20).
- Almost half (nine) of the reviews are not clear about the process for making judgements and there is no explanation of the extent to which the evidence supports the judgements being made. Moreover, gaps and limitations in the data are not discussed (QA 1, 2, 3, 4, 9, 10, 13, 16 and 20).

In addition, we found that the resources available for undertaking the reviews and evaluations are often quite limited. The number of days for the consultancies vary from 13 (QA 2) to 26 (QA 4), with one to two weeks of field visits.⁵⁵

On the positive side we find that 14 of the 20 reviews do collect primary and secondary data in a way that could have been suitable for assessing the results at the output level when taking into account the time and resource constraints. Among those, almost half (six out of 14) describe the data used in a fully satisfactory way and the reader can assess the adequacy of the empirical foundation of the review. Moreover, most of the studies (15 of them) triangulate the data sufficiently for informing about results and half of the reviews get a fully satisfactory quality on this criterion. The typical approach is to triangulate findings from document reviews with those from stakeholder interviews.

⁵⁵ Many reviews do not specify the number of days for the assignment.

Box 11: Example of good practice for short reviews

The final report 'Livelihood Sustainability through Raising Community Capacity for Fisheries/Coastal Management', Kenya was commissioned by WWF and conducted by Naturecom Group during a period of 16 days and can serve as a good practice for reviews with limited time and budget (see full assessment in QA 5). The report outlines the evaluation methodology in detail and the techniques for data analysis are also discussed. The methodology is thoroughly presented (pp.2-3 and Annexes) together with a description of the key questions and the data sources and methods/tools including an elaborate Evaluation Matrix (Annex 4).

The review used the theory of change already established during the project initiation and assessed the achievements of the programme against the logical framework parameters. An assessment is undertaken for the achievements of each output and the underlying assumptions are assessed. The review incorporated additional indicators in the logframe that were not incorporated in the original version (pp.37-42) and made sufficient use of existing data sources, collected own data and provided detailed descriptions of data used. However, the sampling strategy is not explicitly outlined.

The results at the output level were documented and assessed in detail in Annexes 5-7 (pp.35-48) and then summarised in the main text. The findings clearly follow from the data and the conclusions are also derived from the analysis of the findings. The recommendations follow logically from the conclusion (see pp.43-48 and the discussion of the lessons learned pp.20-21). The process of making judgements is clear; however, gaps and limitations in the data are not explained, though the team has cross-checked the reliability of the data.

3.5.2 Factors that influence attention to results in grant management

The role of culture

The findings from the review of 20 grants indicate that good standards of preparation and follow-up for results are not being followed in most instances, but that good examples exist where a high standard is achieved and grants are highly evaluable. To understand why this should be the case, we included questions about the style of management and ways of working in the MFA and Norad. First, we questioned staff perceptions about the culture of the Norwegian approach to working with partners for development cooperation.

Programme officers take time to identify potential partners prior to approving a grant. Once identified, the relationship is managed with a high degree of trust and flexibility. The Norwegian approach is regarded as distinctive. Both the staff of the MFA and Norad, and development partners characterise the Norwegian approach as being 'partner-led', in line with international principles. What makes this unique is that staff put a lot of effort into identifying suitable partners before considering a grant, but having done that, once the grant is approved the relationship follows a 'light touch' with a high degree of trust and flexibility and minimal 'interference'. The fact that Norway only ever considers grant applications from partners, and staff never lead or take

responsibility for design and preparation, distinguishes Norwegian aid from most other major bilateral donors.⁵⁶

This practice is in line with the principle of local ownership and alignment to partners' systems, embedded in the harmonisation agenda outlined in the Paris, Accra and Busan declarations. As Norway is a relatively small donor, programme officers consider it important to take into account any extra administrative burden imposed on partners by expecting them to use Norwegian templates for their projects. The ideology is to build on and support partners own planning systems and monitoring as much as possible. Hence planning and execution is left to the partners.

But this approach is presumably not intended to operate contrary to the quality of the projects. Both the partner country and Norway seemingly want projects to succeed as much as possible and there is a wide international view that planning and managing for results is one way of contributing to that success. It would seem inappropriate for Norway to fund a grant application with poor technical design, whether for health or education or governance. So why should the design of results measurement and evaluability be treated differently?

Rather than being seen as an administrative burden, designing for evaluability is just one aspect of good practice. Dialogue with partners about evaluability can focus on principles and is a form of capacity building that fits well with what is sometimes called 'recipient responsibility' for its own development.

An underlying weakness in results specification might be poor capacity among partners. There is little time to develop that capacity when a new application is being considered. To examine a specific manifestation of this relationship we asked participants in the focus group discussion whether the fact that projects are designed by partners means it is sometimes difficult to ensure results are clearly measured /described: 8 out of 9 people from Norad in the focus group agreed with that, but staff from the MFA were less certain, with 6 agreeing and 5 disagreeing.

The role of leadership

Writers on results-based management argue that leadership is an important facet of success.⁵⁷ It is necessary to set the tone and direction of management and clarify aspirations for the organisation.

56 This oversimplifies the relationship a little. Many programme officers liaise closely with their grant partners, while the partner is preparing an application and the degree of interaction can sometimes be high.

57 See, for example: OECD-DAC (2008) Sourcebook on Emerging Good Practice in Managing for Development Results 3rd Edition p.94 *et seq.*; Mackay, K. (2006). *Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management EDC Working Paper Series - No. 15*. Washington, DC: Independent Evaluation Group, World Bank; Mayne, John (2008) Building an evaluative culture for effective evaluation and results management. Institutional Learning and Change (ILAC) Initiative Brief 20. CGIAR, p.4.

There is evidence of a high-level interest to demonstrate results from Norway's aid programme. This is manifested in public statements from ministers and political leaders over the period reviewed in this evaluation. It is notable that the budget submissions to parliament in both 2012 and 2013 have include a supplementary annex on results from development aid. And the incoming government elected in October 2013 included issues about results from development aid in their 'political platform' statement.

Although the need to demonstrate results is stated by the political leadership, many staff believe it does not translate into a drive for a results-focus in the aid programme by senior management in the MFA and Norad. This topic provoked lively discussion in the focus groups and a contradictory range of responses in the staff survey. In the focus group discussions, 15 out of 19 participants disagreed that senior leadership within the MFA and Norad are key drivers of the results agenda. This was much more pronounced for the MFA where 10 out of 11 in the focus group disagreed, than for Norad where five out of eight in the focus group disagreed.

But in the survey, a majority within Norad and the Embassies thought senior leadership was driving the results agenda while, by a small majority, more staff in the MFA disagreed. Small sample sizes and confidentiality in the responses means that no further analysis could be done to ascertain the distribution of holders of opposing views. A plausible explanation for this inconsistency is that the communication of the wishes of top management varies among the organisational units. Box 12 illustrates the nature of responses in the discussions and interviews.

Box 12: Comments on the results orientation of the leadership

"Leadership on results is reactive: we are told to show results when the leadership needs to justify the aid budget. It's not done proactively."

"Leaders are not driving a results agenda, its 'just talk'. Reports go nowhere, hence there are no incentives. Twice a year a minister and a director [say] in a speech, 'if results are too poor, then drop the partner', but no one ever takes any interest or follows up."

"There is leadership in speeches, but not in practical policy."

"Serious leadership from the ambassador is key – we have one example of an embassy from the grant management reviews that did very well. The reason for this was that the ambassador prioritised grant management and results, and there was an employee at embassy level that had the right competencies and had the time."

"You hear ministers and senior leaders talking about the need to show results, but it is another thing to make it happen. The practical implications of the results agenda are not recognised. The leadership are not willing to invest the time and resources needed."

The role of staff capacity and resources

Staff involved with grant management consider that they possess adequate skills to review and monitor grant performance. Despite the preponderance of evidence that the majority of grants do not have good results frameworks and cannot be evaluated for outcomes, perceptions of staff indicate that they consider themselves to be in possession of the necessary skills to review and monitor grant performance. Table 13 summarises results from six questions in the staff survey.

Table 13: Summary of results from six questions in the staff survey⁵⁸

Survey question	Agree	Disagree	n=
Q 4.7 I am confident in my ability to assess partners' results frameworks to identify areas for improvement(e.g. goal hierarchy, indicators and baseline); (strongest at embassies)	68%	27%	126
Q 4.8 I have a clear sense of the minimum requirements that a partner's results framework should have; (weakest in MFA)	83%	14%	126
Survey question	Familiar	Less familiar	n=
Q 7 How familiar are you with the policies and procedures for examining an application for approval?; (rather lower at MFA)	87%	13%	97
Survey question	Mostly	Somewhat	n=
Q 11 To what extent do you feel you have the necessary skills and knowledge to make an informed assessment on a grantee's results framework?	75%	25%	95
Q 17 Do you think you have the necessary skills and knowledge to make an informed assessment on whether a partner's annual reports/progress reports are of a satisfactory quality?	87%	13%	93
Q 23 Do you feel you have the necessary skills and knowledge to make an informed judgement on whether a partner's final reports are of a satisfactory quality?; (highest Norad, lowest MFA)	84%	16%	89
Survey question	Satisfied	Not satisfied	n=
Q 20 Are you satisfied that your partner's final report provides an acceptable final statement on results?; (highest at embassies, lowest MFA)	73%	27%	89

Note: For Q 4.7: 5% replied 'Don't know / Not applicable'. For Q 4.8: 2% replied 'Don't know / Not applicable', however, the total does not sum to 100 % due to rounding of numbers.

⁵⁸ For simplicity of presentation, we have combined 'Strongly Agree' and 'Agree', similarly with 'Disagree' and the same with the other categories. The separation between 'Agree' and 'Disagree' is sufficiently stark that combining them does not lose any nuance in the answers.

In most instances, more than two thirds and often four out of five staff responding consider themselves in possession of the necessary skills. Results were examined separately for the MFA, Norad and the embassies. There is a trend that responses were less positive from staff in the MFA, but the disaggregated sample sizes are too small for confident analysis. It is notable in Q 20, that staff are content with the quality of their partner's reporting, yet surveys of grants and findings from evaluations predominantly conclude that reporting is of poor quality.

Staff do not put into practice all that they know about good results frameworks, mainly owing to pressures of time and limited resources.

Discussions with staff indicate that they feel pressures of time and constraints to resources limit the extent of interaction that can take place. Overall, more than 70 percent of Norad and MFA staff responding to the survey disagree that they have adequate resources (time and money) to ensure the measurement of results is put into practice (n=76). In comparison, staff at the embassies are more positive, with only 45 percent sharing the views of their headquarters-based colleagues, and 41 percent thinking they do have adequate time and resources (n=49).

Box 13 reproduces comments from the staff survey and interviews. Results from the interviews, focus groups and staff survey indicate that time is prioritised during the initial stage of grant application to approval. During follow-up, 69 percent of staff responding from the embassies (n=42) thought they had adequate or plenty of time, whereas a majority of staff responding from both the MFA and Norad (58 percent, n=50) thought they had minimal or no time to discuss with partners how they are measuring results and how this can be improved during implementation.

By the stage of final reporting, 54 percent of embassy staff who responded still think they have enough time (n=41), but 66 percent of the MFA and Norad respondents (n=47) say they have only minimal or no time to discuss with partners about project results, implications and lessons when the project has been completed.

Box 13: Comments on time available to work on results

“The leadership in both MFA, Norad and at the Embassies should recognise that it is resource demanding and time consuming to ensure good results frameworks. There is a lot of talk about the importance of results, but the focus is mainly on reporting and not on the very foundation for reporting; the preparation phase. When it comes to preparation, there is basically always a lack of time to ensure quality at entry.”

“Implementation is somewhat neglected and unfortunately overshadowed by disbursement levels.”

“Attention to detail is at its best prior to approval (in order to get grant approved) because this is a very visible stage. Attention then falls rapidly afterwards resulting in poor follow-up and lack of attention at completion.”

“Results are expected, but encouraging the investment of time for results dialogue is missing. Some partners can cope, others need capacity development.”

“Lack of resources is another problem; staff are overworked and don’t have the time to properly review applications. Some embassies have very large portfolios...”

“Staff don’t have time. And for many embassies it’s not a priority – their emphasis is on political work.”

“Even more emphasis should be given on results. Much more resources should be allocated to this task if this is to be taken seriously.”

The pressure to disburse funds often trumps thoughtful consideration of results frameworks. Another factor to consider in explaining why policies and procedures are not followed is that programme officers are under pressure to disburse funds. This came up in a number of the interviews and focus group discussions. While staff may want to take the time to develop a robust results framework with partners the drive to disburse funds can undermine that intention. Box 14 summarises the responses to interviews and discussions.

Box 14: Comments on the pressures to disburse funds

“There is a pressure to shift money and spend resources. There are also no incentives or sanctions around measuring results. Sometimes there is also a political need to spend money. Might be a demand from the Prime Minister’s office that country x gets resources.”

“The incentives have been about getting money out; spending rather than sitting down and planning grants where we will be able to measure results achieved.”

“It’s challenging. A programme officer can ask a partner to review their application multiple times, but after a while if there is still not improvement what do you do. Often you fund them anyway. There might be pressure externally to do so. Other donors might be putting pressure on you to disburse the money. Also, sometimes it’s very political. Your partner might be the office of the prime minister. How many times can you go back and forth asking them to review an application?”

The role of the number, size and duration of project and programme grants

There is a trend in recent years for the distribution of project and programme grants to have shifted towards fewer grants, more grants of a larger size and for grants to be given for longer durations. Some staff argue that this change will create more opportunities for staff to spend more time to work on results frameworks. One argument advanced by several staff is that there is time to interact more closely over results with the larger grants rather than smaller grants. This seems intuitively plausible, but we are unable to test that hypothesis from the sample of 20 grants. The argument was taken further, that there has in fact been a trend towards fewer, larger grants given for longer durations and that therefore it is likely there will be an improvement in the results-focus over coming years.

We decided to test that by examining the time series of Grant Regime I project and programme grants⁵⁹ with data kindly supplied by AMOR's Statistics Unit. Tables are in Chapter 11 in Annex 4. The data are for the years 2009-12, and though not entirely consistent give a relatively stable population for the project and programme grants.⁶⁰

There has been a steady reduction in the number of Type 1 (project and programme) grants since 2009, from 1,698 to 1,192 in 2012. Total number and value of all Norwegian grants has fallen over the period and the share of Type 1 grants has risen both in number and value. But the average size of grant has remained constant after a rise between 2009 and 2010.

As far as the distribution of grants is concerned, there has been a small fall in the share of small grants less than NOK 1 million and a doubling of grants larger than NOK 25 million. But the total percentage of these is only 6-7 percent. The total share of grants greater than NOK 10 million has risen from 12 percent in number and 74 percent in value in 2009 to 19 percent in number and 84 percent in value in 2012. Most of the change took place between 2009 and 2010 owing partly to some structural changes.⁶¹

59 Grants are categorised by the MFA under a number of grant management regimes. The Grant Management Manual deals with Regimes I to IV although there are VI plus a number of sub-categories. The focus of this study has been project and programme grants which are handled under grant regime, or 'type' I.

60 The data are collected from the case processing system PTA. This system was not used by all desk officers before 2009. Lower numbers were recorded in 2008 so this year was excluded from our analysis. Not all development aid is included. Aid through Norfund, SIU and the FK Norway is not registered in PTA, and is therefore not included. Some agreements are split into sub-units, with one main agreement and several sub agreements. For these agreements, we only have data for the main agreement. The classification of the types of grants have changed over time. The grant types starting with numbers (I- VI(S)) are based on the new grant types, while the other two are old (Agreement with sub-agreements and Standard Agreement).

61 There are several reasons for the comparatively high levels of commitments in 2010. One of the main reasons is that changes to the case processing system PTA were gradually introduced through 2009 and 2010. This resulted in high commitment levels for these two years with a peak in 2010. Another important reason for the high commitments in 2010 is that an agreement on payment of refugee costs over several years was signed in 2010. This agreement has a total value of approximately NOK 5.7 billion.

The distribution of duration of grants is analysed in a third table and indicates a reduction in the number of grants of less than one year, an increase in the numbers and value of grants given for 2-3 years and an increase in numbers given for more than three years, though a fall in their value. Overall, grants are being given for longer durations. Whether these trends will work through to influence the attention given to results will be interesting to see.

