

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: 978-82-7548-790-0

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

Table of contents

Annex 4 Data and instruments 2

Annex 5 Comparator agency study 49

Annex 6 Methodology and analytical framework 68

Annex 4 Data and instruments

This annex contains 11 sections:

1.	Process mapping of the grant management cycle	3
2.	Quality assessment of policies and guidelines.....	7
3.	List of sampled grants that were reviewed	8
4.	List of key documents consulted during the grant review	9
5.	Checklist for assessing grant evaluability.....	15
6.	Staff survey	27
7.	Focus group discussion outline	34
8.	Quality checklist for EVAL report	36
9.	Evaluator competencies survey.....	40
10.	Summary assessment of Grant Scheme Rules results' requirements	45
11.	Grant commitment tables	47

1. Process mapping of the grant management cycle

The following two figures outline the key steps in the grant management cycle. The first identifies the specific activities that are required during each phase (preparatory, follow-up and completion), the second specifies the requirements around results measurement for each phase. The colours represent whether something is mandatory or not.

 If appropriate

 Mandatory according to some grant scheme rules or in cases referred to in the activity description

 Mandatory

Key activities in the grant management cycle

Grant Management Regime	PREPATORY PHASE			FOLLOW-UP PHASE	COMPLETION PHASE
	Receiving the application	Assessment and decision			
		Assessment	Decision		
I Project and programme support	Hold meeting A11, S31				
	Receive expert guidance A02				
	Assess the need for more information A01, S01	Receive external appraisal A03	Prepare and sign agreements A06, V06	Make disbursement/issue debit authorisation A08, A09, V05	Receive payment of unused funds, A18, V05
				Receive confirmation of receipt, A10, V05	Receive process and final report, A19, V04, report from S21
		Receive and process financial report A12, V05, Report from S11		Prepare completion letter, A20	
		Receive and process audit reports A13, V05			
		Receive and process other plans/report A15			
		Receive and process progress report, V04, S11		Complete PTA agreement, A21	
		Conduct review A16			
		Conduct project/field visits A17			
II General grants	Hold meeting A11, S31				
	Receive expert guidance A02				
	Assess the need for more	Receive external appraisal A03	Prepare and sign agreements A06, V06	Make disbursement/issue debit authorisation A08, A09, V05	Receive payment of unused funds, A18
				Receive confirmation of receipt, A10, V05	

	information A01, S01		(Legal quality assurance is mandatory in many cases)	Receive and process financial report A12, V05, Report from S11	Prepare completion letter
		Prepare decision documentation A04, S03, S05, V04, V05		Receive and process audit reports A13, V05	
				Receive and process other plans/report A15, V04, V05	
				Conduct review A16	Complete PTA agreement, A21
		Conduct project/field visits A17, S32			
IV Direct disbursements				Hold meeting A11, S31	Complete PTA agreement, A21
				Make disbursement A08	
				Receive confirmation of receipt, A10	
				Receive and process financial reports A13, V05	

Results requirements at each step of the grant management cycle

Grant Management Regime	PREPATORY PHASE			FOLLOW-UP PHASE	COMPLETION PHASE
	Receiving the application	Assessment and decision			
		Assessment	Decision		
I Project and programme support	Hold meeting A11, S31				
	<ul style="list-style-type: none">Part of the follow-up and control functionFollow-up and control should focus on the efficiency of and quality of implementation and on results achievementContent and frequency of meetings should be specified in the agreement				
	Receive expert guidance A02 (not evaluation specific)				
	<ul style="list-style-type: none">Can be used in preparing for and follow-up grantsMandatory in some cases specified by GSRs, and may be appropriate in other cases where the unit does not itself have sufficient expertise				
	Assess the need for more information A01 - Minimum content in application defined by -Additional requirements in GSRs S01 - Description of a) current situation (baseline) b) activities c) milestones (if used) d) project goals at different levels (goal hierarchy) e) planned/expected goals f) indicators	Receive external appraisal A03	Prepare and sign agreements A06 - Highly recommended to use agreement templates provided by the MFA - In preparing an agreement, the need for legal quality assurance (QA) must be clarified (see GSR and V06) - QA is mandatory in some GSR V06	Make disbursement/issue debit authorisation A08, A09, V05 See prepare decision documentation cell	Receive payment of unused funds A18 V05 See prepare decision documentation cell
		Prepare decision documentation A04, S03, V04 - Grant manager is responsible for the collection, assessment and follow-up on results - Describes key concepts in results management a) results: output, outcome, impact		Receive confirmation of receipt A10, V05 Receive and process financial report A12 To control expenditure, progress and compliance with agreement GMR I: Presented in the form of project accounts with corresponding notes V05 See prepare decision documentation cell Report form S11	Receive process and final report A19 Recipient own presentation and assessment of the project's results Should contain: a) presentation and assessment of outputs and outcome compared