The role of incentives

In view of the findings that staff consider themselves to be adequately trained and competent to manage for results, but subject to competing pressures on financial management and disbursement and a lack of consistent or firm direction from their managers, we asked about the nature of any incentives or sanctions associated with the results agenda.

There are no incentives to reward or sanctions to penalise performance in managing for results. Even if managers have to grapple with competing priorities for staff time, there might still be scope to reinforce attention to results by rewarding staff through simple mechanisms such as praise, or more formally through the annual performance assessment. Equally, sanctions could be imposed where grants do not meet with policy objectives for a focus on results. But the response to the staff survey showed that 70 percent of respondents consider there are no incentives and 63 percent say there are no sanctions. In both cases, staff at the embassies were slightly less firm in their views than staff at headquarters.

This is potentially an important finding. Under the current system of multiple policies, guidelines and rules and inconsistent leadership, staff largely have the autonomy to make their own decisions and in view of the pressure of work, mostly that means accepting what a partner proposes rather than working with them to improve the results framework.

Whilst the surveys do not illuminate a strong distinction between Norad and the MFA, comments in the focus groups and in responses to the survey indicate a clear contrast in results culture between the MFA and Norad. Box 15 highlights the comments received from both staff and some development partners.

Box 15: Comment on the incentives and sanctions around results measurement

“The MFA is very different from Norad – MFA has no standard application form. You have a meeting with them and they give you a million kroner. I’m exaggerating, but you know what I mean...it’s more political funding. For example, it is about making sure Norway has a presence in [name of country]. They are interested in results, but it’s just more political...MFA funding is much more political than Norad funding. We don’t have lots of discussion around indicators, but that doesn’t mean they are not interested in results.”

“There are no consequences for supporting organisation that can’t show results.”

“There are no consequences for not having a results framework....Our partners really struggle putting these together. They don’t understand how to do it and need support.”

“The pace of change with the MFA means that results are not a priority. By the time it comes to measuring results we have already moved onto the next policy issues. Its yesterday’s news....the incentive isn’t there to measure results.”

“Incentives in MFA are around bureaucratic compliance and putting things in the archive. The problem is that you can’t tick a box on impact/results...We don’t look at quality.”

“People see themselves as political staff, not grant managers. But these political staff are going to embassies and are managing very large sums of money.”

“Grant management is viewed as separate from the political and it’s perceived as cumbersome.”

“Political aspects of work in MFA brings prestige, so people would prioritise attending a seminar on politics rather than something on results; there is a perception that the top of the ministry doesn’t expect more than a basic knowledge on results.”

“Interestingly, AMOR staff are the most wanted by other departments. Because of their results skills. This is a carrot to people. Coming to the department is good for their career.”

“[The] MFA recruit[s] diplomats; grant management is not high on the list of things what people want to do. Sometimes people don’t even know they are doing grant management. Leaders in the system need to give these things attention.”

Major findings about implementation to ensure evaluability and results measurement

Implementation of procedures on results in grant management

Of the 20 grants reviewed:

- The majority had poorly developed results frameworks with particular weaknesses in relation to the clarity of objectives and indicators.
- Outcome-level reporting, despite it being a minimum requirement for final reports, was almost completely absent. Most progress and final reports dealt with implementation of activities and financial status rather than with results.
- While there was evidence that programme officers engaged with partners around the grant application, it was mostly about general issues of programme design, delivery and finance rather than evaluability and results measurement.

Use of reviews and evaluations in grant management

- Of the 20 grants reviewed, few made use of mid-term or end reviews or evaluations to measure results and/or to learn from experiences.
- Of the 20 reviews/evaluations assessed, few met basic quality standards. In particular, of the reports reviewed, all failed to provide a robust measurement of change at outcome level.
- The resources made available for reviews and evaluations are typically very limited. This may be a contributing factor to their variable quality.

Factors influencing attention to results

- Although the need to demonstrate results is stated by the political leadership of MFA and Norad, many staff believe it has not translated into a clear focus on measuring results by senior management.
- Limited time and resources present barriers to staff putting into practice good results management, with the pressure to disburse funds often trumping thoughtful consideration of results frameworks.
- There are no incentives to reward or sanctions to penalise performance in managing for results.
- The recent trend towards fewer, larger and longer duration grants may create opportunities for staff to spend more time working with partners on results frameworks and reporting.

4. Results management and evaluability in EVAL commissioned reports

This chapter is focused on the functioning of Norad's Evaluation Department (EVAL). It is divided into two parts: it first (4.1) explores how evaluation processes are managed and whether this supports (or hinders) results measurement; then (4.2) it approaches the issue from the perspective of those conducting the evaluation and explores whether the evaluators EVAL contracts have the necessary competencies to deliver on the evaluations. The overall aim is to come up with recommendations that can improve the work of EVAL in generating high-quality evaluation reports in terms of measuring results and provide suggestions that would be helpful to accomplish its mission in this regard.

4.1 Planning, commission and quality assurance (QA) of evaluations

The objective of this part is to explore whether **the planning, commissioning and Quality Assurance of evaluations places an emphasis on measuring results** (Hypothesis 4). The evidence for this section is drawn from a number of sources: 1) our quality assessment of six EVAL-commissioned evaluation reports;⁶² 2) our assessment of the process associated with the evaluation reports; 3) interviews with the EVAL managers involved in the reviewed evaluations; 4) a review of 26 EVAL ToRs; and 5) our comparative analysis of the evaluation functions from other development agencies.

EVAL is mandated to initiate and perform independent evaluations of activities under the administration of ODA-reportable expenditure in the MFA's budget. It commissions a range of different types of evaluations with a broad coverage of Norwegian development cooperation and does not evaluate individual grants. Both process and results evaluations are conducted. Some results evaluations are focused on documenting outputs, others are about assessing what higher-level results have been achieved with Norwegian funds – typically outcomes. Given that the rationale for this study is to understand the challenges around measuring outcomes, it is primarily the latter type of EVAL reports that we are concerned with.⁶³

62 For a list of the selected reports, see section 2.3 in Annex 6. The titles of the reports are also found in table 16 later in this chapter.

63 We have therefore paid particular attention to the results at the outcome and impact levels, though a complete testing of the hypotheses necessitates the inclusion of results at the output level, as well. This broadening of the hypotheses was explained in the original proposal to the assignment, and also elaborated in the Inception report.

The EVAL Evaluation Report series has a wide focus, which is commendable and in line with EVAL's mandate, but is demanding for EVAL staff. When going through the ToRs of the 26 reports eligible for our sample, we found that EVAL's Evaluation Report series consists of a wide range of studies, both methodologically and thematically, analysing different aid modalities and often combining process evaluations and results evaluations. Going more in-depth, we found from the six selected evaluations that there is indeed a very large span in methods and types of evaluation questions posed.⁶⁴

In our interviews with EVAL staff, all confirmed that they are required to cover a wide spectrum of methodologies and themes.⁶⁵ This diversity is fully in line with EVAL's mandate. Nevertheless, the diversity poses challenges to EVAL staff since they need to handle detailed evaluations across disciplines (health, climate, conflict prevention/rehabilitation, private sector development, etc.) and a wide range of methods. One available support measure is that EVAL uses external experts as advisers to EVAL staff – both thematic and methodological – when deemed necessary.

There are large numbers of evaluation questions in the ToRs and this could lead to reduced quality of the results assessment. Our review of the 26 EVAL ToRs indicates that there is usually a very long list of questions that evaluations are asked to cover.⁶⁶ The number of questions in the selected six evaluations are indicative: There are 29⁶⁷ evaluation questions in each of the ToRs for the more mainstream output evaluations (Evaluation Report 7.10 and 4.12), while the ToR for Evaluation Report 4.11, which is more of a literature review, contains eight⁶⁸ questions. For the ToRs that are outcomes and/or impact focused, there are on average 22 evaluation questions.⁶⁹

The large number of evaluation questions may affect the quality of the evaluation reports in a number of ways. First, the large number of questions can lead to resources being spread too thinly. An evaluation that has too many questions is unlikely to generate in-depth analysis and consequently a robust assessment of results may suffer.

Secondly, it may give room for consultants to direct their efforts towards the questions they are more interested in answering, or which are less demanding of fieldwork. This, in turn, can lead to lower emphasis on higher results since these are usually the more challenging part of the evaluation - especially the task of documenting impact.

64 Recall that this is also a result of the selection process where we purposively selected to cover a wide range of aid modalities, sectors and result levels.

65 EVAL has informally divided responsibilities of thematic areas between staff to try to ensure at least a minimum level of specialisation as often as possible.

66 In fact evaluation teams are often invited to add to the original list of evaluation questions outlined in the ToRs.

67 For 7.10, see Final Report Volume 1 pp.123-124, for 4.12, see Final Report pp.165-166.

68 Final Report, p.138.

69 There are 14 evaluation questions in the terms of reference for evaluation 10.11 (see Final Report, Volume 2, pp.38-39) and 30 in the ToR for evaluation 16.10 (see Final Report pp.141-142), while the very long list of possible combinations of questions made it difficult to decide how many evaluation question are actually posed in terms of reference for 6.10. To exemplify, the first question contains at least 14 sub-questions: 'What have been the results at local, national and regional levels of Norwegian project assistance for the partner country, its business sector, institutions, enterprises, and when applicable to local communities and households?' (6.10 Final Report, p.137).

There is sometimes a large discrepancy between the ToRs and what is delivered in the final report. Moreover, the results from the process assessment show that requirements in the ToRs are sometimes not addressed in the final evaluation report. In those instances where this occurred we found limited evidence of discussion between EVAL and the evaluation team around why certain parts of the ToRs could not be delivered. Most striking is the fact that the three evaluations intended to assess outcomes and/or impacts actually did not do so and were still approved.⁷⁰

One explanation for this might be that consultants during the course of the evaluation find other lines of enquiry to pursue than what was specified in the ToRs (with more or less consultation with EVAL). There is clearly room to negotiate during the process of developing the draft report. In the cases of the three evaluations that were supposed to document outcomes and/or impacts, but for some reason did not do so, we questioned whether EVAL should have taken stronger measures to ensure delivery in accordance with the ToRs.

Interviews with EVAL staff confirm that deviations from ToRs seldom lead to any other implications than a request for additional work and contracts are terminated only in very rare cases. However, the final reports have been approved by EVAL and it was confirmed during the interviews that the quality and the usefulness of the content satisfied the minimum standards of the EVAL evaluation adviser managing the evaluation.

A hands-off approach to managing evaluations can make handling deviations between the ToR/proposal and the final report challenging and can affect the quality of the final report. All EVAL staff interviewed as part of our review of the six evaluations confirmed that there had been limited interaction between EVAL and the consultants between the approval of the inception report and the draft report. This was confirmed by the process assessment for each of the evaluations, although this assessment only relied on archived material. However, recent experience suggests a much closer follow-up of consultants in that period of work.⁷¹

A challenge with a hands-off approach is that it means EVAL only has sight of a report after the data have been collected and the analysis is complete. If results have not been measured effectively it is too late to amend the problem without incurring additional costs. EVAL staff suggested it would be very difficult to get a study back on track if there were large deviations between the inception report and the draft report. The only power they have at the draft report stage is to refuse to accept the draft report and request further work, to deny full payment and/or to terminate the contract. However, termination of contract is only possible if the deliverable deviates considerably from the original specification.

⁷⁰ The process assessment showed that two out of six evaluation questions were ignored without any follow-up from EVAL (Evaluation Report 10.11); collection of primary data was not conducted, despite this being an explicit requirement in the ToR (Evaluation Report 6.10); unintended impacts were ignored, despite EVAL specifically asking for such analysis (Evaluation Report 4.12); and lack of accessibility of main results were not rectified (Evaluation Report 16.10).

⁷¹ This experience is drawn from the EVAL commissioned study of Training for Peace project where CMI and Itad are the consultants, and from the current evaluation.

Also the limited contact with EVAL may be taken advantage of by the consultants. They may submit a proposal that promises a certain deliverable to win the contract and then backtrack at a later stage when it is too late for EVAL to terminate the contract (unless there is a serious breach of contract). This hands-off approach to managing evaluations is not common place. Evaluations undertaken by the Independent Evaluation Group in the World Bank are usually led by a staff member, even if consultants are team members. Practice in DFID is changing, but prior to the creation of ICAI, consultants would be expected to submit progress reports and interact with the task manager before submission of a draft final report often through delivery of intermediate analytical products.

Negotiations between the responsible EVAL officer and the consultant during the evaluation lead to a divergence between the ToR and the deliverables in the final evaluation report. Interviews with EVAL staff indicated that during the evaluation process, there will usually be a negotiation between EVAL and consultants at the inception report stage and at the draft report stage. One view was that the ToR is just the starting point where a wide range of topics and questions are included, and then the process (proposal, contract, inception report and draft report) will and should reveal new knowledge that then could be included in the assignment and be reflected in the final report. In this view, it is useful for EVAL not to take a strict approach in order to get the best evaluation, and then it is also acceptable for there to be substantial changes among the deliverables under the assignment. However, this view was not shared by other EVAL staff and it was confirmed in several interviews that it is not in line with the agreed approach in the department.⁷²

None of the evaluations that were commissioned to assess the achievement of outcomes provided a credible analysis of attribution/contribution. None of the three evaluations that aimed to identify outcomes and/or impacts deliver analytical work that provides information about the results at those levels.⁷³ As documented in their respective Quality Assessment Checklists these reports do not identify a counterfactual and they do not provide a credible attribution of the projects to results at the outcome or impact level. Moreover, the process analysis shows that EVAL did not attempt to rectify the issue at the inception phase,⁷⁴ which would be the only entry point for raising the issue before the design of the study is finalised and field work is undertaken (see the Process Map, Chapter 1 of Annex 4). Similarly, none of ToRs for the three evaluations specified with any level of detail the need to show attribution or contribution through the evaluation.⁷⁵ The inability of any of these reports to measure results and consider attribution of the project suggests the EVAL practice of leaving the design of evaluations to consultants is perhaps not advisable.

72 EVAL even have a checklist that is used to indicate that certain aspects of ToR and end report have been checked – and that they are in accordance with each other.

73 See the Quality Checklists of Evaluation Reports 6.10, 16.10 and 10.11: the Overall Quality Scores, but also the detailed discussions on the criteria for conducting an impact evaluation.

74 In 10.11 the impact evaluation questions are not addressed in the inception report or in the planning of field visits. EVAL did not object to this de-prioritisation.

75 None of the three ToRs describe methodologies that could be applied to assess contribution or attribution at the levels of outcome or impact. Moreover, indicators are not described and data requirements necessary for such analysis are not specified. The consultants' competencies are only vaguely specified and would not ensure a team capable of designing the attribution or contribution analysis for an impact/outcome evaluation, an issue that we document later in this chapter.

Many EVAL reports serve other purposes than documenting outcomes, which is also in line with the EVAL mandate. Almost half of the 26 EVAL reports (42 percent) in the initial sample, as stated in their ToRs, were not aimed at documenting outcomes or impact. In 2011, the year that EVAL's annual report stated that none of its report that year reported results at the outcome or impact level, only two reports out of the eight that were part of our sample and published that year (EVAL Report 1.11 and 10.11) requested an assessment of results at either at the outcome or impact level.

Impact evaluations that seek to quantify attribution require a high degree of specialisation both for EVAL and for consultants. This is required for everything from design (pre-assessment of feasibility, development of ToR, specification of competencies) to completion (assessing the quality of the work). Such impact evaluations are very demanding when it comes to competencies of both the commissioning body and consultants and will usually require a lot of highly specialised resources.

Attribution of impact generally requires more demanding evaluation designs that need to be integrated in the design of the project or programme that is to be assessed.⁷⁶ Impact evaluation designed *ex-post* has a much lower likelihood of successfully delivering credible documentation of the effects of the intervention.⁷⁷ There is a need to identify who will do what when it comes to ensuring impact evaluation in the aid administration as the decision to integrate the evaluation should be made at a design stage of the aid projects and thus involves a joint decision with the implementers or partners.

Evaluations lack clearly planned data collection strategies for using existing monitoring data collected by the project and primary data collected by the evaluation team to answer evaluation questions. When conducting an evaluation it is important to draw on both existing monitoring data collected by the project and primary data collected by the evaluation team. An appropriate combination of data types ensures that resources are used effectively for documenting results.