	Refers to 'Results management in Norwegian Development Cooperation'	<p>b) measuring results: baseline, indicators</p> <p>c) choice of follow-up measures: min. described in GSR</p> <p>d) documentation of outcome and probable impact should be part of final report</p> <p>e) progress/final report must be compared to the described goal hierarchy, baseline values and indicators.</p> <p>The output's effect on the target group – outcome – should be measured and compared to baseline. Other follow-up measures (field visits/review) are described in relevant activity descriptions</p> <p>V05</p> <p>Financial management is an essential aspect of the grant manager's obligation process to assess, follow up, and control that funds have been used efficiently, that result reports, financial accounts, and other information account satisfactorily for the project</p>	Guide on how to enter and follow up on grant agreements	<p>Receive and process audit reports</p> <p>V05</p> <p>See prepare decision documentation cell</p>	with the goal hierarchy and estimation of effect on society (impact)
				<p>Receive and process progress report</p> <p>A14</p> <p>Mandatory in connection with multiannual agreements</p> <p>V04</p> <p>See prepare decision documentation cell</p> <p>S11</p>	
				<p>Receive and process other plans/report</p> <p>A15</p> <p>Often required for general budget support and sector support</p>	
				<p>Conduct review</p> <p>A16</p> <p>Reviews are part of the formalised follow-up of a project or programme (mid-term review), as well as a final process to assess the effect (outcome, impact) of the project programme (end review). A review is a thorough assessment with focus on implementation and follow-up of plans. (...) A review is initiated by one or more parties to an agreement and, when mandatory, this must be stipulated in the agreement. The applicable grant scheme rules may specify requirements for reviews.</p> <p>V04</p> <p>See prepare decision documentation cell</p>	<p>Prepare completion letter</p> <p>A20</p>
				<p>Conduct project/field visits</p> <p>A17</p> <p>Part of the follow-up and represents a useful control measure.</p> <p>A project/field visit can be used for assessment and control of a project's or programme's progress and delivered outputs in relation to the reports submitted by the grant recipient. The findings shall be described in a report, and/or discussed in formal meetings or followed up in some other way. A project/field visit may be conducted at any time by the unit responsible or by a review team of other internal or external experts. A field visit may be an important measure to detect and prevent misuse of funds, and unannounced field visits should be considered.S32</p>	<p>Complete PTA agreement,</p> <p>A21</p>

II General grants	Hold meeting A11, S31 (Same as GMR I)				
	Receive expert guidance A02 (Same GMR I)				
	Assess the need for more information A01, S01 (Same as GMR I)	Receive external appraisal A03	Prepare and sign agreements A06 (Same as GMR I) V06 (Legal QA is mandatory in many cases)	Make disbursement/issue debit authorisation A08, A09, V05 See prepare decision documentation cell	Receive payment of unused funds A18, V05 (See prepare decision documentation cell)
		Prepare decision documentation A04, S03, S05, V04, V05 (Same as GMR I)		Receive confirmation of receipt A10, V05	
				Receive and process financial report A12 GMR II: Normally in the form of annual report and accounts Report from S11	
				Receive and process audit reports Mandatory in connection with multiannual agreements V05 See prepare decision documentation cell	Prepare completion letter A20
				Receive and process other plans/report A15 Often required V04, V05 See prepare decision documentation cell	
				Conduct review A16, V04 (same as GMR I)	Complete PTA agreement A21
		Conduct project/field visits A17S32 Same as GMR I)			
		Hold meeting A11, S31			
IV Direct disbursements	<ul style="list-style-type: none">Can be used in preparing for and follow-up grantsMandatory in some cases specified by GMRs, and may be appropriate in other cases where the unit does not itself have sufficient expertise				
		Make disbursement A08	Receive payment of unused funds A18, V05 See prepare decision documentation cell		
		Receive confirmation of receipt A10	Receive process and final report, A19. V04 (Same as GMR I)		
		Receive and process financial reports A13, V05	Prepare completion letter A20		
			Complete PTA agreement A21		

2. Quality assessment of policies and guidelines

	PREPATORY PHASE Step 1. Designing the intervention					FOLLOW UP Step 2. Implementing the intervention					COMPLETION Step 3. Conducting a review or evaluation			
	Developing a theory of change	Defining targets/ goals and indicators	Reviewing existing evidence	Planning and budgeting for reviews and evaluations	Seeking technical advice on how to design interventions to improve evaluability	Monitoring the intervention	Reporting on the progress of the intervention	Designing the ToR for a mid-term review	Commissioning a mid-term review	Seeking technical advice on designing a mid-term review	Designing a ToR for an end-term review or evaluation	Commissioning an end-term review or evaluation	Seeking technical advice on an end-term review or evaluation	Managing and quality assuring the final report
Grant management manual (0.0)	V04	V04		A16, A17	A02		A14		A16, A17	A02		A16, A17		A19
Development cooperation manual (0.4)		p. 14		p. 14, 32-34		p. 19, 29-30	p. 22	p. 33, 34	p. 32		p. 34	p. 32		p. 37-39
Results management in development cooperation (0.5)	p. 10-11, 13, 21	p. 14-15	p. 23	p. 16, 28		p. 15-16	p. 26	p.28			p. 28			
The agreement manual (0.7)		p. 16, templates		p. 16, templates		p. 16, templates	p. 16, templates		templates			templates		templates
Establishing and managing Grant Schemes (6.3)		p.14-15		p. 17, refers to SSØ		p. 16	p. 16							
Management of Grants by the MFA and Embassies (6.4/6.1)						p. 18, 53-55, 75	p. 18, 53-55, 75							
Managing delegated Grant Schemes (6.2)						p. 52, 55, 73	p. 52, 55, 73							p. 59
The logical framework approach (8.3)	Whole document	Whole document				p. 88	p. 88				p. 89-90			
Evaluering av statlige tilskuddsordninger (SSØ)	p. 18-24	p.25-26		p. 13-18										