Only one EVAL report (4.12) out of the five that focused on project results made use of existing M&E data in a way that informed their analysis of results.⁷⁸ Similarly, the same report was the only one among those five evaluations that

⁷⁶ Use of a quasi-experimental design, such as difference in difference requires information to be collected from 'treatment' and 'control' groups before and after an intervention, hence should be planned in advance of a project.

⁷⁷ But causal effects can be tested using generative or comparative frameworks to explain how and why change occurs. These too require careful design and skilled management.

⁷⁸ Report 7.10 does not use project M&E data for result measurement purposes, only interviews, findings from other reports and project documents. Report 10.11 does not use project M&E data so it is uncertain what the inputs of the project are. Report 6.10 uses M&E data from one of the sub-projects, but in an *ad hoc* way and not in a way to document results of the Norad support. However, reviews and mid-term reports are used as evidence base. Report 16.10 collected information through stakeholder interviews, literature surveys, document reviews, reports, and policy documents but there are no indications that existing M&E systems were used. However, programme outputs were used for the evaluation, but not in a systematic way.

collected primary data appropriate for answering the evaluation questions.⁷⁹ In all four remaining cases the documentation of outputs would have benefited from a better use of project M&E data and primary data and in particular to determine attribution.⁸⁰ Moreover, in these four evaluations we found that the lack of application of project M&E data and collection of basic primary data led to a failure to comply with basic requirements for assessing results.⁸¹

None of the three evaluations aiming to analyse outcomes and/or impacts collected, or attempted to collect, data in a way that could be useful to document results at that level. All three evaluations were bound to collect primary data in order to document outcomes/impacts and no suitable secondary data existed for such analysis.

The ToRs do not guide the consultants sufficiently on data collection. In the ToRs for the four evaluations, whether primary data collection is to be conducted is left open for the consultants to decide (Evaluation Report 7.10, p.125),⁸² not mentioned at all (Evaluation Report 16.10), required but not delivered in the final report (Evaluation Report 6.10), or required but inadequate for the identification of the envisaged results documentation (Evaluation Report 10.11, see Quality Checklist Criterion 1 and sub-criteria).⁸³

Logic models, theories of change or results chains are not being used effectively in EVAL evaluations. A good evaluation clearly describes and assesses the logic that underlines an intervention.⁸⁴ It should test the assumptions and causal links between inputs, activities, outputs and outcomes. A well-developed logic model can be a useful tool for designing an evaluation, particularly for identifying key evaluation questions. It can also help identify the performance dimensions most critical to the programme's success, which in turn may also be those that are of most importance for the evaluation of results.

Only one⁸⁵ out of the six reports used or reconstructed the intervention logic to structure the evaluation. In the one instance where it was used EVAL specified its use in the ToR and had to reiterate the request in their response to the

79 See Evaluation Criterion 3 and sub criteria in the Quality Assessment Checklist of each of the four evaluations. Reports 6.10, 7.10 and 10.11 did not collect any own primary data while in 16.10 the methodology section does not include a review of whether data collection is designed appropriately for the evaluation questions to be answered. No sampling strategy is mentioned.

80 The exception is Evaluation Report 4.12 which gets a good score on all criteria for making use of existing project and non-project data and own data collection. They also make a clear distinction between the output that can be attributed to the intervention from that of other contributors. For details of the findings and references, see Evaluation Criterion 1.2 in the Quality Assessment Checklist of each of the six evaluations.

81 In Evaluation Report 16.10, it is not possible to find evidence based outputs, in 10.11 the description of inputs are not covering the interventions, nor are they linked to the activities or outputs described, in 6.10 almost none of the outputs are documented sufficiently, in 7.10 the projects selected for the review, the evaluation does not specify the outputs in any detail necessary to discuss results at the level of outputs. For details and references, see Evaluation Criterion 1.2 in the Quality Assessment Checklist of each of the six evaluations.

82 For the case studies where results at outcome and impact levels were to be documented, the following is stated: 'Data collection: A document review of the Norwegian Assistance to the Western Balkans during the period (1991-2007) is provided together with this document. Further data collection is the responsibility of the evaluation team.'

83 Data collection was in the form of anecdotes in relation to the interventions that were to be evaluated.

84 OECD-DAC Quality Standards for Development Evaluation, p.12. See also the SSØ 2007 "Evaluering av statlige tilskuddsordninger", p.21 and the Donaldson and Lipsky (2006) in The Sage Handbook in Evaluation.

85 This was Evaluation Report 4.12. Two others used the intervention logic in a sufficient way for reporting about results (Evaluation Report 4.11 and 7.10) while the remaining three scored 'poor' on those criteria. For details of the assessment and references, see the Quality Criterion 2 in the Quality Checklist for each of the respective evaluations.

inception report.⁸⁶ While an absence of programme logic may in part be a reflection of their absence at the grant level, their omission from the majority of evaluations suggests an under-appreciation of their value to quality evaluation from both EVAL and the consultants.

Although ToRs for the six reports in our sample include requests for results measurement, several do not provide the detail for directing the consultants towards approaches that could lead to high quality results assessment. All results evaluations should apply a theory of change, logical model, results chain or similar (SSØ 2007, p.21). It is particularly useful for guiding the development of the best research questions for the assignment, which could include development of testable hypotheses, to identify and test assumptions for the intervention and for developing the data collection strategy.

That said, only two of the six ToRs request the inclusion of such models in the assignment with sufficient detail and guidance for the consultants to develop such models (Reports 4.12 and 10.11).⁸⁷ However, the ToR of Report 6.10 requests this in two sentences – ‘the assessment shall identify the results chains for long-term effects,’ (p.137) and ‘The consultant will reconstruct the intervention logic for the main policy instruments used in Norwegian Business assistance in consultations with the stakeholders involved in the policy development.’ (p.139) but it is not specified, explained or followed up.⁸⁸ None of the ToRs provide any requirements or guiding on how the theory should be applied.

There is no requirement in the ToRs for testing the assumptions of the intervention, something that can give important inputs to the results analysis. Deriving testable hypotheses is a more advanced instrument that could be useful. Still, this is never suggested to be included in any of the ToRs, which would be particularly relevant for outcome and impact evaluations. Adding the above findings on the lack of adequate specifications of data requirements in the ToRs, the conclusion is that EVAL has not been sufficiently focused on specifying the main principles that are more likely to lead to a high-quality results assessment in the EVAL commissioned reports.

The overall assessment of the EVAL reports indicates that they are generally of a sufficient quality for informing about outputs but not about causality. Despite the shortcomings documented in the review, five out of the six evaluations contain useful and interesting information about results at the output level and responds to many of the relevant evaluation questions. Only Report 6.10 is rated poor implying that the quality of the results documentation is deficient in important respects; one is rated ‘fully satisfactory quality’ (Report 4.12); while four are rated ‘sufficient for informing about results’ (Reports 16.10, 4.11, 10.11 and 7.10).

⁸⁶ Source: Document reference 1100822-46 ‘Response to comments inception report HRITF’, dated 3 Feb 2012, p.3.

⁸⁷ Report 4.11 is more of a literature review and not directly comparable, and still it seems to incorporate the theory of how corruption works (typologies and ‘the functioning of corruption’, see p.138).

⁸⁸ The ToR does provide a background paper with the main programme theories in an annex, but there is no request for the consultants to use any of it in the evaluation (see p.85 in the report).

The findings from the review indicate that the quality of the five evaluations could have been improved by relatively simple means such as: developing a standard evaluation methodology capable of delivering results documentation; establishing the likely causal links between the intervention and its expected results; making logical use of intervention M&E data in conjunction with collecting primary data; and provide a clear and logical analysis with a clear line of evidence to support the conclusions.

Major findings about planning, commissioning and Quality Assurance of evaluations

Design

- The ToRs for EVAL commissioned evaluations contain a large number of evaluation questions; this could lead to reduced quality of the results assessment.
- The EVAL practice of leaving the design of evaluations to consultants may be a contributing factor to EVAL reports failing to measure results and consider attribution of the project.
- The ToRs for EVAL commissioned evaluations do not guide consultants sufficiently on data collection.
- The ToRs for EVAL reports do not guide consultants sufficiently on how programme theories and models should be applied to inform an evaluation.
- There is no requirement in the ToRs for EVAL commissioned evaluations for testing the assumptions of the intervention, something that can give important inputs to the results analysis.
- All the ToRs for the six reports in our sample request for results measurement, however several do not identify the approaches that could lead to high quality results assessment.

Management

- There is sometimes a large discrepancy between EVAL's ToRs and what is delivered in the final reports.
- Reports were still approved by EVAL despite substantial deviations from the original ToRs.
- Challenges to ensuring quality may be exacerbated when a hands-off approach to managing evaluations is taken.
- EVAL commissioned evaluations lack clearly planned data collection strategies that effectively combine existing monitoring data collected by the project with primary data collected by the evaluation team to answer evaluation questions.

- Logic models, theories of change, or results chains might help plan data collection and support analysis of results but are not being used well in EVAL evaluations.
- The absence of programme logic may in part be a reflection of their absence at grant level but their omission from the majority of evaluations suggests an under-appreciation of their value to quality evaluation from both EVAL and the consultants.

Outcome and impact evaluation

- None of the evaluations that were commissioned to assess the achievement of outcomes provided a credible analysis of attribution/contribution.
- The overall assessment of the EVAL reports indicates that they are generally of a sufficient quality for informing about outputs but not about causality.
- The ToRs for evaluations focused on outcomes and impacts do not have questions, methodologies, data requirements and/or competencies set out explicitly for that purpose.

Purpose

- The EVAL report series has a wide focus, which is commendable and in line with EVAL's mandate, but is demanding for EVAL staff. Many EVAL reports serve other purposes than documenting outcomes, and this is also in line with the EVAL mandate.

4.2 Competencies of evaluators

The objective of this part is to explore whether the **evaluators commissioned by EVAL have adequate competencies to effectively measure results and find/use evidence** (Hypothesis 5). Whereas section 4.1 assessed the extent to which EVAL plans, commissions and quality assures evaluations in such a way as to ensure results are being adequately assessed, this section explores the role that evaluators play in ensuring that results are measured and the extent to which they have the knowledge, skills and experience to enable them to do this effectively.

The analysis of consultants' competences draws on four sources of evidence: 1) quality assessment of the six evaluation reports and the process leading up to the approved reports; 2) the specification of consultants competencies as described in the ToRs for the evaluations; 3) interviews with the respective EVAL evaluation managers; and 4) an online self-assessment survey of the consultants who participated in the six evaluations. Responses were received from 16 of the evaluation consultants.⁸⁹

⁸⁹ Five out of six evaluation teams answered. In some cases, correspondence was done through the team leaders and the total number of requests was not recorded as administrative team members and assistants were excluded by the team leaders. More details can be found in section 9 of Annex 4.

The main source of evidence for this section is the results of consultants' assessments of their own capacities. This, of course, has clear limitations for the conclusions that can be drawn from the data. That said, the level of openness demonstrated by the respondents is revealing and we believe it is possible to draw out a number of key findings.

A majority of the consultants have substantial experience with evaluation, and have formal qualifications in the discipline. The survey shows that half of the consultants have more than 10 years' experience in evaluation and only two have less than three years. Moreover, 10 out of the 16 consultants have undertaken courses in project/programme or policy evaluation. These include: university-level long-term education as part of a degree (nine consultants); university-level short courses of less than three months (seven consultants); professional development short courses of less than three months within past five years (eight consultants). The findings on consultants' competencies were supported in the interviews with EVAL staff where the evaluators' basic knowledge of evaluation theory and practice was generally considered to be high.

A majority of the consultants indicate that they have a solid foundation in the application of core evaluation approaches and tools. Table 14 below shows that all of the consultants indicated that during the past five years they have undertaken participatory evaluation and most have also done results-based evaluations (15 consultants) and mixed-methods evaluations (14 consultants). Similarly, a high share of consultants have used generic social science tools such as key informant interviews, focus group discussions, case study, etc. Again, this finding suggests a basic level of competency in evaluation skills among the consultants that delivered the studies in our sample.

Table 14: Consultants' self-assessment of their expertise by evaluation approach

Consultants' expertise by evaluation approach	Number of consultants using the approach (past 5 years)	Number of responses
Participatory evaluation	16	16
Results-based evaluation	15	16
Mixed methods	14	16
Utilisation focused	8	12
Theory based	8	14
Realist	1	9
Experimental and quasi experimental	4	12

Table 15: Consultants' self-assessment of their expertise by evaluation methods

Consultants' expertise by evaluation methods	Number of consultants who have used the method (past 5 years)	Number of responses
Generic tools ⁹⁰	15	16
Theory of change	11	15
Value for money/Cost benefit	9	13
Logic models	8	12
Contribution analysis	4	11
Advanced statistical analysis	2	10

Most consultants indicate that they do not have the competencies to conduct more advanced results analysis. Table 15 reveals that few consultants have experience in using more advanced tools for establishing causal inference. Only four out of the 11 consultants who answered that question have used contribution analysis, while even fewer, two out of 10, have experience with multivariate statistical analysis. Similarly, around half of the consultants claim familiarity with the use of theory-based evaluations and only every third consultant says the same for experimental and quasi-experimental evaluations. Two-thirds of all consultants also failed to identify the correct description of a counterfactual.

It is not possible to conclude from this evidence that the consultants who conducted the three evaluations aiming to document impacts included in our sample did not have sufficient competencies, as there may have been team members that had the necessary skills but that did not complete the survey. What the evidence does point towards, however, is a general skills gap among the consultants who responded to the survey.

EVAL ToRs do not adequately describe the competencies that are needed for teams to deliver more rigorous impact evaluations. For the evaluation reports that focus on attributing results at the outcome and impact level we found that the ToRs consistently failed to specify the team competencies that would be needed to deliver on the contract. In the ToR for Evaluation Report 10.11, for example, there is only a requirement for 'statistical analysis' expertise in the team, when in fact what was needed were skills in econometrics (see p.41 in the Final Report).

⁹⁰ Document review and analysis, Individual key informant interviews, Selection of indicators, Design of data collection for baseline, Use of the logical framework, Interviews with small groups, Focus group discussions, Participatory methods, Qualitative methods, Case study, Ratings for assessments, Sampling for surveys.

Likewise, in Evaluation Report 6.10 there is an explicit requirement to have at least one member with competence in ‘impact assessment methods’, but no details on which methods or what level of experience (see pp.141-2 in the Final Report). Interviews with EVAL staff supports the finding that the required expertise for evaluating results and particularly outcomes and impacts could have been better articulated in the ToRs for the three evaluations. The lack of specification around the evaluation skills that are required of teams cuts across the ToRs for the six evaluations in our sample.⁹¹

Table 16: Specification of consultants' expertise in Norad evaluation ToRs

Evaluation title	Evaluation competency specified in ToR
6.10 – Evaluation of Norwegian Business Related Assistance: Uganda Case study	No specific evaluation expertise is required to apply.
7.10 – Evaluation of Norwegian Development Cooperation with Western Balkans	Relevant evaluation expertise is required but not further specified.
16.10 – Real Time Evaluation of Norway's International Climate and Forest Initiative. Country report: Indonesia	No specific evaluation expertise is required to apply, but it is stated that a consortium will compile a team with the necessary competencies.
4.11 – Contextual Choices in Fighting Corruption: Lessons Learned	No specific evaluation expertise is required to apply.
10.11 – Evaluation of Norwegian Health Sector Support to Botswana	The competencies for the team leader are very broad and open to interpretation. For team member, no specific evaluation expertise is required to apply.
4.12 – Evaluation of the health results innovation trust fund	The competencies for the team leader are very broad and open to interpretation. For team member no specific evaluation expertise is required to apply.

Major findings about competencies of evaluators

- A majority of the consultants have substantial experience with evaluation, and have formal qualifications in the discipline.
- A majority of the consultants indicate that they have a solid foundation in the application of core evaluation approaches and tools.
- Most consultants indicate that they do not have the competencies to conduct more advanced results analysis.
- EVAL ToRs do not adequately describe the competencies that are needed for teams to deliver more rigorous impact evaluations.

⁹¹ This includes three evaluations with outcome/impact focus and three others. We also did a review of the ToR in the 10 most recent EVAL reports published, and the same pattern emerges: Almost no specification of the evaluation competency required for the assignment.

SECTION C:

Conclusions and recommendations

5. Conclusions

This chapter draws conclusions from the evaluation, building on the findings presented in each chapter. The structure of the analysis was based around the five hypotheses introduced in Chapter 1. Four of the conclusions lead to a simple rejection or acceptance of the hypotheses and these are set out in boxes in the text.

Hypothesis 1: Internal policies, systems and procedures to ensure evaluability and results documentation in the grant management process provide appropriate and comprehensive guidance.

The current versions of manuals and scheme rules identify many of the right issues to help ensure that results are measured, but they are not as effective as they could be. They lack a number of key requirements central to ensuring evaluability and the depth of treatment of key issues is insufficient to support staff in turning theory into practice. In addition, too many requirements related to results measurement vary across grant schemes, are optional, left to the discretion of programme officers and poorly supported with technical guidelines. This leads to inconsistencies in the system around how results are measured.

- **Gaps in the minimum standards required from grant applicants about how they plan to measure results makes it difficult for programme officers to make an informed judgement on the quality of their approach and to ensure evaluability.** The basic level of detail required from grant applicants about their approach to results measurement is not sufficient to make a solid assessment of how well a partner has thought through its approach to measuring results. Major omissions include: theory of change, details of the evidence base that supports the intervention and details of its overall results management system, including how they plan to use reviews and/or evaluations. These are all requirements that peer agencies have embedded in their result management systems and sets Norway apart from current good practice. A strengthened focus on results measurement during the preparatory stages of the grant cycle is where there is a need for the most improvement within the Norwegian system.

- **Development of grant scheme rules and a new Grant Management Manual could have improved consistency and coherence around results measurement. But the non-mandatory use of templates means there are no clear standards.** Templates exist for applications, progress reports and final reports, that reflect these requirements, but their use is optional. The Grant Scheme Rules are supposed to provide more specific requirements around results measurement; however, in practice, most are ambiguous and leave the decisions around what to do to the individual programme officer.
- **Staff do not have access to appropriate detailed information to implement the basic guidance given in the Grant Management Manual.** The current guidance fails to provide practical support. There is no coherent body of practical reference available to staff on either how to appraise partners' results frameworks and reporting, or strengthen their capacity in results measurement. In particular, there is no guidance for staff to make clear judgements on whether something is good enough.
- **The absence of a clear approach to and guidance on reviews and evaluations means their use within the grant management cycle is fragmented and their quality is variable.** There is an absence of any clear guidance on when to commission reviews/evaluations and how best to design and manage them. The decision is for the most part at the discretion of the individual desk officer and partner. The fact that only a small number of the 20 grants we reviewed commissioned evaluations or reviews, and of the reviews and evaluations that are commissioned most failed to meet basic quality standards suggests the current approach is not working (see Overall conclusion, below). Reviews and evaluations present an important source of results information for Norway and are central to learning; there needs to be a more informed and strategic approach to their use at the grant level. This should be grounded in a clear understanding of the intervention being funded and the quality of the evidence base that underpins its design. This should in turn shape the level of investment that is needed in evaluation and the type of evaluation that should be commissioned. The DFID experience is instructive in this regard.

Hypothesis 1 is Rejected

Hypothesis 2: Staff receive appropriate training and technical advice/ support to effectively ensure evaluability and results documentation as part of the grant management process.

The current training programme, while providing a basic level of support on how to measure results, is not providing staff with the breadth and depth of skills and knowledge needed to ensure grants are managed and partners supported in such a way as to ensure evaluability. Coupled with this, the current institutional arrangement for technical support and Quality Assurance are not set up in a way that supports a consistent approach to how results are measured. Whether technical advice is sought is frequently at the discretion of individual programme officers, and those providing the technical advice do so in the absence of standardised templates or checklists, or in-depth support in order to build specialist expertise in results measurement.

- **Courses currently on offer are of a good quality, and cover many of the issues needed to appraise partners' results frameworks and staff who attend are satisfied with them. However, the courses fail to cover key areas for ensuring evaluability** (developing theories of change, reviewing the quality of evidence, assessing partners' results measurement systems) **and uptake of the courses is low.** The low uptake is likely to result from staff not viewing skills in results measurement as supporting career progression and managers not valuing skills in results measurement. The short duration of all of the courses is also of concern. It is our view that courses on results measurement of between three hours and one day are of insufficient length to develop staff skills to the necessary level for them to undertake their roles effectively. The absence of other supportive practical guidance on how to put results measurement into practice that sits alongside the training is also a weakness.
- **The institutional arrangements are not set up in a way that ensures consistent quality across the system.** The use of technical advice is only mandatory in a minority of grant schemes and is little used, as is formal Quality Assurance by the AMOR Legal Section. When technical support is provided by Norad advisers our assessment indicated that it can often lack depth and fail to provide a clear judgement on whether results information is 'good enough.' A reason for this could be that, while Norad advisers bring strong sector skills, very few have specialist expertise in results measurement. In the case of AMOR's Results Section, the technical input it provides is helpful, but sometimes too formalised and lacking the practical guidance programme officers need to effectively support partners in developing robust results frameworks. Lastly, the independent QA by AMOR at embassy level deals only with issues of process compliance, not with content, the quality of results documentation or evaluability of grants.

Hypothesis 2 is Rejected

Hypothesis 3: The policies, systems and procedures that are in place (to ensure interventions are evaluable and robust results data is being collected) are being correctly and adequately implemented.

Implementation of the current rules is inconsistent and the standards of results management and evaluability that are reached are low. This arises from a combination of weak specification of minimum standards, a low uptake of technical skills in evaluation throughout the Ministry and Norad, a low perceived emphasis on results by management, and no incentives for stronger results management. The large number of grants to be managed and a culture by which projects are designed by the applicant reduces the time and scope for staff to intervene and strengthen arrangements.

- **The current rules and guidance on grant management are ineffective at providing staff with the means of ensuring that grants have adequate results frameworks for subsequent evaluation.** Based on the review of 20 grants, the majority do not have sound frameworks of results and would be difficult to evaluate using established methods and criteria. Many of the key steps needed to support evaluability in the design phase are not being taken consistently. The large volume of grants committed every year present a considerable burden for management and the absence of a requirement for an M&E plan leaves little basis to decide where efforts should be concentrated.
- **Staff are under pressure and in many instances find themselves unable to devote the time and attention necessary to improve specification of results.** Staff consider themselves to be competent to manage for results, but with few exceptions they do not put that into practice for an assortment of mutually reinforcing reasons: pressure of work that restricts the time available for interaction with applicants; and perceptions of a low priority for results by managers, especially in the MFA, despite high-level leadership over the need for results. There are neither incentives nor sanctions to reinforce good practice on managing for results.
- **The culture in the MFA and Norad is permissive towards a grant applicant's proposal rather than ensuring that applications meet with minimum standards before being approved.** Inconsistent direction and difficulties in securing technical advice, coupled with the long-established system whereby applicants lead the grant design process all combine to result in a low standard of results specification. Concern to follow the post-Paris harmonisation agenda and use partners' own systems appears to have held back constructive dialogue about results measurement.

Hypothesis 3 is Rejected

Hypothesis 4: EVAL puts sufficient emphasis on results measurement in the planning, commissioning and Quality Assurance of evaluations.

- **Evaluations designed by EVAL does not put sufficient emphasis on gathering evidence about results.** EVAL fails to put sufficient emphasis on results documentation during the planning and commissioning phase of the evaluation. In particular, ToRs contain too many evaluation questions; there is ambiguity in the required methodologies; no guidance on data collection requirements; and most of the ToRs do not require the development of a project logic or theory of change. Specification of the required skills among consultants is also poor.
- **EVAL does not have a systematic active management towards ensuring evaluation consultants remain focused on measuring results.** The hands-off approach to managing evaluations has not been conducive to ensuring consistent quality. Limited interaction during the evaluation lead to missed opportunities for correcting deviations between the ToR/proposal and the final report.
- **Evaluations directed towards outcomes and impacts do not have the necessary design features to ensure an outcome/impact assessment is delivered.** The specification of methodologies, data requirements and competencies of evaluators are not in line with the requirements for outcome and impact evaluations.

Hypothesis 4 is Rejected

Hypothesis 5: Evaluators have adequate competencies to effectively measure results and find/use evidence.

- **Necessary competencies are not expressed clearly in the ToRs.**
- **A majority of the consultants have substantial experience with evaluation, and have formal qualifications in the discipline.** Most of the consultants indicate that they have a solid foundation in the application of core evaluation approaches and tools. However, there is a large gap between what the consultants say they have competency to do and their application of this competency in practice. Since consultants may have the skills, but may not be able to apply them for different reasons, we cannot conclude whether the evaluators' competencies influenced the quality of results measurement in the EVAL reports.
- **Few evaluators have adequate competencies to measure results at the outcome and impact levels, but evaluators have been assessed as competent for the ToRs as tendered.** The evaluators lack skills and expertise in the application of more sophisticated evaluation methodologies. Contracts are awarded through a competitive tendering process and competency is a key criterion. All the evaluators involved in the evaluations were judged as competent by EVAL.

Hypothesis 5 is Unproven

Overall conclusion

The conclusions presented above indicate that the hypotheses developed for our ToR correctly identified the main issues. With one exception, the findings have produced a balanced, triangulated and comprehensive body of evidence that leads to firm conclusions. Where it has not been possible to accept or reject a hypothesis, as in the case of competency of consultants, nonetheless the evidence suggests that competency may be an issue, but that it needs to be tackled through the process of improved ToRs and clearer specification of required skills.

The title of this evaluation is the question: 'Can we demonstrate the difference that Norwegian aid makes?' The overarching conclusion from our findings must be that, as a result of a combination of a lack of incentives, poor processes for planning and monitoring grants, and weaknesses in the procedures for evaluations, this cannot be demonstrated. Although there are some elements of good foundations for better results measurement, current arrangements lack the strength of leadership, depth of guidance and coherence of procedures necessary for effective evaluation of Norwegian aid.

6. Recommendations

The extensive findings in this evaluation lead to a potential plethora of recommendations that respond to each of the detailed elements in the analysis. However, we are concerned that proposing extensive changes across many facets of grant management might run counter to the Norwegian approach to development cooperation. Many examples can be cited where good practice as followed by some of the comparator agencies described in Annex 5 would bring improvements to the systems currently in use. But some of those systems reflect a different scale of operations and different values held by those organisations that are not consistent with Norway's approach and principles.

For that reason, we present our recommendations in three parts. First, we present a list of technical changes that we believe would resolve the shortcomings or gaps in current grant policies, guidelines and operations. Secondly, we make recommendations dealing with the work of Norad's evaluation department. Lastly, we consider some of the bigger, more structural changes that would need to be put in place to address the challenges raised in this evaluation. To do this, we present two options for how Norad and the Ministry of Foreign Affairs could proceed, and list recommendations under each.

6.1 Detailed recommendations on grant management systems

Our recommendations on the functioning of the grant management system are grouped under three themes and derive from findings in the report. These recommendations are directed primarily at the Grant Management Unit in the Ministry of Foreign Affairs and Noard's Department for Quality Assurance.⁹²

Minimum requirements on results measurement

- **As part of the planning of grants, partners should be required to outline in greater detail how they plan to measure results.** The details requested of partners during the planning of grants and indicated on the proforma templates should include, as well as goal hierarchy: indicators and baseline; a theory of change; a review of the evidence base that underpins the programme design and an assessment of its quality; details of their

⁹² The table in Annex 3 shows the connection between findings, conclusions and recommendations in this report.

results management system; and their plan for the use of reviews and evaluations. Dialogue about results measurement should be seen as part of support to build capacity in partners' systems rather than a rigid imposition.

Additional guidance and checklists to support evaluability

- **Guidance should be developed for staff on how to put results into practice, specifically how to appraise results frameworks and support partners in developing effective measurement systems.** This guidance should cover:
 - How to develop and appraise theories of change;
 - How to identify and assess the quality of evidence that support a programme design;
 - Planning and managing grant-level evaluations and reviews;
 - How to appraise the quality of results management systems.Examples can be found from the comparator agencies, but a good start would be to update the 2008 publication on Results Management in Development Cooperation.
- **Standard Quality Assurance checklists should be developed for staff to use when appraising results frameworks, progress reports and final reports.** These should require staff to make a clear judgement on quality, possibly using a red-amber-green rating system. Good practice examples could draw on QA checklists developed by DFID for the Business Case, but would need to be adapted to Norway's grant systems. These checklists would in effect provide a clear specification of what the minimum requirements on results measurement detailed in the Grant Management Manual mean in practice and how staff can make a judgement on what is considered good enough. At the planning stage, for example, the QA checklist should cover: the quality of the theory of change and whether there is evidence for the critical assumptions and linkages from input to impact; quality of the evaluation plan and whether it follows logically from the evidence supporting the intervention; and the quality of the indicators, baselines and data sources.
- **Partners should be required to use the standard templates that have been developed, rather than using their own approaches.** To create greater consistency within the system and to ensure that an appropriate level of detail is provided by a partner in their application, progress reports and final reports, the standard templates that currently exist for applications (S01/51), progress reports (S11/S61) and final reports (S21/81) should be made mandatory for partners to use and not just as a checklist. As an interim step, when working with partners with very low capacity, programme officers might use information provided by partners to fill in the forms themselves, and clearly some agreement would need to be reached when projects are applying for joint support from several donors.

We recognise that some may feel that this recommendation runs counter to the principles of the Paris and Busan declarations; however, the content of these templates follows wide international practice and in many instances is unlikely to be very different from the approach partners are using themselves, and approaches being advocated by other donors. The absence of standardisation is undermining effective results measurement by Norway.

- **Add a rating assessment to the Progress and Final Report templates for partners to complete themselves.** Reporting currently makes little use of indicators and does not require any judgement about performance by the grantee or programme officer. This is in contrast to the practice of the agencies reviewed for comparison and contributes to weak evaluation awareness among staff. A rating is something that can be done even where the partner has poor indicators and limited data collection. Agreement or otherwise with the rating would be indicated by grant managers as part of their Quality Assurance checklist. These data should be entered into a performance database, where they would enable comparative analysis across sectors and over time and would provide a strong foundation for analysis to target thematic evaluations.

Strengthening the quality and use of reviews and evaluations at grant level

- **A more considered and strategic approach to the use of evaluations and reviews at the grant level should be developed.** As part of the preparatory phase of a grant, greater consideration should be given to whether reviews and/or evaluations will be commissioned and why, and the budget implications. In order to inform whether and what type of evaluation should be undertaken a review of the evidence base supporting a proposed intervention (and its quality) should be undertaken. This will provide a logical foundation for deciding the type and details of evaluation to be planned and will help ensure a more strategic use of evaluations. Where the evidence is strong a Final Report may suffice, where an intervention is new and innovative and the evidence base is weaker an attribution or contribution evaluation may be necessary. This detail may not be necessary at the initial application stage, but should be undertaken before a contract is signed.
- **Standard checklists should be developed for quality assuring grant-level evaluations and reviews.** These should cover both quality at entry (ToR, inception report etc.) and exit (final evaluation/review report). These checklists would help staff to make clear judgements on quality and more easily identify when a report is of insufficient quality. Examples are available from both the World Bank and DFID.

Strengthening the technical support provided to staff on results measurement

Staff training

- **Develop a more comprehensive training programme to support staff capacity in results measurement.** The training programme should offer more in-depth and longer-term training for those who want to deepen their skills in results measurement. This more in-depth training should be mandatory for Norad advisers who are providing technical advice on results frameworks. The programme should also combine e-learning and self-study modules for staff too busy to attend. In the case of DFID for example, a one week training course on commissioning and managing evaluations is run in collaboration with the UK Evaluation Society. This is available to all evaluation and results advisers. Evaluation and results advisers (in common with other sectoral disciplines) also attend a professional meeting once a year and have a ‘head of profession’ appointed to look after their professional interests. The revised training programme should cover the following issues in greater depth: the development and appraisal of theories of change; how to appraise the quality of evidence that underpins a programme’s design; and reviewing and appraising a partner’s results measurement systems.

Technical support

- **The requirements on technical assistance and Quality Assurance should be harmonised across all grant scheme rules.** The current diversity in results requirements across the grant scheme rules should be harmonised. Rather than have each scheme set its own requirements, we suggest using budget thresholds to determine at which point it is mandatory to get technical assistance and QA (see below for more details). Within each budget category the requirements would be the same.