3. List of sampled grants that were reviewed

No	Grant	Unit	budget (NOK mill)	Start year	Planned end year
08-1082267	Liberia CIVPOL	Embassy	24.99	2008	2008
08-1084031	UNICEF/Basic Education. Gender	MFA	43	2008	2008
SDN-07/055	Nuba Mountains Integrated Development	Norad	25.5	2008	2011
NFD0511-308	EXIM Bank (Tanzania)	Norfund	33.25	2008	2012
GLO-07/387-2	Integrated Programme for Xingu	Norad	11.64	2008	2012
LKA-08/075	IFC Enterprise Development	Embassy	15.09	2009	2012
GLO-09/993	FTI Secretariat	MFA	10	2009	2010
PAL-09-025	Augusta Victoria Hospital	MFA	11	2009	2011
MDA-09/003	Rule of law NORLAM	MFA	42.65	2009	2012
GLO-09/853	Disaster Risk Reduction and Preparedness	MFA	73	2009	2011
GLO-08/449-24	African Rift Valley Lakes	Norad	17.9	2009	2011
QZA-09/220	Pilot Programme for Climate Resilience	Norad	45	2009	2010
NPL-10/0070	rehabilitation of Maoist Combatants	Embassy	12.44	2010	2011
QZA-10/0485	Climate and Forest Initiative 2009	Norad	11.49	2010	2012
CIV-11/0005	IDPs in Ivory Coast	MFA	10	2011	2011
MEU-11/0077	EBRD Southern and Eastern Mediterranean	MFA	16	2011	2012
QZA-11-1031	WFP 2011 + Sudan	MFA	43	2011	2012
QZA-12/0140	DREF. IFRC Disaster	MFA	40	2011	2011
QZA-11-0333	Slum Dwellers International	MFA	19.2	2011	2012
RAF-11/0021-5	South Sudan. ICRC	MFA	24.99	2011	2011

4. List of key documents consulted during the grant review

Key documents reviewed as part of grant assessments

The following section provides a list of the key documents that were reviewed as part of our assessment of the 20 grants. Additional documents were consulted; however, what is listed below reflect those that were relevant for the purposes of this evaluation.

08-1082267 – Liberia CIVPOL

- Grant decision, report
- Grant decision, contract
- Progress report, July 2006
- Progress report, December 2006
- Progress report, 2007
- Progress report, 2008
- Progress report, 2009
- Final report, 2009

LKA-08/075 – IFC Enterprise Development

- Grant application, 2009
- Grant decision, contract, 31 Aug 2009
- Grant decision, 24 June 2009
- Progress report, 2009
- Progress report, 2009–2010
- Progress report, 2010–2011
- Progress report, 2011–2012
- MTR report, 2011

MDA-09/003 – Rule of Law NORLAM

- Signed MoU NORLAM – Memorandum of Understanding, May 2007
- Grant decision – contract, 2010
- Debit authorisation to Ministry of Justice for 2007, 05.2007
- Debit authorisation to Ministry of Justice for 2009, 08.2009
- Debit authorisation to Ministry of Justice for 2008, 11.2008
- Debit authorisation to Ministry of Justice for 2010, 06.2010
- Progress report, 2007 (2007 Activity report, 07/01582–21)
- Progress report, 2008 (Quarterly report, January 2008)
- Progress report, 2008 (Quarterly report, 07/01582–25 June 2008)
- Progress report, 2008 (Quarterly report, 07/01582–33 October 2008)
- Progress report, 2008 (2008 Annual report, 07/01582–35)
- Progress report, 2009 (2009 Annual report, 07/01582–47)
- Progress report, 2009 (2009 Exit plan)
- Final report, 2008 (Final report on improvement of probation 2008, 07/01582–2606, June 2008)
- 2007 record of a review meeting and agenda

RAF-11/0021-5 – South Sudan ICRC

- Grant app – report (Application for Financial Support, 18 February 2011)
- Grant decision – contract add (ICRC Africa operations – ‘RAF-11/0021 ICRC Africa operations’ 11/05446-5, 23 March 2011)
- Final report – correspond 18 February 2013
- Final report – correspond (Norwegian RC Annual report 2012 – ‘The Norwegian Red Cross Annual report ICRC Emergency Appeal’, QZA-12/0016, 29 August 2013)

GLO-07/387-2 – Integrated Programme for Xingu

- Grant decision – report (Programdokument Regnskogfondet 2008–2012.pdf (PROGRAMDOKUMENT FLERÅRIG SØKNAD OM SAMARBEIDSAVTALE)
- Grant decision – report (Revidert Programdokument (...) m budsjett.pdf (PROGRAMDOKUMENT FLERÅRIG SØKNAD OM SAMARBEIDSAVTALE)
- Budget: 'Programdokument Regnskogfondet 2008–2012 Budsjett'
- Grant decision – correspond (Vurdering av søknad fra Regnskogsfondet for 2008-DRC.pdf «Samarbeidsavtale mellom Norad og Regnskogsfondet»)
- Tech advice (Bestilling av Faglige råd fra MIE og FLID.pdf (GLO-0850 GLO-07/387 Bestilling: Faglig vurdering (...) 0701528-3, 19Oct2007)
- Progress report – report [2012] (SIVSA 2012 anbef (10) GLO-07-387 (1).pdf 'Forvaltningsgjennomgang SIVSA 2012 – Hovedanbefalinger for oppfølging»
- Evaluation report ('Evaluation of the Xingu programme in Amazon Brazil', June2010)