Additional support material

- **An online resource hub should be developed that provides staff access to examples of good practice in results measurement and pools sector-specific resources.**
- **Develop the capacity of grant recipients to measure results, such as by e-learning, but perhaps also through a ‘partners’ guide to managing for results’.** Provide outreach to partners to raise their skills in results measurement.

6.2 Recommendations for Norad's Evaluation Department

The second set of recommendations deal with the work of EVAL. These have been grouped under three headings: improving the design of an evaluation; improving the management of an evaluation; and conducting impact evaluations.

Designing an evaluation

- **Tighten the design specifications for evaluations.** Draft ToRs with tighter specifications for the purpose, objective and scope of evaluations so it is clear when outcome or impact is to be evaluated in addition to outputs.
- **Keep evaluation questions focused.** Reduce the number of evaluation questions that are to be covered by an evaluation so that resources are clearly prioritised to key results.
- **Require evaluators to clearly describe the programme logic of the intervention being evaluated.** All evaluations should be required to specify the programme logic or reconstruct it if necessary as a basis for the design.
- **Be more specific in ToRs about the required consultants' skills.** More consideration should be given to the specific skills and expertise required for either the team leader or core team members. This would require EVAL to do more preparation up front around which evaluation designs and methods are best suited to answer the evaluation questions.

Managing an evaluation

- **Monitor the progress of evaluations more closely.** Once an inception report has been agreed, EVAL should plan periodic check-ins with evaluation teams to ensure the process is on track, and delivering according to what has been agreed. Providing this is conducted with appropriate sensitivity, this should not affect the independence of the evaluation.

Conducting impact evaluations

- **Develop a clear process for deciding and managing impact evaluations.** To be most effective, impact evaluations should be set up at the design stage of an aid project, and will require a joint decision with the implementers or partners.
- **When conducting impact evaluations that seek to quantify attribution, ensure the appropriate competencies exist both among the EVAL staff managing the evaluation and the consultants.**
- **Ensure the specification of methodologies, data requirements and competencies of evaluators are in line with the requirements for outcome and impact evaluations.**

6.3 Options for implementation to address more structural issues

The detailed recommendations in Chapter 6.1 tackle a wide range of technical issues, but they do not address the findings that emerged in the course of the evaluation about skills and ways of working. Different combinations of reforms are possible to tackle what we regard as a core weakness in the system: the fact that current guidelines are not being followed.

We consider two approaches that can be used to address this problem: by concentrating expertise or by broadening it. In the following section, we describe what both of these two approaches entail and list the recommendations that would need to be taken forward to put each into practice. Cutting across each approach are a number of common recommendations. They provide the foundations of both approaches and we deal with them first.

Cross-cutting recommendations

- **Strengthen the support at senior management level for results measurement.** There needs to be more visible action among the leadership of the MFA and Norad to insist on better results management and evaluability. Changing processes or structures are not enough to change behaviour and culture. Experience with results management systems indicates that senior management incentives and direction to staff are key elements. Strengthening a results-focus in the grant management process will require staff to spend more time with partners, developing and reviewing results frameworks, reporting and evaluations; senior management needs to create the space and resources for this to take place.
- **Improve staff incentives for measuring results and ensuring evaluability.** At present, training is not given a high priority by managers and there are no career rewards for expertise in results management. A strong incentive would be to incorporate continuing professional development in results management as a positive career attribute to be recorded on personnel files and factored into career development. This recommendation is addressed to top management in the MFA and Norad.
- **Establish a consistent basis, such as size of grant, for which technical assistance and Quality Assurance of partners' results frameworks, reporting and evaluation plans is mandatory.** As part of the harmonisation of Grant Scheme Rules, there should be a clear requirement for mandatory technical assistance for all projects greater than a certain amount and sample assessment of mid-size grants. We do not think it is feasible to replicate this process for small grants.^{93,94}

93 The precise limits would need to be agreed, but considering the distribution of project and programme grants in 2012, setting the mandatory requirement at greater than NOK 25 million would mean reviewing 83 grants and cover 64 percent of the committed value. Depending on the available staff and workload, a limit of NOK 15 million might be practical. Mid-range grants could then be NOK 5 million to NOK 15 million. These are the same limits that have been used under past arrangements.

94 We are aware that this is the approach taken in the old Development Cooperation Manual. This is a consistent and easier approach for staff to work with.

Approach 1: Concentrating expertise

This would involve re-designing the approach to QA and expanding the role of AMOR so that all eligible grants receive QA on their content before approval and upon completion. This approach concentrates actions to improve quality, to build a model that reflects high international standards of good practice among comparable donors. It would bring a high degree of consistency, but has substantial implications for staffing in the short time span necessary to process grants.

There is also evidence from both DFID and the World Bank that an excessive application of QA and review can lead to risk-averse behaviour, which is contrary to the way Norway approaches development aid, and undesirable.⁹⁵ For that reason, we propose mandatory QA on only a small proportion of grants. This approach would bring radical change to the grant management system. The recommendations that would need to be taken forward are:

- **Resource AMOR to provide more comprehensive support to all eligible grants on their results measurement frameworks, and evaluation plans and reports at approval and completion.** AMOR should be provided with additional resources to conduct QA prior to grant approval and at grant completion of all large projects, over the agreed amount. This could be contracted out or could be staffed as temporary teams of peer reviewers, such as were used in the World Bank Quality Assurance Group. As part of this process, AMOR could also analyse and validate a sample of grant-level reviews and evaluations and feed this into the Annual Results Report.

Approach 2: Broadening expertise

This would involve mainstreaming skills to assist the quality of planning for results. A programme to transfer specialist skills to a large number of volunteer staff across Norad, the MFA and the embassies would bring specialist advice closer in space and time to grant managers. This approach is modelled on the evaluation focal points adopted by DFID. It has the advantage of building broad capacity and not requiring a substantial increase in staff numbers. It would need to be supported by changes such as the use of performance ratings indicated elsewhere.

This approach is designed more to work within current practices and build through progressive change rather than radical change. It broadens the approach by developing staff capacity and working more closely with partners, which fits with local ownership and capacity building of partners' systems. It is less resource-intensive than Approach 1. The recommendations that would need to be taken forward are:

- **Build a cadre of staff specialised in results measurement and evaluation.** Norad and the MFA should deepen the skills of its staff in measuring results. We recommend, alongside an improved training

⁹⁵ See Annex 5.

programme, that a small number of staff from across the organisations be given intensive training in evaluation and results measurement. This would broaden skills within the organisations and create a cadre of results and evaluation advisers across the organisation to give local/regional advice and train others. Embedded advisers would offer a more flexible and informal form of support to staff around how to measure results. They would bring results advice closer to the ground. Focal points could specialise in specific sectors. To encourage participation in the scheme, an incentive system should be developed that allowed for career rewards and development within the cadre.

- **Design a new approach to reach out to partners, with a combination of improved technical guidance, some direct capacity building and access to the self-study materials.** More focus should be put on developing the capacity of partners to improve their measurement and reporting of results.

The two approaches represent different strategies but are not necessarily mutually exclusive. Both (or elements of both) could be taken forward in tandem. For example, greater resources could be provided to AMOR to undertake more detailed Quality Assurance of results frameworks and evaluation plans and reports, and a network of results advisers could be created that provided more informal support to staff managing grants on measuring results.

References

Norad and MFA policies, guidelines and procedures

Norad (1999), The Logical Framework Approach: Handbook of objectives orientated planning (fourth edition), Norad, Oslo

Norad (2012), Overordnede retningslinjer for tilskuddsforvaltning i Norad, Norad, Oslo

Norwegian Ministry of Foreign Affairs and Norad (2005), Development Cooperation Manual, MFA/Norad, Oslo

Norwegian Ministry of Foreign Affairs and Norad (2006), The Agreement Manual, MFA/Norad, Oslo

Norwegian Ministry of Foreign Affairs (2008), Håndbok i tilskuddsforvaltning Om å forvalte tilskudd selv, MFA, Oslo

Norwegian Ministry of Foreign Affairs (2008), Håndbok i tilskuddsforvaltning Om å etablere og å forvalte ordninger, MFA, Oslo

Norwegian Ministry of Foreign Affairs (2008), Håndbok i tilskuddsforvaltning Om å forvalte delegerte ordninger, MFA, Oslo

Norwegian Ministry of Foreign Affairs and Norad (2008), Results Measurement in Norwegian Development Cooperation: A Practitioner Guide, MFA/Norad, Oslo

Norwegian Ministry of Foreign Affairs (2009), Grant Management Manual: Management of Grants by the Ministry of Foreign Affairs and embassies, MFA, Oslo

Norwegian Ministry of Foreign Affairs (2010), S51 - Application form for project/programme support, MFA, Oslo

Norwegian Ministry of Foreign Affairs (2010), Guide to filling in S51 – Application form for programme and project support, MFA, Oslo

Norwegian Ministry of Foreign Affairs (2010), S61 - Progress report form for project/programme support, MFA, Oslo

Norwegian Ministry of Foreign Affairs (2010), S81 – Final report form for project/ programme support, MFA, Oslo

Norwegian Ministry of Foreign Affairs (2012), Risk management: methods and terminology at the Ministry of Foreign Affairs, MFA, Oslo

Norwegian Ministry of Foreign Affairs (2013), Grant Management Manual: Management of Grants by the Ministry of Foreign Affairs and Norad, MFA, Oslo

Utenrikstjenestens Kompetansesenter (2013) Kurskatalog 2. Halvår 2013, MFA, Oslo

State-level rules and guides

Senter for Statlig Økonomistyring (2010), Evaluering av statlige tilskuddsordninger, SSØ, Oslo

Senter for Statlig Økonomistyring (2010), Resultatmåling, Mål- og resultatstyring i staten, SSØ, Oslo

Det Kongelige Finansdepartement (2013), Reglement for økonomistyring i staten Bestemmelser om økonomistyring i staten Fastsatt 12. desember 2003 med endringer, senest 18. september 2013, Det Kongelige Finansdepartement, Oslo

Grant management reviews

Norad and Norwegian Ministry of Foreign Affairs (2012), Guide for completing the checklist in a Grant Management Review (GMR), Norad/MFA, Oslo

Norad and Norwegian Ministry of Foreign Affairs (2013), Grant Management Review of the Norwegian Embassy in Abuja, 10th to 15th of October 2012, Norad/MFA, Oslo

Norad and Norwegian Ministry of Foreign Affairs (2013), Grant Management Review at the Norwegian Embassy in Beijing, 14th to 21st January 2013, Norad/MFA, Oslo

Norad and Norwegian Ministry of Foreign Affairs (2013), Grant Management Review at the Norwegian Embassy in Addis Ababa, 5th to 12th March 2012, Norad/MFA, Oslo

Norad and Norwegian Ministry of Foreign Affairs (2013), Grant Management Review at the Norwegian Embassy in Luanda, 9th to 15th April 2013, Norad/MFA, Oslo

Norad and Norwegian Ministry of Foreign Affairs (2013), Grant Management Review of the Norwegian Embassy in Khartoum, 10th to 16th of June 2013, Norad/MFA, Oslo

EVAL references

Norad (2010), Evaluation of Norwegian Business Related Assistance: Uganda Case study, Evaluation Report 6/10, Norad, Oslo

Norad (2010), Evaluation of Norwegian Development Cooperation with Western Balkans, Evaluation Report 7/10, Norad, Oslo

Norad (2010), Real Time Evaluation of Norway's International Climate and Forest Initiative. Country report: Indonesia, Evaluation Report 16/10, Norad, Oslo

Norad (2011), Annual Report 2011: Evaluation of Norwegian Development Cooperation, Norad, Oslo

Norad (2011), Contextual Choices in Fighting Corruption: Lessons Learned, Evaluation Report 4/11, Norad, Oslo

Norad (2011), Evaluation of Norwegian Health Sector Support to Botswana, Evaluation Report 10/11, Norad, Oslo

Norad (2012), Evaluation of the Health Results Innovation Trust Fund, Evaluation Report 4/12, Norad, Oslo

Other references

OECD-DAC (2010), DAC Guidelines and Reference Series Quality Standards for Development Evaluation, OECD, Paris

OAG (Office of the Auditor General) (2011), The Office of the Auditor General's Investigation into Results Orientation in Norwegian Development Cooperation (Norwegian translation), Document 3:4 (2010-2011), 13 January 2011, Oslo

OECD/DAC (2013), Peer Review: Review of the development co-operation policies and programmes of Norway, 6 November 2013.

Grant-level reviews and evaluations

Baptist Union of Norway (2008), Baptist Community in Northern Congo (CBCN) Health Project

Norad (2008), Final evaluation of the UVETA project

Norwegian Humanist Association/Humanist (2008), Review of Integrated Health Camp for Southern Humla district

Plan Norway (2008), Final Evaluation of the Maternal and Child Health Project

Norad (2009), End Review of the Norwegian People Aid Mine Action Programme in Ethiopia 2005-2007

Norad (2009), Norwegian Refugee Council Review & Appraisal – Strengthening Child Protection Mechanisms through Emergency Education Support to Northern Districts of Kitgum and Pader

Norad (2009), Review of Norwegian support to National Democratic Institute in Kenya

Norwegian Church Aid (2009), NCA Ethiopia and Partners Engagement in Abandoning HTPs/FGM in Ethiopia Review of 9 partners contribution (2002-2008)

Norwegian Church Aid/Kirkens Nødshjelp KN (2009), Final Evaluation Report for the State of Eritrea Administration of Northern Red Sea Shebah-Demas Integrated Development Programme (SDIDP)

Union of Education Norway (2009), Union of Education Norway Evaluation Report of the Cooperation on Trade Unions between Education International, Union of Education Norway & Federation of Mongolian Education and Science Unions

ADRA Norge (2010), Budi Education Support Project – Final Evaluation

Norwegian Red Cross (2010), Impact evaluation: Capacity Building, Organizational Development and Disaster Preparedness Project in Vietnam

Save the Children Norway (2010), Final Evaluation of Community Base Rehabilitation (CBR) Programme

Save the Children Norway (2010), Final Evaluation of the ECDE Project in Karamoja

Pym – the Pentecostal Foreign Mission of Norway (2011), Evaluation of Various Components of the Kuatine e Project

Save the Children Nepal country office (2011), final evaluation of safer environment for girls project

Save the Children Norway (2011), Thematic Evaluation of Save the Children Norway's Cooperation with Partners – Zimbabwe Case Study

WWF-Norway (2011), Enhancing Livelihood Sustainability through Raising Community Capacity for Coastal Management (RaCCCoM) in Lamu Archipelago – Final evaluation report

Norwegian Church Aid (2012), We have faith – Act now for Climate Justice

Save the Children (2012), CRC Advocacy through Radio – Television programmes for children and youth – End of Project Evaluation

Annexes



Annex 1 Terms of Reference

Evaluation of the Norwegian Aid Administration's systems and practices to ensure Evaluability¹ of Norwegian Grants

Norad's Evaluation Department (EVAL) will commission an evaluation to assess the Norwegian aid administration's (i.e. the MFA, Norad and the embassies) internal processes for grant management² of Norwegian funded interventions. The evaluation will focus on how and to what degree the current regulations and practices ensure evaluability³ of the interventions, including results measurement.

1. Rationale and Evaluation Hypotheses

There has been an increasing demand globally for development aid to demonstrate results. The demand for results measurement and reporting is also clear in central documents guiding the Norwegian aid administration⁴.

None of the evaluations and studies commissioned by EVAL⁵ and finalised in 2011 could report sufficiently on results at the level of outcomes or impact⁶. This is also a recurrent finding in previous evaluation reports commissioned by the department⁷. The independent evaluations commissioned by EVAL build on the findings reported in the internal reviews⁸ of interventions, which are commissioned by the section/ department/unit providing financial support to the interventions.

1 Evaluability refers to the degree to which an intervention is possible to evaluate in a reliable and credible manner, i.e. availability and reliability of data.

2 The term "grant management" is used for the Norwegian "tilskuddsforvaltning" and includes the entire grant management process; appraisal/approval/decision documents, results measurement, quality assurance, reviews and evaluations.

3 See footnote 1 for a definition of "evaluability".

4 In *Climate, Conflict and Capital*, Report nr. 13 to the Storting (p 110, first two bullet points), it is stated that: "*The Government will: strengthen focus on results and aid effectiveness ; ensure transparency and access to information with regard to the Norwegian aid administration*".

5 Norad's evaluation department (EVAL) is mandated to evaluate interventions financed over the aid budget (03-area of the State budget). EVAL's mandate is to: "document the extent to which the Ministry's, the embassies' and Norad's efforts to realise the objectives laid down for Norwegian development policy are effective, relevant and achieve the intended outcomes". The Instruction in its entirety can be found at <http://www.norad.no/en/evaluation> : "Instructions for Evaluation Activities in Norwegian Aid Administration" (2006).

6 The statement that one has not been able to report on findings at *outcome or impact level* does *not* imply that there are no outcomes or impact from the interventions (only that they have not been sufficiently measured), nor does it imply that there is no reporting on *output level* (e.g. number of schools built).

7 Some examples from recent evaluations commissioned by EVAL: a) The report from the *Evaluation of the Strategy for Norway's Culture and Sports Cooperation with Countries in the South* points to both a lack of criteria for assuring quality and assessing performance and the absence of data for evaluating success or failure; b) The *Evaluation of Results of Development Cooperation through Norwegian NGOs in East Africa* reported that 'most projects lacked the data and information required to be able to measure changes in indicators for key results accurately'; c) The *Evaluation of Norwegian Development Cooperation to promote Human Rights* could report on the increased number of university-based human rights centers in Indonesia as a result of Norwegian support. Still the report had to conclude that 'there is no clear indication of what impact this has had on actual human rights compliance'.

8 «Gjennomganger».

This deficiency of reporting on results at the level of outcomes and impact in EVAL-commissioned evaluations might thus be due to a variety of reasons (evaluation hypotheses);

- 1) The evaluators commissioned by EVAL are not looking hard enough for alternative sources of data in order to report on results on outcome and impact levels, which might be due to:
 - a. Insufficient requirements on the subject of results measurement and reporting in the planning and follow-up of evaluations (i.e. in Terms of References and quality assurance of reports) from EVAL; and/or:
 - b. Insufficient competencies of evaluators with regard to results measurement and evidence; and/or:
- 2) Insufficient requirements for evaluability of interventions in the grant management process in the Aid Administration (for appraisal/approval, quality assurance, and results measurement, reviews and evaluations), which might be due to:
 - a. Insufficient requirements to ensure evaluability, including results measurement and reporting in guidelines and standard documents such as handbooks, grant scheme rules and templates in planning, implementation and follow-up of interventions; and/or:
 - b. Insufficient training, technical advice and/or quality assurance; and/or
 - c. Insufficient implementation, i.e. routines in practice, and staff skills in the planning, appraisal and follow-up process (appraisal/approval, results measurement, reviews and evaluations) in the aid administration.

2. Purpose and objectives

The purpose of the evaluation is to contribute to further learning and progress in the aid administration's follow-up of the government's demand for a strengthened focus on results⁹ by identifying reasons for the insufficient results documentation and provide evidence-based recommendations for improvements in this area.

9 Norad's Department for Quality Assurance (AMOR) undertook a review of the results-focus and risk analysis in selected cases of technical advice from Norad in 2009; "Gjennomgang av resultatfokus og risikohensyn i Norads faglige rådgivning". The review concluded that "results" and "risks" do not receive sufficient attention to the degree that Norad's senior management would expect. The conclusion was that the weaknesses seemed to be more systematic than dependent on the individual staff member involved. The review provided ten recommendations, among them a follow up of the review 2-3 years later to monitor the progress of this work. Such a follow up of the review has not yet happened. The theme of this current evaluation is similar to the review undertaken by AMOR in 2009, but this evaluation focuses on grant management processes and not on the content of Norad's technical advice.

The main users of the evaluation's findings will be the Norwegian Ministry of Foreign Affairs (MFA) and other stakeholders who have direct or indirect interest in the subject of this evaluation. In this context, the *Ministry of Foreign Affairs, MFA* refers to its political leadership, its officials, the Norwegian Embassies and the Norwegian Agency for Development Cooperation, Norad. Other *stakeholders* include other development agencies, their evaluation units and governmental twinning partners.

The objective of the evaluation is to test the following hypotheses:

Hypothesis 1a) Insufficient requirements for results measurement and reporting in the planning and follow-up of evaluations (i.e. in ToRs and quality assurance of reports) from EVAL;

The evaluation shall review the planning and implementation process of five of EVAL's recent evaluations to assess the degree to which the evaluation department have been sufficiently focussing on results measurement in ToRs and in the quality assurance of reports (clarity on theory of change, testing of assumptions/hypotheses, requests for results measurement and reporting on outcome and impact levels, etc.). Based on the findings, the evaluation shall provide recommendations for improvements.

Hypothesis 1b) Insufficient competencies of evaluators with regard to results measurement and finding/using evidence

The evaluation shall map out what kind of data the evaluators were looking for/not finding/not using to report on results at outcome and impact levels in five randomly selected recent EVAL-commissioned evaluations. This will include an outline of types of available data (reported and not reported) of the five evaluations as well as information from the evaluators regarding their evaluation approaches and methods.

Hypothesis 2a): Insufficient requirements for evaluability of interventions in the grant management process

The evaluation shall review guidelines/handbooks/regulations/standard documents¹⁰ and assess to what degree there are written requirements for documentation of theory of change, identification of assumptions, assessment of available evidence including assessment of the quality and reliability of evidence, development of indicators, systems of results measurement, etc. in the grant management process. The evaluation shall also investigate to what degree the current guidelines and standard documents facilitate evaluability, results measurement and impact evaluations. Furthermore, the evaluation shall examine to what degree standard procedures for quality assurance of grant management processes, reviews and evaluations are in place.

10 See list of relevant documents (not exhaustive) in Annex 4.

Hypotheses 2b) Insufficient training, technical advice and/or quality assurance

- i. The evaluation shall assess the training provided to staff to ensure evaluability including results measurement, i.e. the type of training offered and whether it is compulsory, its frequency, content, relevance and usefulness.
- ii. The evaluation shall furthermore assess the technical advice provided by Norad to the MFA and the embassies in the planning and review process with a focus on ensuring evaluability (theory of change, identification and assessment of evidence that the intervention will lead to the goal, indicators, results measurement and embedded evaluation).
- iii. The evaluation shall also assess the procedures and systems in place for *quality assurance* of
 - a. approval of interventions (i.e. in the planning process) with a focus on evaluability (theory of change, identification and assessment of evidence that the intervention will lead to the goal, indicators, results measurement and embedded evaluation), and;
 - b. reviews (mid-term-/end-) with a focus on results measurement and reporting.

Hypothesis 2c): Insufficient implementation of the rules and regulations in the aid administration

The evaluation shall examine how the guidelines, handbooks and regulations are being implemented in the aid administration, i.e. routines, systems for quality assurance, and staff skills (in the grant management process; appraisal/ approval, results measurement, reviews and evaluations). The evaluation shall specifically look into to what degree the theory of change (programme theory) is documented, assumptions are accounted for, the availability and quality (strength) of evidence is assessed, relevant indicators developed, and sufficient funds for reviews and evaluations allocated (including for baseline studies with or without control groups), with the view to ensure the evaluability of Norwegian-funded interventions. This shall include an assessment of to what degree results measurement and evaluations are planned for within Norwegian funded interventions prior to implementation, i.e. “embedded in programme planning”.

Reasons for any deviation in the written regulations and practice shall be explored. This could potentially include – but not be limited to - issues concerning the guidelines and regulations themselves (clarity/relevance/level of ambition/abundance?), or more systemic issues like incentives/sanctions, time and/or personnel constraints, competing requirements, political and administrative demand to spend the money within the annual budgets, operations in fragile states and unstable environments; or it could be due to

insufficient training/lack of competencies among individual staff members (front-line staff and their line managers).

The testing of the hypotheses shall result in :

- a) an overview of regulations and routines (standard requirements and common practices) to ensure evaluability including results measurement and quality assurance in the Norwegian aid administration (i.e. both for EVAL's work and for the grant management processes within the MFA, Norad and the embassies), including an analysis of possible main impediments for sufficient reporting on results;
- b) a comparable overview of routines and regulations for ensuring evaluability and results measurement in the grant management processes of three-four similar aid organisations (DfID and WB, and one-two others) to identify alternative practices that might offer relevant lessons; and
- c) evidence-based recommendations for improvements of the results measurement and reporting in the Norwegian aid administration and in EVAL.

3. Scope and Suggested methods

The evaluation shall focus on the current situation (of currently endorsed standard documents, rules, regulations and practices). For the assessment of the grant management process from appraisal to closure, the evaluators should identify interventions that have terminated recently (i.e. within the past couple of years), and interventions that have been approved during the past 3-5 years. The assessment of the reviews should also focus on reviews undertaken and completed during the past 3 years.

The tenderers are encouraged to provide sound solutions for the application of methods in this evaluation. The outline below is only suggestive and meant to spark off ideas. The methods should be presented in an evaluation framework matrix (including identified evaluation questions, information need and sources of information (remember triangulation), methods for information/data collection (including selection methods and sample sizes), etc.

- 1) Assessment of the result-focus in a selection of recent ToRs from EVAL, and in the follow-up (e.g. comments to draft reports and subsequent correspondence). Desk review of five recent ToRs.
- 2) Assessment of evaluators' competencies in five randomly selected recent EVAL-commissioned evaluations;
 - a. Assessment of the replicability and reliability of five randomly selected EVAL-commissioned evaluations (i.e. assess the extent to which they are consistently replicable; the degree to which methods

and data are transparently presented so other evaluators would come to the same conclusions with the available data and evidence).

- b. Assessment of the results reporting in the evaluations; Desk review mapping available data, i.e. the kind of data the evaluators reported on, what kind of data they were looking for and not finding and what kind of available data they were not using (not looking for or not finding) to report on results at outcome and impact levels.
 - c. Assessment of the consultant's results-focus and competencies. Interviews with the evaluators of the selected evaluations to identify alternative reasons for the lack of reporting on results at outcome and impact levels.
- 3) Assessment of the degree to which relevant documents in the Norwegian aid administration such as guidelines, handbooks, standard contracts, grant scheme rules¹¹, instructions, etc, focus on- and set a standard for evaluability and results measurement. See attached list of relevant documents in Annex 4 (the list of documents is partly in Norwegian). To what degree the documents are clear in requesting/recommending a theory of change¹² as part of the planning process shall also be assessed. This should also include an consideration of the rules and regulations as a whole; the overall number of documents, their ease of reference and use; degree of clarity of the documents; and their hierarchical position. Desk review and interviews with relevant staff working with quality assurance in Norad and the MFA.
- 4) Assessment of courses/training/capacity building provided to ensure evaluability and including results measurement. Content, relevance and quality of the course material shall be assessed in terms of to what degree they facilitate learning regarding evaluability issues. Desk review and interviews.
- 5) Assessment of the procedures and systems in place for quality assurance of deliveries (like technical advice), of the planning and approval-processes of supported interventions, and of the reviews.
- 6) Assessment of the grant management process from a to z (from appraisal to closure) of 20 randomly selected Norwegian-funded interventions – (from MFA, Norad and the embassies); to examine to what extent the part of the guidelines/handbooks/regulations that deal with evaluability, results measurement and quality assurance have been followed. Included here is an assessment of technical advice provided by Norad staff. Reasons for

11 «Ordningsregelverk».

12 That is, to what degree the documents request/set a standard for (a) clear goal(s), identification and quality assessment and rating of available evidence that the planned intervention will lead to the goal, and an assessment of the need for evaluations depending on the availability and reliability of evidence.

deviances from the regulations shall be explored and described in the final report¹³. Desk review of relevant documents, interviews with relevant staff.

- 7) Assessment of similar kinds of standard documents (described under point 3) and grant management procedures (planning/approval/implementation/follow-up/evaluation), including systems for quality assurance in other organisations (DfID, WB and 1-2 more) to identify relevant lessons. Desk reviews and interviews.
- 8) Assessment of the degree to which reviews¹⁴ of Norwegian funded interventions make use of a theory of change, (including the testing of assumptions), results measurement and reporting at output, outcome and impact levels. Included here is an assessment of technical advice provided by Norad staff. Desk review of 20 randomly selected reviews from the MFA, the embassies and Norad.
- 9) Assessment of how front-line staff (staff responsible for grant management) in MFA, the embassies and Norad perceive the guidelines and how they operationalise and use them in their work. Gaps in staff skills and qualifications regarding result based management, theory of change, assessment of evidence, and results measurement, reviews and evaluations shall also be assessed. Following a preliminary analysis of the desk reviews and interviews with relevant staff, a questionnaire survey will be developed and sent to front-line staff, and the results will be analysed in the report.

¹³ This could potentially include – but not be limited to – issues concerning the guidelines and regulations themselves (clarity/relevance/level of ambition), time and/or personnel constraints, competing requirements, insufficient training/lack of competencies; political and administrative demand to spend the money within the annual budgets, operations in fragile states and environments, etc.

¹⁴ "Gjennomganger"; mid-term-reviews and end-reviews.

4. Composition of the Team

The team shall cover the following competencies (these must be documented in the tender):

Competence	Team leader	At least one team member (can also be the team leader)
Academic	Higher academic degree	Higher academic degree
Discipline	Economics and/or Development and/or Evaluation/and or Social Sciences	Economics and/or Development and/or Evaluation and/or Social Sciences
Evaluation	Significant experience in managing corporate evaluations and multi-disciplinary teams. Thorough understanding of theory-based evaluations Knowledge of and familiarity with different institutional set-ups for development evaluation.	Working experience with evaluation methodologies and DAC evaluation quality standards Experience with results measurement, and familiarity with UNEG's framework for professional peer reviews of evaluation units Thorough understanding of theory based evaluations Documented knowledge of monitoring's role in evaluation Experience in analysing incentive structures
Methodologies	Working experience with mixed method designs	Experience in designing and conducting impact evaluations
Methodologies		Working experience with qualitative methods
Development cooperation	Extensive experience with evaluations in/of international development cooperation	Familiarity with DfID's and IEG's (WB) Evaluation Work and institutional set-up for operations and evaluations and/or other relevant organisations' evaluation systems
Development Cooperation		Thorough knowledge of Norwegian Development Cooperation including the institutional set-up for operations and evaluations
Language	Excellent oral and written English	Excellent oral and written English
Language		Reading ability/understanding of Norwegian or Swedish or Danish

Tendering institution

The tendering institution is expected to have experience with delivering multi-disciplinary evaluations contracted preferably through competitive procurement process during last three years.

5. Budget, time frame and deliverables

A maximum of 2,2 million NoK has been allocated to this evaluation. The evaluation shall start in March 2013 and shall be completed before the end of the year. Deadline for reports are as outlined in the table on page 2 of this Tender Document.

The deliverables in the consultancy consist of the following outputs:

- **Inception Report** not exceeding 15 pages shall be prepared and discussed with the reference group before final approval by EVAL.
- One work-in-progress reporting **seminar**.
- **Draft Final Report** for preliminary approval by EVAL for circulation to the stakeholders. The stakeholders shall provide feedback that will include comments on structure, facts, content, and conclusions.
- **Final Draft Evaluation Report**.
- **Policy brief** (not exceeding 2 pages).
- **Seminar for dissemination** of the final report. Direct travel-cost related to dissemination in the case countries if any, will be covered separately on need basis, and shall not be included in the budget.

Poorly substantiated findings will not be accepted in the reports. In connection with questions where the team does not find sufficient information to make meaningful assessments, the team will list the sources sought and not found and / or describe the type of information sources they would have required to carry out such an assessment.

All presentations and reports (to be prepared in accordance with EVAL's guidelines given in *Annex A-3 Guidelines for Reports* of this document) are to be submitted in electronic form in accordance with the deadlines set in the time-schedule specified under *Section 2 Administrative Conditions* in *Part 1 Tender specification* of this document. The data collected during the study shall be submitted in EXCEL format. EVAL retains the sole rights with respect to all distribution, dissemination and publication of the deliverables.