NFD0511 – 308 – EXIM Bank Norfund

- Grant decision – contract add (Notat.docx 'Notat – Sammendrag AV beslutningsprosesser og rapportering')
- Grant decision – report (Final approval D 0511 EXIM Bank main document (2).pdf 'Final approval D 0511 EXIM Bank (Tanzania) Ltd', 30 March 2006)
- Grant decision – contract ('Subordinated Loan Agreement', 16 October 2008)
- Evaluation – report (120529 FI Review.pdf 'Financial Institutions Annual Review', May 2012)
- Progress report – report (FI lessons learned...).pdf 'Memo', 8 May 2012
- Progress report – report (Norfund 2010 report ESG.pdf 'Form of Annual Monitoring Report', June 2011)

QZA-11-1031 – World Food Programme 2011 + Sudan

- WFP, 2012: WFP's emergency operations in Sudan
- MFA, 2011: QZA-11-1031. Norwegian support to WFP's emergency operations and protracted relief and recovery operations 2011/12, including Sudan and Zimbabwe. Hoveddokument

MEU-11-0077 – EBRD Mediterranean

- EBRD, 2012: 2012-11-28_SEMED_Fiche_ISE Morocco Energy Efficiency Capacity Building
- EBRD: Meeting of the Assembly of Contributors to the Southern and Eastern Mediterranean Multi-donor Account, 2013 (1300268-21-Mote I London 26.06.2013/EBRD bakgrunnsrapport)
- MFA, 2012: 1300268-4-Beslutningsnotat
- EBRD, 2013: 2013-04-24_SEMED_Fiche secured transactions reforms in Morocco
- EBRD, 2013: 2012-12-07_SEMED MDA Assembly
- EBRD, 2012: Southern and Eastern Mediterranean Multi-donor Account. Semi-annual Narrative Report for the period ended 31 December 2012
- MFA, 2012: SEMED Multi-donor account project for approval (1300268-17-Vurdering)

QZA-12/0140 – DREF IFRC

- IFRC, 2012: update on use of DREF fund during 2012 (1201740-5-DREF update)
- Red Cross, 2013: Letter dated 26 June 2013 (1201740-8-FRC Appeal 2012 DREF)

GLO-09/993 – FTI Secretariat

- EFA-FTI, 2009: Catalytic Fund: resource mobilisation background note, June 2009
- EFA-FTI, 2009: EFA-FTI: input for the redesign of EPDF, February, 2009
- EFA-FTI, 2009: EFA-FTI Catalytic Fund: resource mobilisation background note, June 2009
- EFA-FTI, 2009: FTI Catalytic Fund: quarterly financial update, July 2009
- EFA-FTI, 2009: External quality reviews – background and summary findings – Rwanda, March 2009

CIV-11/0005 – IDPs in Ivory Coast

- NRC, undated – Project proposal to the Norwegian Ministry of Foreign Affairs (163953 – CIFM 1101 Proposal and LFA to NMFA)
- MFA, undated: NRC 149272 (1112614-8-Final Report)

SDN-07/055 – Nuba Mountain Integrated Development Programme

- Søknad om midler fra overgangsbistands -ordningen til kirkens nødhjelps integrert utviklingsprogram i nubafjellene – 3 årig programsøknad (November 2008 until November 2011): prosjektdokument 'Nuba Mountain Integrated Development Programme'; budsjett 2008–2011; conflict sensitivity analysis
- Bevilgningsdokument integrert rehabiliterings-og utviklingsprogram i sør-kordofan 2008–2011
- Contract between the Norwegian Ministry of Foreign Affairs (MFA) and Norwegian Church Aid (NCA) regarding Nuba Mountain Integrated Development Programme (December 2008)
- SDN-2019 sdn-07/055 kirkens nødhjelps integrerte utviklingsprogram i nubafjellene – 3-årig program (15 November 2008–15 November 2011)
- Kirkens ødhjelpsintegrerte utviklingsprogram i nubafjellen 20 September 2011. søknad om 6 måneders forlengelse av prosjektperiode og godkjenning av revidert budsjett 2011–2012
- Sdn 07/055 – Nuba Mountains Integrated Programme anmodning om 'no cost extension'
- Oversendelse av statusrapport og anmodning om utbetaling (audit and annual progress report for 2008/09, plus expenditure spreadsheet. Report follows contract specification for content)
- Back to office – Sudanreise 2008
- Part of progress report 2009 (no title page)
- NCA annual activity plan and report in logframe matrix 2010
- 2010 annual progress report for SDN 2019 SDN 07/055: Nuba Mountain Integrated Development Programme 2009–2011
- NCA Khartoum, project notes 2010
- Progress reports and financial statements for 2010
- Kirkens ødhjelpsintegrerte utviklingsprogram i nubafjellen 2009–2011. Arbeidsplan og budsjett for 2011
- Email sdn-07/055 godkjent arbeidsplan og budsjett 2011
- Beslutningsdokument – 3-årig programsøknad 11.2008–11.2011 brief background of project
- Brief assessment of intervention
- Describes a field visit by Norad's executive officer (elin eikeland) in november/december 2008, and refers to a field visit by a larger delegation in march 2006.
- Email godkjenning av revidert budsjett: SDN 07/055 Nuba Mountains Integrated Development 2011
- 2011 Annual Progress Report for SDN 2019 SDN 07/055: Nuba Mountain Integrated Development Programme 2009–2011
- Overforingsdokument: Nuba Mountains Integrated Development Programme 2008–
- Email oversendes Norad for behandling – ref. overforingsnotat
- Sdn-2019 sdn-07/055 kn godkjenning av rapport og regnskap 2011
- Mid-term Review of Norwegian Church Aid's Integrated Development Programme in Nuba Mountains October 2010
- ToR and covering email for MTR
- Email about the MTR and follow-up
- Final report, 2012