Annex 2 List of persons consulted

Family name	Given name	Post	Unit
Norad			
Aakre	Erik	Assistant Director	Department for Quality Assurance, Results Management Section, Norad
Bæra	Svein	Director	Civil Society Department, Norad
Brandtzæg	Marit E.	Director	Department for Quality Assurance, Norad
Buli	Ingrid	Senior adviser	Civil Society Department, Section for Development Initiatives, Norad
Bull	Beate	Senior Adviser	Evaluation Department, Norad (at the time)
Hansen	Kari Edvardsdal	Senior Adviser	Department for Quality Assurance, Aid Management Section, Norad
Herstad	Bente	Policy director	Department for Climate, Energy and Environment, Section for Sustainable Development and Environment, Norad
Jansen	Eirik Gjøstein	Senior Adviser	Evaluation Department, Norad (now retired)
Lillehammer	Giske Charlotte	Adviser	Department for Economic Development, Gender and Governance, Section for Development Strategy and Governance, Norad
Lindøe	Kjersti	Senior Adviser	Civil Society Department, Section for Development Initiatives, Norad
Lillestøl	Siv J.	Senior Advisor	Evaluation Department, Norad
Melby	Hans Peter	Senior Advisor	Evaluation Department, Norad
Olsen	Arne	Senior Adviser	Department for Economic Development, Gender and Governance, Oil for Development Section, Norad
Padmanabhan	Sheila	Consultant	Department for Economic Development, Gender and Governance, Section for Rights and Gender Equality, Norad

Family name	Given name	Post	Unit
Halvorsen	Vigdis	Assistant Director	Civil Society Department, Section for Civil Society Strengthening, Norad
Hansen	Camilla Solvang	Adviser	Department for Economic Development, Gender and Governance, Section for Private Sector Development, Norad
Singh	Balbir	Senior Adviser	Evaluation Department, Norad
Sjøveian	Solbjørg	Senior Adviser	Civil Society Department, Section for Civil Society Strengthening, Norad
Skanche	Gunvor Wittersø	Assistant Director	Civil Society Department, Section for Civil Society Strengthening, Norad
Stave	Jørn	Senior Adviser	Evaluation Department, Norad (now Multiconsult)
Strand	Marit Marie	Senior Adviser	Department for Economic Development, Gender and Governance, Section for Development Strategy and Governance, Norad
Thonstad	Hanne	Senior Adviser	Department for Quality Assurance, Statistics Team, Norad
Vestbøstad	Katrine	Advisor	Department for Climate, Energy and the Environment, Section for Renewable Energy, Norad
MFA/Embassies			
Austad	Håvard	Senior Adviser	Section for Russia, Eurasia and regional co-operation , MFA
Azedo	Ann-Helen Perez	Senior Adviser	Department for Regional Affairs and Development, Section for the Management of Subsidiary Agencies and Development Funds, MFA
Bleken	Sidsel	Assistant Director	Department for UN, Peace and Humanitarian Affairs, Section for Budget and Administration, MFA
Berge	Gunnvor	Senior Adviser	Department for Regional Affairs and Development, Section for Sub-Saharan Africa, MFA
Boye	Thorvald	Senior Adviser	Royal Norwegian Embassy in Monrovia, Liberia
Brisis	Pierre Michel De	Adviser	Department for UN, Peace and Humanitarian Affairs, Section for Budget and Administration, MFA
Eikeland	Elin	Counsellor	Royal Norwegian Embassy in Nairobi, Kenya.

Family name	Given name	Post	Unit
Fladby	Berit	Policy Director	Department for UN, Peace and Humanitarian Affairs, Section for Multilateral Development Assistance and Global Economic Issues, MFA
Fyhri	Torgeir	Senior Adviser	Department for UN, Peace and Humanitarian Affairs, Section for Humanitarian Affairs, MFA
Helstrøm	Wera	Senior Adviser	Department for Regional Affairs and Development, Section for Sub-Saharan Africa, MFA
Hjertvik	Stine	Assistant Director General	Finance Section, Grant Management Unit, MFA
Ilseas	Per Albert	Senior Adviser	Department for UN, Peace and Humanitarian Affairs, Section for Peace and Reconciliation, MFA
Langlete	Toril	Senior Adviser	Department for Regional Affairs and Development, South Asia and Afghanistan Section, MFA
Malvik	Henrik	Senior Adviser	Department for UN, Peace and Humanitarian Affairs, Section for Climate Change, Global Health and Sustainable Development, MFA
Pettersen	Marit Viktoria	Senior Adviser	Department for UN, Peace and Humanitarian Affairs, Section for Climate Change, Global Health and Sustainable Development, MFA
Salvesen	Hilde	Senior Adviser	Department for UN, Peace and Humanitarian Affairs, Section for Humanitarian Affairs, MFA
Stormark	Kåre	Director	Department for UN, Peace and Humanitarian Affairs, MFA
Syse	Christian	Deputy Secretary General	MFA
Søegaard	Vibeke G.	Senior Adviser	Department for UN, Peace and Humanitarian Affairs, Section for Budget and Administration, MFA
Sørum	Rolf Erik	Senior Adviser	Human and Financial Resources Department, Finance Section, Ministry of Foreign Affairs
Vijayanathan	Loganathan	Senior Adviser	Royal Norwegian Embassy in Colombo, Sri Lanka
Øksnevad	Arild	Senior Adviser	Regional Department, Section for Sub Agencies Development Funds, MFA

Family name	Given name	Post	Unit
Øveraas	Ole	Senior Adviser	Department for Regional Affairs and Development, Section for Development Policy, MFA
External			
Fitzgibbon	Andrew	Conservation Director, Programmes	WWF-Norway
Keler	Adrien	Principal Manager Donor Co-Financing	European Bank for Reconstruction and Development
Nøhr	Henning	CTA/Deputy Evaluation Department (EVAL)	Evaluation Department, Ministry of Foreign Affairs, Denmark
Ruud	Helene	Senior Adviser, Disaster Risk Reduction	Norwegian Refugee Council
Scobey	Rick	Deputy to Director General	Independent Evaluation Group, The World Bank
Urheim	Kristoffer Beer	Investment Manager	Financial Institutions and Funds, Norfund
Volden	Margrethe	Area Team Leader Middle East, Afghanistan and Pakistan	Norwegian Church Aid

Annex 3 Table summarising findings, conclusions and recommendations

Finding	Conclusion	Recommendation
- The minimum requirements on results measurement outlined in the Grant Management Manual are not adequate to ensure evaluability. In particular, the level of information required of partners during the planning of grants on how results will be measured and evaluation used is not sufficient. Key gaps include: - Partners are not required to unpack the underlying logic of their proposed project and explain how activities link to outputs, to outcomes and eventual impact and the assumptions that underpin these relationships. - Partners are not required to describe the overall systems that are in place for results measurement, including their planned use of evaluations. - Partners are not required to present the available evidence base that underpins their intervention design or assess its strength. - The minimum requirements on results measurement outlined in the Grant Management Manual are not consistently understood by staff.	Gaps in the minimum standards required from grant applicants about how they plan to measure results makes it difficult for programme officers to make an informed judgment on the quality of their approach and to ensure evaluability. The basic level of detail required from grant applicants about their approach to results measurement is not sufficient to make a solid assessment of how well a partner has thought through its approach to measuring results. Major omissions include: theory of change, details of the evidence base that supports the intervention and details of its overall results management system, including how they plan to use reviews and/or evaluations. These are all requirements that peer agencies have embedded in their result management systems and sets Norway apart from current good practice. A strengthened focus on results measurement during the preparatory stages of the grant cycle is where there is a need for the most improvement within the Norwegian system.	As part of the planning of grants partners should be required to outline in greater detail how they plan to measure results. The details requested of partners during the planning of grants and indicated on the proforma templates should include as well as goal hierarchy, indicators and baseline, a theory of change, a review of the evidence base that underpins the programme design, and an assessment of its quality, details of their results management system and their plan for the use of reviews and evaluations

Finding	Conclusion	Recommendation
- The number of Grant Scheme Rules and the variation in requirements around goal achievement, quality assurance and evaluation present an inconsistent set of procedures for staff to follow. This creates unnecessary confusion around results measurement. - Norway does not require the use of mandatory reporting templates by partners. This produces high levels of variability in how partners present results frameworks and report on results. This creates inconsistency in quality and undermines compliance with the minimum requirements on results measurement.	Development of grant scheme rules and a new Grant Management Manual could have improved consistency and coherence around results measurement. But the non-mandatory use of templates means there are no clear standards. Templates exist for applications, progress reports and final reports, that reflect these requirements, but their use is optional. The Grant Scheme Rules are supposed to provide more specific requirements around results measurement, however in practice, most are ambiguous and leave the decisions around what to do to the individual programme officer.	- The requirements on technical assistance and QA should be harmonised across all grant scheme rules. The current diversity in results requirements across the grant scheme rules should be harmonised. Rather than have each scheme set its own requirements we suggest using budget thresholds to determine at which point it is mandatory to get technical assistance and QA. Within each budget category the requirements would be the same. - Establish a consistent basis such as size of grant, for which technical assistance and QA is mandatory. As part of the harmonisation of Grant Scheme Rules, there should be a clear requirement for mandatory technical assistance for all projects greater than a certain amount and sample assessment of mid-size grants. We do not think it is feasible to replicate this process for small grants - Partners should be required to use the standard templates that have been developed, rather than using their own approaches. To create greater consistency within the system and to ensure that appropriate level of detail is provided by a partner in their application, progress reports and final reports, the standard templates that currently exist for applications (S01/51), progress reports (S11/S61) and final reports (S21/81) should be made mandatory for partners to use and not just as a checklist. We recognise that some may feel that this recommendation runs counter to the principles of the Paris and Busan declarations; however we feel that the absence of any standardisation is undermining effective result measurement.

Finding	Conclusion	Recommendation
- While results measurement is referenced in numerous policies and guidelines, there is not a coherent body of guidance material that supports staff in the practical task of appraising results frameworks, supporting partners in measuring results and ensuring the evaluability of grants. - There is an absence of guidelines and checklists that support staff in making judgements on whether partner's results information or results frameworks are of 'good enough' quality.	Staff do not have access to appropriate detailed information to implement the basic guidance given in the Grant Management Manual. The current guidance fails to provide practical support. There is no coherent body of practical reference available to staff on either how to appraise partners' results frameworks and reporting, or strengthen their capacity in results measurement. In particular, there is no guidance for staff to make clear judgements on whether something is good enough.	- Guidance should be developed for staff on how to put results into practice, specifically how to appraise results frameworks and support partners in developing effective measurement systems. This guidance should cover: - How to develop and appraise theories of change; - How to identify and assess the quality of evidence that support a programme design; - Planning and managing grant-level evaluations and reviews; - How to appraise the quality of results management systems. - Examples can be found from the comparator agencies but a good start would be to update the 2008 publication on Results Management in Development Cooperation. - Standard QA checklists should be developed for staff to use when appraising results frameworks, progress reports and final reports. These should require staff to make a clear judgement on quality, possibly using a red-amber-green rating system. Good practice examples could draw on QA checklists developed by DFID for the Business Case but would need to be adapted to Norway's grant systems. These checklists would in effect provide a clear specification of what the minimum requirements on results measurement detailed in the Grant Management Manual mean in practice and how staff can make a judgement on what is considered good enough. - Add a rating assessment to the Progress and Final Report templates for partners to complete themselves. Reporting currently makes little use of indicators and does not require any judgement about performance by the grantee or programme officer. This is in contrast to the practice of the agencies reviewed for comparison and contributes to weak evaluation awareness among staff. A rating is something that can be done even where the partner has poor indicators. Agreement or otherwise with the rating would be signified by the grant manager as part of their quality assurance checklist. These data should be entered to a performance database where they would enable comparative analysis across sectors and over time and would provide a strong foundation for analysis to target thematic evaluations.

Finding	Conclusion	Recommendation
• Of the 20 grants reviewed few made use of mid-term or end reviews or evaluations to measure results and /or to learn from experiences. • Of the 20 reviews/evaluations assessed few met basic quality standards. In particular, of the reports reviewed, all failed to provide a robust measurement of change at outcome level. • The resources made available for reviews and evaluations are typically very limited. This may be a contributing factor to their variable quality. • There is limited guidance available to staff on how to design and manage grant level reviews or evaluations. This is likely to be a contributing factor to the low quality of review and evaluations.	• The absence of a clear approach to and guidance on reviews and evaluations means their use within the grant management cycle is fragmented and their quality is variable. There is an absence of any clear guidance on when to commission reviews/evaluations and how best to design and manage them. The decision is for the most part at the discretion of the individual desk officer and partner. The fact that only a small number of the 20 grants we reviewed commissioned evaluations or reviews, and of the reviews and evaluations that are commissioned most failed to meet basic quality standards suggests the current approach is not working (see conclusion below). Reviews and Evaluations present an important source of results information for Norway and are central to learning; there therefore needs to be a more informed and strategic approach to their use at the grant level. This should be grounded in a clear understanding of the intervention being funded and the quality of the evidence base that underpins its design. This should in turn shape the level of investment that is needed in evaluation and the type of evaluation that should be commissioned. The DFID experience is instructive in this regard.	• A more considered and strategic approach to the use of evaluations and reviews at the grant level should be developed. As part of the preparatory phase of a grant, greater consideration should be given to whether reviews and/or evaluations will be commissioned and why, and the budget implications. In order to inform whether and what type of evaluation should be undertaken a review of the evidence base supporting a proposed intervention (and its quality) should be undertaken. This will provide a logical foundation for deciding the type and details of evaluation to be planned and will help ensure a more strategic use of evaluations. Where the evidence is strong a Final Report may suffice, where an intervention is new and innovative and the evidence base is weaker an attribution or contribution evaluation may be necessary. This detail may not be necessary at the initial application stage, but should be undertaken before a contract is signed. • Standard checklists should be developed for quality assuring grant level evaluations and reviews. These should cover both quality at entry (ToR, inception report etc.) and exit (Final evaluation/review report). These checklists would help staff to make clear judgements on quality and more easily identify when a report is of insufficient quality. Examples are available from both the World Bank and DFID.

Finding	Conclusion	Recommendation
- The training courses on results measurement are generally of a good quality with staff satisfaction high. However, there are gaps in content in a number of areas that are central to effective results measurement and evaluability: – There is an absence of training on how to appraise a partner's results measurement systems and plans. – There is insufficient attention given to reviewing the quality of evidence as part of appraising programme design. This is touched upon in the Reviews and Evaluations in Grant Management course, but only to a limited extent. – There is insufficient attention given to how to develop and appraise theories of change. This is touched upon in the Reviews and Evaluations in Grant Management course, but only to a limited extent. - The short duration of the training courses (between a couple of hours and a day) also raises questions about their effectiveness in building appropriate levels of staff capacity in results measurement. - Attendance levels for training courses focused on results measurement are also low and there is a perception among staff that building skills in results measurement neither supports career progression nor is valued by managers. - Outside of the training there is also little other practical guidance available to staff on how to put results measurement into practice.	Courses currently on offer are of a good quality, and cover many of the issues needed to appraise partner's results frameworks and staff who attend are satisfied with them. However, the courses fail to cover key areas for ensuring evaluability (developing theories of change, reviewing the quality of evidence, assessing partner's results measurement systems) and take up of the courses is low. The low take up is likely to result from staff not viewing skills in results measurement as supporting career progression and managers not valuing skills in results measurement. The short duration of all of the courses is also of concern. It is our view that course on results measurement of between 3 hours and 1 day are of insufficient length to develop staff skills to the necessary level for them to undertake their roles effectively. The absence of other supportive practical guidance on how to put results measurement into practice that sits alongside the training is also a weakness.	- Develop a more comprehensive training programme to support staff capacity in results measurement. The training programme should offer more in-depth and longer-term training for those that want to deepen their skills in results measurement. This more in-depth training should be mandatory for Norad advisers that are providing technical advice on results frameworks. The programme should also combine e-learning and self-study modules for staff too busy to attend. In the case of DFID for example, a one week training course on commissioning and managing evaluations is run in collaboration with the UK Evaluation Society. This is available to all evaluation and results advisers. Evaluation and results advisers (in common with other sectoral disciplines) also attend a professional meeting once a year and have a 'Head of Profession' appointed to look after their professional interests. The revised training programme should cover the following issues in greater depth: the development and appraisal of theories of change; how to appraise the quality of evidence that underpins a programme's design; and reviewing and appraising a partner's results measurement systems. - An online resource hub should be developed that provides staff access to examples of good practice in results measurement and pools sector specific resources. - Develop the capacity of grant recipients to measure results such as by e-learning, but perhaps also through a 'partners' guide to managing for results'. Provide outreach to partners to raise their skills in results measurement

Finding	Conclusion	Recommendation
• There is an inconsistent use of technical advice by programme officers to support the appraisal of results frameworks and reporting. A review of 20 grants indicated that in the majority of cases technical advice was not used. • The inconsistent use of technical support results from the fact that the decision whether to request technical input rests with the individual programme officer; only in a limited number of cases is it mandated by Grant Scheme Rules. • This approach to technical support is at odds with the practices in other peer agencies such as DFID and the World Bank. It is where Norwegian practice on results measurement differs the most from comparator organisations. • A review of 20 grants indicated that when technical advice is provided by Norad to MFA and embassies its quality is mixed. The advice frequently lacks depth and fails to provide a clear judgement on whether the results framework or report is of 'good enough' quality. • The technical support provided by AMOR Results Management Section, while helpful, is sometimes viewed as too formalised and lacking the practical guidance programme officers need to effectively support partners in developing robust results frameworks. • The Norad staff responsible for providing technical advice on results frameworks and reporting are often sector specialists and are not given additional support to build their skills in results measurement.	• The institutional arrangements are not set up in a way that ensures consistent quality across the system. The use of technical advice is only mandatory in a minority of grant schemes and is little used, as is formal QA by the AMOR Legal Section. When technical support is provided by Norad Advisers our assessment indicated that it can often lack depth and fail to provide a clear judgment on whether results information is 'good enough.' A reason for this could be that while Norad Advisers bring strong sector skills, very few have specialist expertise in results measurement. In the case of AMOR's Results Section, the technical input it provides is helpful, but sometimes too formalised and lacking the practical guidance programme officers need to effectively support partners in developing robust results frameworks. Lastly, the independent QA by AMOR at embassy level deals only with issues of process compliance, not with content, the quality of results documentation or evaluability of grants.	• The requirements on technical assistance and QA should be harmonised across all grant scheme rules. The current diversity in results requirements across the grant scheme rules should be harmonised. Rather than have each scheme set its own requirements we suggest using budget thresholds to determine at which point it is mandatory to get technical assistance and QA. Within each budget category the requirements would be the same. • Establish a consistent basis such as size of grant, for which technical assistance and QA is mandatory. As part of the harmonisation of Grant Scheme Rules, there should be a clear requirement for mandatory technical assistance for all projects greater than a certain amount and sample assessment of mid-size grants. We do not think it is feasible to replicate this process for small grants Approach 1: Concentrating expertise • Resource AMOR to provide more comprehensive support to all eligible grants on their results measurement frameworks and systems, and evaluation plans and reports at approval and completion. AMOR should be provided with additional resources to conduct quality assurance prior to grant approval and at grant completion of all large projects, over the agreed amount. This could be contracted out or could be staffed as temporary teams of peer reviewers such as were used in the World Bank QAG. As part of this process AMOR could also analyse and validate a sample of grant level reviews and evaluations and feed this into the Annual Results Report. Approach 2: Broadening expertise • Build a cadre of staff specialised in results measurement and evaluation. Norad and the MFA should deepen the skills of its staff in measuring results. We recommend, alongside an improved training programme, that a small number of staff from across the organisations is given intensive training in evaluation and results measurement. This will broaden skills within the organisations and create a cadre of results advisers across the organisation to give local/regional advice and train others. Embedded advisers will offer a more flexible and informal form of support to staff around how to measure results. They will bring results advice closer to the ground. Focal points could specialise in specific sectors. To encourage participation in the scheme an incentive system should be developed that allows for career rewards and development within the cadre.

Finding	Conclusion	Recommendation
Of the 20 grants reviewed: • The majority had poorly-developed results frameworks with particular weaknesses in relation to the clarity of objectives and indicators. • Outcome level reporting, despite it being a minimum requirement for final reports, was almost completely absent. Most progress and final reports dealt with implementation of activities and financial status rather than with results. • While there was evidence that programme officers engaged with partners around the grant application, it was mostly about general issues of programme design, delivery and finance rather than evaluability and results measurement.	• The current rules and guidance on grant management are ineffective at providing staff with the means of ensuring that grants have adequate results frameworks for subsequent evaluation. The majority of grants reviewed do not have sound frameworks of results and would be difficult to evaluate using established methods and criteria. Many of the key steps needed to support evaluability in the design phase are not being taken consistently. The large volume of grants committed every year present a considerable burden for management and the absence of a requirement for a monitoring and evaluation plan leaves little basis to decide about where efforts should be concentrated.	Note: The recommendation is the same as in previous section, page 120.

Finding	Conclusion	Recommendation
- Although the need to demonstrate results is stated by the political leadership of MFA and Norad, many staff believe it has not translated into a clear focus on measuring results by senior management. - Limited time and resources present barriers to staff putting into practice good results management with the pressure to disburse funds often trumping thoughtful consideration of results frameworks - There are no incentives to reward or sanctions to penalise performance in managing for results - The recent trend towards fewer, larger and longer duration grants may create opportunities for staff to spend more time working with partners on results frameworks and reporting.	- Staff are under pressure and in many instances find themselves unable to devote the time and attention necessary to improve specification of results. Staff consider themselves to be adequately competent to manage for results, but with few exceptions they do not put that into practice for an assortment of mutually-reinforcing reasons: pressure of work that restricts the time available for interaction with applicants; and perceptions of a low priority for results by managers, especially in the MFA, despite high-level leadership over the need for results. There are neither incentives nor sanctions to reinforce good practice on managing for results. - The culture in MFA and Norad is permissive towards a grant applicant's proposal rather than ensuring that applications meet with minimum standards before being approved. Inconsistent direction and difficulties in securing technical advice, coupled with the long-established system whereby applicants lead the grant design process all combine to result in a low standard of results specification.	- Strengthen the support at senior management level for results measurement. There needs to be more visible action among the leadership of the MFA and Norad to insist on better results management and evaluability. Changing processes or structures are not enough to change behaviour and culture. Experience with results management systems indicates that senior management incentives and direction to staff are key elements. Strengthening a results-focus in the grant management process will require staff to spend more time with partners, developing and reviewing results frameworks, reporting and evaluations; senior management needs to create the space and resources for this to take place. - Improve staff incentives for measuring results and ensuring evaluability. At present training is not given a high priority by managers and there are no career rewards for expertise in results management. A strong incentive would be to incorporate continuing professional development in results management as a positive career attribute to be recorded on personnel files and factored into career development. This recommendation is addressed to top management in the MFA and Norad. - Design a new approach to reach out to partners, with a combination of improved technical guidance, some direct capacity building and access to the self-study materials. More focus should be put on developing the capacity of partners to improve their measurement and reporting of results.

EVALUATION REPORTS

6.00	Making Government Smaller and More Efficient.The Botswana Case	5.08	Evaluation: Evaluation of the Norwegian Research and Development Activities in Conflict Prevention and Peace-building
7.00	Evaluation of the Norwegian Plan of Action for Nuclear Safety Priorities, Organisation, Implementation	6.08	Evaluation: Evaluation of Norwegian Development Cooperation in the Fisheries Sector
8.00	Evaluation of the Norwegian Mixed Credits Programme	1.09	Evaluation: Joint Evaluation of Nepal's Education for All 2004-2009 Sector Programme
9.00	"Norwegians? Who needs Norwegians?" Explaining the Oslo Back Channel: Norway's Political Past in the Middle East	1.09	Study Report: Global Aid Architecture and the Health Millennium Development Goals
10.00	Taken for Granted? An Evaluation of Norway's Special Grant for the Environment	2.09	Evaluation: Mid-Term Evaluation of the Joint Donor Team in Juba, Sudan
1.01	Evaluation of the Norwegian Human Rights Fund	2.09	Study Report: A synthesis of Evaluations of Environment Assistance by Multilateral Organisations
2.01	Economic Impacts on the Least Developed Countries of the Elimination of Import Tariffs on their Products	3.09	Evaluation: Evaluation of Norwegian Development Cooperation through Norwegian Non-Governmental Organisations in Northern Uganda (2003-2007)
3.01	Evaluation of the Public Support to the Norwegian NGOs Working in Nicaragua 1994-1999	3.09	Study Report: Evaluation of Norwegian Business-related Assistance Sri Lanka Case Study
3A.01	Evaluación del Apoyo Público a las ONGs Noruegas que Trabajan en Nicaragua 1994-1999	4.09	Evaluation: Evaluation of Norwegian Support to the Protection of Cultural Heritage
4.01	The International Monetary Fund and the World Bank Cooperation on Poverty Reduction	4.09	Study Report: Norwegian Environmental Action Plan
5.01	Evaluation of Development Co-operation between Bangladesh and Norway, 1995-2000	5.09	Evaluation: Evaluation of Norwegian Support to Peacebuilding in Haiti 1998-2008
6.01	Can democratisation prevent conflicts? Lessons from sub-Saharan Africa	6.09	Evaluation: Evaluation of the Humanitarian Mine Action Activities of Norwegian People's Aid
7.01	Reconciliation Among Young People in the Balkans An Evaluation of the Post Pessimist Network	7.09	Evaluation: Evaluation of the Norwegian Programme for Development, Research and Education (NUFU) and of Norad's Programme for Master Studies (NOMA)
1.02	Evaluation of the Norwegian Resource Bank for Democracy and Human Rights (NORDEM)	1.10	Evaluation: Evaluation of the Norwegian Centre for Democracy Support 2002-2009
2.02	Evaluation of the International Humanitarian Assistance of the Norwegian Red Cross	2.10	Synthesis Study: Support to Legislatures
3.02	Evaluation of ACOPAM An ILO program for "Cooperative and Organizational Support to Grassroots Initiatives" in Western Africa 1978 - 1999	3.10	Synthesis Main Report: Evaluation of Norwegian Business-related Assistance
3A.02	Évaluation du programme ACOPAM Un programme du BIT sur l'« Appui associatif et coopératif aux Initiatives de Développement à la Base » en Afrique de l'Ouest de 1978 à 1999	4.10	Study: Evaluation of Norwegian Business-related Assistance South Africa Case Study
4.02	Legal Aid Against the Odds Evaluation of the Civil Rights Project (CRP) of the Norwegian Refugee Council in former Yugoslavia	5.10	Study: Evaluation of Norwegian Business-related Assistance Bangladesh Case Study
1.03	Evaluation of the Norwegian Investment Fund for Developing Countries (Norfund)	6.10	Study: Evaluation of Norwegian Business-related Assistance Uganda Case Study
2.03	Evaluation of the Norwegian Education Trust Fund for Africa in the World Bank	7.10	Evaluation: Evaluation of Norwegian Development Cooperation with the Western Balkans
3.03	Evaluerings av Bistandstøtters Evalueringsnettverk	8.10	Evaluation: Evaluation of Transparency International
1.04	Towards Strategic Framework for Peace-building: Getting Their Act Together. Overview Report of the Joint Utstein Study of the Peacebuilding.	9.10	Study: Evaluability Study of Partnership Initiatives
2.04	Norwegian Peace-building policies: Lessons Learnt and Challenges Ahead	10.10	Evaluation: Democracy Support through the United Nations
3.04	Evaluation of CESAR's activities in the Middle East Funded by Norway	11.10	Evaluation: Evaluation of the International Organization for Migration and its Efforts to Combat Human Trafficking
4.04	Evaluerings av ordningen med støtte gjennom paraplyorganisasjoner. Eksempifisert ved støtte til Norsk Misjons Bistandsnemda og Atlas-alliansen	12.10	Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative (NICFI)
5.04	Study of the impact of the work of FORUT in Sri Lanka: Building Civil Society	13.10	Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative. Country Report: Brasil
6.04	Study of the impact of the work of Save the Children Norway in Ethiopia: Building Civil Society	14.10	Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative. Country Report: Democratic Republic of Congo
1.05	-Study: Study of the impact of the work of FORUT in Sri Lanka and Save the Children Norway in Ethiopia: Building Civil Society	15.10	Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative. Country Report: Guyana
1.05	-Evaluation: Evaluation of the Norad Fellowship Programme	16.10	Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative. Country Report: Indonesia
2.05	-Evaluation: Women Can Do It - an evaluation of the WCDI programme in the Western Balkans	17.10	Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative. Country Report: Tanzania
3.05	Gender and Development - a review of evaluation report 1997-2004	18.10	Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative
4.05	Evaluation of the Framework Agreement between the Government of Norway and the United Nations Environment Programme (UNEP)	1.11	Evaluation: Results of Development Cooperation through Norwegian NGO's in East Africa
5.05	Evaluation of the "Strategy for Women and Gender Equality in Development Cooperation (1997-2005)"	2.11	Evaluation: Evaluation of Research on Norwegian Development Assistance
1.06	Inter-Ministerial Cooperation. An Effective Model for Capacity Development?	3.11	Evaluation: Evaluation of the Strategy for Norway's Culture and Sports Cooperation with Countries in the South
2.06	Evaluation of Fredskorpset	4.11	Study: Contextual Choices in Fighting Corruption: Lessons Learned
1.06	- Synthesis Report: Lessons from Evaluations of Women and Gender Equality in Development Cooperation	5.11	Pawns of Peace. Evaluation of Norwegian peace efforts in Sri Lanka, 1997-2009
1.07	Evaluation of the Norwegian Petroleum-Related Assistance	6.11	Joint Evaluation of Support to Anti-Corruption Efforts, 2002-2009
1.07	- Synteserapport: Humanitær innsats ved naturkatastrofer: En syntese av evalueringsfunn	7.11	Evaluation: Evaluation of Norwegian Development Cooperation to Promote Human Rights
1.07	- Study: The Norwegian International Effort against Female Genital Mutilation	8.11	Norway's Trade Related Assistance through Multilateral Organizations: A Synthesis Study
2.07	Evaluation of Norwegian Power-related Assistance	9.11	Activity-Based Financial Flows in UN System: A study of Select UN Organisations Volume 1 Synthesis Volume 2 Case Studies
2.07	- Study Development Cooperation through Norwegian NGOs in South America	10.11	Evaluation of Norwegian Health Sector Support to Botswana
3.07	Evaluation of the Effects of the using M-621 Cargo Trucks in Humanitarian Transport Operations	1.12	Mainstreaming disability in the new development paradigm. Evaluation of Norwegian support to promote the rights of persons with disabilities.
4.07	Evaluation of Norwegian Development Support to Zambia (1991 - 2005)	2.12	Hunting for Per Diem. The uses and Abuses of Travel Compensation in Three Developing Countries
5.07	Evaluation of the Development Cooperation to Norwegian NGOs in Guatemala	3.12	Evaluation of Norwegian Development Cooperation with Afghanistan 2001-2011
1.08	Evaluation: Evaluation of the Norwegian Emergency Preparedness System (NOREPS)	4.12	Evaluation of the Health Results Innovation Trust Fund
1.08	Study: The challenge of Assessing Aid Impact: A review of Norwegian Evaluation Practice	5.12	Real-Time Evaluation of Norway's International Climate and Forest Initiative. Lessons Learned from Support to Civil Society Organisations.
1.08	Synthesis Study: On Best Practice and Innovative Approaches to Capacity Development in Low Income African Countries	6.12	Facing the Resource Curse: Norway's Oil for Development Program
2.08	Evaluation: Joint Evaluation of the Trust Fund for Environmentally and Socially Sustainable Development (TFESSD)	7.12	A Study of Monitoring and Evaluation in Six Norwegian Civil Society Organisations
2.08	Synthesis Study: Cash Transfers Contributing to Social Protection: A Synthesis of Evaluation Findings	8.12	Use of Evaluations in the Norwegian Development Cooperation System
2.08	Study: Anti-Corruption Approaches. A Literature Review	9.12	Evaluation of Norway's Bilateral Agricultural Support to Food Security
3.08	Evaluation: Mid-term Evaluation of the EEA Grants	1.13	A Framework for Analysing Participation in Development
4.08	Evaluation: Evaluation of Norwegian HIV/AIDS Responses	2.13	Local Perceptions, Participation and Accountability in Malawi's Health Sector
		3.13	Evaluation of the Norwegian India Partnership Initiative
		4.13	Evaluation of Five Humanitarian Programmes of the Norwegian Refugee Council (NRC) and of the Standby Rpster NORCAP
		5.13	Real-Time Evaluation of Norway's International Climate and Forest Initiative Contribution to Measurement, Reporting and Verification

Norad

Norwegian Agency for
Development Cooperation

Postal address
P.O. Box 8034 Dep. NO-0030 OSLO
Visiting address
Ruseløkkveien 26, Oslo, Norway

Tel: +47 22 24 20 30
Fax: +47 22 24 20 31

postmottak@norad.no
www.norad.no