GLO-08/449-24 – African Rift Valley Lakes Priority Place

- Grant app – strategy – 2009-11.pdf
- Grant app – report – submission.pdf
- Grant app – report.pdf
- Grant app (Norwegian) – correspond – 2009.pdf
- Grant decision (Norwegian) – report.pdf
- Grant decision (Norwegian) – correspond – 031208.pdf
- Progress report – report – 2009, 2010.pdf
- Progress report – correspond – 310809.pdf
- Progress report (part Norwegian) – correspond – 2010.pdf

- Progress report – correspond – 011110 – annual plan.pdf
- Progress report – correspond – 180811.pdf
- Progress report – 2011 – annual plan.pdf
- Monitoring – correspond – no date.pdf
- Monitoring – correspond – annual meeting – 160610.pdf
- Monitoring – correspond – 170910.pdf
- Monitoring – correspond – 151110.pdf
- Monitoring – correspond – 100211.pdf
- Monitoring – correspond – 090710.pdf
- Monitoring – correspond – 020211.pdf
- Monitoring – annual meeting – 2009.pdf
- Monitoring – annual meeting minutes – 2009.pdf
- Monitoring – annual meeting 2010.pdf
- Monitoring – annual meeting minutes – 291010.pdf
- Follow-up contract (Norwegian) – correspond – 120410 – funding extension.pdf
- Follow-up contract (Norwegian) – 090211 – funding extension.pdf
- Follow-up (Norwegian) – gender policy correspond – 020710.pdf
- Follow-up – gender policy – 2010.pdf
- Follow-up – strategy – 2011 – strat objectives.xlsx
- Follow-up – strategy – 2009-11- logframe analysis.pdf
- Follow-up – strategy – 2012-16.pdf
- Final report – report – 2011.pdf
- Final report – correspond – 310512.pdf

PAL-09-025 – Augusta Viktoria Hospital

- Final report for year 2009 Strategic Cancer Care Initiative Augusta
- Årsrapport 2010, revidert budsjett 2011. Norwegian Church Aid
- Årsplan og budsjett for 2011 – PAL-09025
- Støtte 2011 – Samarbeidsavtale UD og Kirkens Nødhjelp Augusta
- Arsregnskap og årsberetning til Brønnøysundregistrene – Augusta
- Final Report Augusta Victoria Hospital
- Approved final report with remarks – Augusta Viktoria Hospital

QZA-11-0333 – Slum Dwellers International (SDI)

- Oppsummering fra SDI Building cities through partnership 11 November 2010 (Summary from SDI conference 'Building cities through partnership')
- SDI proposal – first draft (27-page proposal for the 'Seven Cities Project Phase II' covering letter and draft text. No milestones or detailed budget included)
- Questions to the grant proposal 2011–2013
- Further Questions to the grant proposal 2011–2013
- Reply to questions to the grant proposal 2011–2013
- Contract 2011 – support to the Seven Cities Project phase II (draft contract for two years funding 2011 and 2012. No reporting requirements)
- Beslutningsnotat – norsk støtte til SDI phase II (decision memorandum – Norwegian support for SDI, phase II 2011–2012)
- Signed contract 2011 – support to the Seven Cities Project phase II
- Minutes – SDI Board of Governors 11 November 2011 – Cape Town
- SDI progress report (2011)
- SDI dispatch 2011 UPFI annual report
- Copy of the SDI annual report 31.03.2012
- Progress report – SDI basket funding report (2012)
- Minutes of the Board of Governors' meeting (Sept 2012)
- Oppsummering – SDI – Board of Governors' meeting (comments)
- UPFI report 2011
- Beslutningsnotat – SDI – (decision memorandum – Norwegian support for Slum Dwellers International, 2012)
- Proposal support The Transformation of Cities in the Global South

- Signed contract 05.12.2012 support to Uniting and Empowering Urban Poor Women
- SDI annual progress report QZA-120926 (2013)

08-1084031 – UNICEF basic education/gender

- Programme agreement, 2007
- Programme agreement, 2008/09
- Programme agreement, 2010/11
- Comments on 2007 programme agreement
- Memo, to minister of development from department for UN, peace and humanitarian questions, on 25 September 2008
- Thematic report 2006/07
- Thematic report 2007/08
- Thematic report 2008
- Comments on 2006/07 thematic report
- Comments on UNICEF programme, 2007
- Comments on 2007/08 thematic report
- Agenda for annual review meeting on basic education and gender equality, 2007
- Annual review meeting minutes, 2007
- UNICEF 2006/07 Medium Term Strategic Plan
- MFA review of its support to UNICEF, 2011.
<http://www.regjeringen.no/upload/UD/Vedlegg/FN/profilark2013/Profilark2013-eng/UNICEF.pdf>
- UNICEF Executive Board minutes compendium of decisions, 2013

GLO-09/853 Norwegian Red Cross – Disaster risk reduction and preparation

- Annual report, 2009
- Interim report, 2009
- Comments on annual report, 2009
- Annual report 2010
- Interim report 2010
- Comments on 2010/11 annual report
- Agreed Programme summary, 2010
- Three year cooperation agreement on disaster risk reduction and preparedness: programme portfolio and budget overview for 2009
- Framework document: disaster risk reduction and preparedness 2009–2011
- Regional disaster preparedness and disaster risk reduction
- Programme Agreement Final report
- Framework agreement about support for crisis prevention through the Norwegian Red Cross – juridical quality assurance
- Programme evaluation and lessons learned, June 2012
- Management response to evaluation

QZA-09/220 – Pilot programme for climate resilience (PPCR)

- Programme contract
- PPCR Disbursement Report, June 30, 2011
- PPCR Donor Call Agenda call notes
- PPCR 3 Semi-Annual Operational Report
- PPCR 9 Semi-Annual Operational Report
- Revised PPCR Results Framework
- Beslutningsnotat støtte til PPCR for 2012
- Bestilling_ innspill til møte i underkomiteen til **MDBs** PPCR
- Faglig vurdering av CIF dokumenter for Istanbul møte, Norad
- World Bank Trust Fund for PPCR – juridical quality assurance of agreement

NPL-10/0070 – Rehabilitation of Maoist combatants

- Decision document, 2011

- Donation agreement, 2010
- Decision document, 2011
- Implementation agreement
- Key findings of the opportunity mapping for rehabilitation of Maoist army combatants
- Rapid assessment of potential labour market rehabilitation of Maoist army combatants
- Report for field trip, 24–25 June 2011
- Revised proposal, Support activities in Maoist army combatants and the adjoining communities, Nov 2012

QZA-10/0485 – Climate and forest initiative

- Annual report, 2009
- Annual report, 2010/11
- Logframe, year 1
- Norad appropriation document, 2010
- Comments on annual report, 2009
- Comments on annual report, 2010/11
- Project summaries and budgets Peru and Madagascar
- Logframe Peru project
- Logframe Madagascar project
- Grant application
- Contract 2010
- Justification for extension of the project
- Progress report 2012
- Progress report 2011

5. Checklist for assessing grant evaluability

Framework for assessing the evaluability of Norad/MFA grants					
		Excellent	Good	Satisfactory	Poor
Justification for score					
Preparatory phase					
Clear design	Did the programme have a clearly defined problem that it aimed to change? Did the proposed programme activities lead to goals and objectives? Were the objectives clear, realistic and measurable?				
Programme logic	Did the programme have a clear theory of change/logic model? Did the outputs, outcomes and goal follow results chain logic? Were assumptions unpacked?				
Indicators	Did the programme identify SMART ¹ indicators on key areas of intervention?				
Justification for design	Was the available evidence reviewed to shape and inform the programme design?				
Resources for monitoring	Were adequate resources allocated in the budget to support the proposed data collection, analysis and evaluation?				
Monitoring systems	Did the programme have a monitoring system in place to gather the information with defined responsibilities, sources and periodicity?				
Baseline data	Was appropriate baseline data collected?				
Technical advice	Did the grant manager/programme officer request support from Norad or external consultants to assess the evaluability of the grant? What was the quality of the advice? How satisfied were they with the support that was provided? [Development Cooperation Manual: Technical advice from Norad is mandatory for interventions greater than NOK 15 mill]				

¹ Indicators that meet a quality standard to be Specific; Measurable; Attainable; Relevant; and Timely

Framework for assessing the evaluability of Norad/MFA grants						
		Excellent	Good	Satisfactory	Poor	Justification for score
Changes to intervention based on programme officer feedback	Is there evidence of changes between the original submission and the version put forward for approval in any of the following areas: a) Project activity and work planning b) Project budgeting and financial management and reporting c) Compliance with policies and strategies of Norwegian aid d) Setting a baseline and defining results and indicators e) Ensuring projects can be evaluated	Extensive	Moderate	Minimal	None	
Interaction between programme officer and grantee	Is there a record of interaction between the grant manager/programme officer and the applicant concerning any of the following: a) Specification of outputs delivered b) Financial management and reporting c) Comparison between planned and actual indicator targets d) Actual experience compared with the intervention logic/theory of change e) Measurement of progress towards outcomes	Extensive	Moderate	Minimal	None	
Follow-up phase						
Ongoing monitoring and reporting	Did the grantee provide monitoring data at output and outcome level in its progress reports?					
ToR for a mid-term review	Was the ToR for the MTR focused and achievable given the available budget for the review?					
MTR report	Quality of mid-term review [use Quality Framework as in EVAL assessment to review reports] [Development Cooperation Manual: Mid-term are mandatory for interventions areater than NOK 50m]					

Framework for assessing the evaluability of Norad/MFA grants						
		Excellent	Good	Satisfactory	Poor	Justification for score
Managing and quality assuring the MTR	Did the programme officer Quality Assure the MTR?					
Interaction between programme officer and grantee	Is there a record of interaction between the grant manager/programme officer and the grant recipient concerning analysis and reporting of any of the following: a) Specification of outputs delivered b) Financial management and reporting c) Comparison between planned and actual indicator targets d) Actual experience compared with the intervention logic/theory of change e) Measurement of progress towards outcomes	Extensive	Moderate	Minimal	None	
Technical advice	Was external and/or Norad support requested in the design, management and/or QA of the MTR? What was the quality of the advice? How satisfied were they with the support that was provided? <i>Development Cooperation Manual: Technical advice from Norad is mandatory for interventions greater than NOK 15m]</i>					
Completion phase						
End reports from grantee	Did the grantee provide high-quality monitoring data at output and outcome level in its final report?					
Designing a ToR for an end-term review or evaluation	Were the ToRs focused and achievable given the available budget for the review? <i>[Development Cooperation Manual: mid-term are mandatory for interventions greater than NOK 50m]</i>					
Commissioning an end-term review or evaluation	Quality of final review/evaluation [use Quality framework used for EVAL assessment to review report]					

Framework for assessing the evaluability of Norad/MFA grants

		Excellent	Good	Satisfactory	Poor	Justification for score
Managing and quality assuring the final report	Did the programme officer provide support in the management of the final review and quality assure the final report?					
Technical assistance	<p>Was external and/or Norad support requested is the design, management and/or QA of the MTR? What was the quality of the advice? How satisfied were they with the support that was provided?</p> <p><i>[Development Cooperation Manual: Technical advice from Norad is mandatory for interventions greater than NOK 15 mill]</i></p>					
Interaction between programme officer and grantee	<p>Is there a record of interaction between the grant manager/programme officer and the grant recipient concerning analysis and reporting of any of the following:</p> <ul style="list-style-type: none"> a) Specification of outputs delivered b) Financial management and reporting c) Comparison between planned and actual indicator targets d) Actual experience compared with the intervention logic/theory of change e) Measurement of progress towards outcomes 	Extensive	Moderate	Minimal	None	

Quality framework for mid-term and final reviews/evaluations

Our framework for assessing whether the reports provides information on the degree of evaluability of the grant applies four criteria. Based on the information in the report, would it have been possible to:

- identify a credible counterfactual and measure the results against this counterfactual;
- establish the theory of causal links between the intervention and its expected results;
- find data suitable for measuring the results;
- document the reliability of the evaluation and validate the conclusions?

Inference: sometimes the answer to these questions would be negative, especially when little resources have been made available for the consultants. A negative answer does not necessarily provide any conclusions to whether the grant can be evaluated (e.g. there is no information about any data in the report => the data could still be available in a manner that would facilitate a good evaluation of the grant). On the other hand, a positive answer does provide clear indications about the grant's evaluability (e.g. the report contains information about data that is available for measuring results => increases the likelihood that the grant can be evaluated in a good way).

Within each of these quality criteria we have identified a number of sub-criteria and for each of these we have developed a statement that describes appropriate practice. Content has been taken from both the OECD-DAC evaluation standards relevant for our assignment and from the research literature on programme evaluations (see Norad 2008 <http://www.norad.no/no/evaluering/publikasjoner/publikasjon?key=109790>). Each evaluation report will be assessed against each of the quality criteria using a four-point rating scale (see below). In order to average across sub-criteria we score each rating. The scoring evidence will be provided alongside, justifying why a particular rating has been given.

Rating	Score	Quality
Poor	0	Quality deficient in important respects
Sufficient	1	Sufficient for informing about results
Good	2	Fully satisfactory quality
Very good	3	Best-practice model, suitable for emulation

Mid-term review

Key quality criteria	Quality sub-criteria	Rating	Implications for evaluability of the grant
Quality criterion 1 Identifies a credible counterfactual and measures the results against this counterfactual			
1.1 Explaining review methodology	The report describes and explains the review methodology and its application for identifying the results of the intervention. The methodology includes specification and justification of the design of the empirical verification approach and the techniques for data analysis and estimation.		
1.2 Identifies the output that stems from the intervention	The review makes a clear distinction between the output that can be attributed to the intervention from that of other contributors.		
1.3 Identifies the counterfactual	Implicitly or explicitly the counterfactual is established. (necessary for outcome and impact evaluation)		
1.4 Measures the difference between the actual situation and the counterfactual in a credible manner	The result of the intervention is measured against the counterfactual in a credible way. Distinguishes between intended and unintended results. (necessary for outcome and impact evaluation)		
Overall rating for the quality criterion 1:			
Overall comments on criterion 1:			
Quality criterion 2 Establishes the theory of causal links between the intervention and its expected results			
2.1 Articulation of intervention logic	The report presents how the implementers of the intervention thought that the intervention would lead to the desired results. Identifies the intervention's results chain, logic model, theory of change or similar. The intervention and its underlying model is clearly defined, described and explained.		

Key quality criteria	Quality sub-criteria	Rating	Implications for evaluability of the grant
2.2 Assessment of intervention logic	The review assesses the intervention logic or theory, including underlying assumptions and factors affecting the success of the intervention. Is it a plausible model for the intervention? Are there alternative models that should have been applied for the intervention?		
Overall score for the quality criterion 2:			
Overall comments on quality criterion 2:			
Quality criterion 3 Makes use of data suitable for measuring the results			
3.1 Drawing on existing M&E data	The review presents and makes use of the existing M&E systems and data collected by the project intervention.		
3.2 Identifying and using available data suitable for the evaluation	The evaluator reviews the available existing data sources, assesses whether they are suited for the review and records the implications (i.e. insufficient data for the review implies that additional data needs to be collected).		
3.3 Collecting own data	Data collection is designed appropriately for the review questions to be answered. Sampling strategy is clear and appropriate for the review questions under scrutiny. Limitations of the representativeness of the samples are identified. Does this data support later efforts to evaluate the grant?		
3.4 Information sources	The report describes the data used (documents, respondents, administrative data, literature, etc.) in sufficient detail so that the adequacy of the information can be assessed. To what degree is this documentation helpful for later evaluations?		
Overall score for the quality criterion 3:			
Overall comments on quality criterion 3:			
Quality criterion 4 Documents the reliability of the evaluation and validates the conclusions (a high-quality MTR can be supportive to later evaluations indirectly by the fact that one would have some conclusions that one could trust)			

Key quality criteria	Quality sub-criteria	Rating	Implications for evaluability of the grant
4.1 Incorporation of stakeholders comments	Relevant stakeholders have been given the opportunity to comment on the draft report <i>and</i> the final evaluation report reflects these comments and acknowledges any substantive disagreements.		
4.2 Triangulation of information and methods	The evaluation triangulates data from different sources and critically assesses the validity and reliability of the information.		
4.3 Clarity of analysis	Findings flow logically from the data, showing a clear line of evidence to support the conclusions. Conclusions are substantiated by findings and analysis. Recommendations and any lessons follow logically from the conclusions. Any assumptions underlying the analysis are made explicit. The relative importance of findings is made clear in the report.		
4.4 Clear basis for judgement	The process for making judgements is clear and there is a clear explanation of the extent to which the evidence supports the judgements being made; gaps and limitations in the data are clearly explained.		
4.5 Reliability	All documentation and data used is either published can easily be made publicly available. If privacy /confidentiality conflicts with such transparency, a procedure for independent assessment of the data should be specified. The evaluation report discusses limitations of the study and its reliability.		
4.6 Replicability	The evaluations should document each step in the process sufficiently for others to be able to replicate their study.		
Overall score for the quality criterion 4:			
Overall comments on quality criterion 4:			
OVERALL QUALITY SCORE			
Overall comments on quality:			

Key quality criteria	Quality sub-criteria	Rating	Implications for evaluability of the grant

Completion report

Key quality criteria	Quality sub-criteria	Rating	Evidence
Quality criterion 1 Identifies a credible counterfactual and measures the results against this counterfactual			
1.1 Explaining evaluation methodology	The report describes and explains the evaluation methodology and its application for identifying the results of the intervention. The methodology includes specification and justification of the design of the empirical verification approach and the techniques for data analysis and estimation.		
1.2 Identifies the output that stems from the intervention	The evaluation makes a clear distinction between the output that can be attributed to the intervention from that of other contributors.		
1.3 Identifies the counterfactual	Implicitly or explicitly the counterfactual is established. (necessary for outcome and impact evaluation)		
1.4 Measures the difference between the actual situation and the counterfactual in a credible manner	The result of the intervention is measured against the counterfactual in a credible way. Distinguishes between intended and unintended results. (necessary for outcome and impact evaluation)		
Overall rating for the quality criterion 1:			
Overall comments on criterion 1:			
Quality criterion 2 Establishes the theory of causal links between the intervention and its expected results			
2.1 Articulation of intervention logic	The report presents how the implementers of the intervention thought that the intervention would lead to the desired results. Identifies the intervention's results chain, logic model, theory of change or similar. The intervention		

Key quality criteria	Quality sub-criteria	Rating	Evidence
	and its underlying model is clearly defined, described and explained.		
2.2 Assessment of intervention logic	The evaluation assesses the intervention logic or theory, including underlying assumptions and factors affecting the success of the intervention. Is it a plausible model for the intervention? Are there alternative models that should have been applied for the intervention?		
Overall score for the quality criterion 2:			
Overall comments on quality criterion 2:			
Quality criterion 3 Makes use of data suitable for measuring the results			
3.1 Drawing on existing M&E data	The evaluation makes use of the existing M&E systems and data collected by the project intervention.		
3.2 Identifying and using available data suitable for the evaluation	The evaluator reviews the available existing data sources, assesses whether they are suited for the evaluation and records the implications (i.e. insufficient data for evaluation implies that additional data needs to be collected).		
3.3 Collecting own data	Data collection is designed appropriately for the evaluation questions to be answered. Sampling strategy is clear and appropriate for the evaluation questions under scrutiny. Limitations of the representativeness of the samples are identified.		
3.4 Information sources	The report describes the data used (documents, respondents, administrative data, literature, etc.) in sufficient detail so that the adequacy of the information can be assessed.		
Overall score for the quality criterion 3:			
Overall comments on quality criterion 3:			
Quality criterion 4 Documents the reliability of the evaluation and validates the conclusions			

Key quality criteria	Quality sub-criteria	Rating	Evidence
4.1 Incorporation of stakeholders comments	Relevant stakeholders have been given the opportunity to comment on the draft report <i>and</i> the final evaluation report reflects these comments and acknowledges any substantive disagreements.		
4.2 Triangulation of information and methods	The evaluation triangulates data from different sources and critically assesses the validity and reliability of the information.		
4.3 Clarity of analysis	Findings flow logically from the data, showing a clear line of evidence to support the conclusions. Conclusions are substantiated by findings and analysis. Recommendations and any lessons follow logically from the conclusions. Any assumptions underlying the analysis are made explicit. The relative importance of findings is made clear in the report.		
4.4 Clear basis for judgement	The process for making judgements is clear and there is a clear explanation of the extent to which the evidence supports the judgements being made; gaps and limitations in the data are clearly explained.		
4.5 Reliability	All documentation and data used is either published can easily be made publicly available. If privacy /confidentiality conflicts with such transparency, a procedure for independent assessment of the data should be specified. The evaluation report discusses limitations of the study and its reliability.		
4.6 Replicability	The evaluations should document each step in the process sufficiently for others to be able to replicate their study.		
Overall score for the quality criterion 4:			
Overall comments on quality criterion 4:			

OVERALL QUALITY SCORE		
------------------------------	--	--

Key quality criteria	Quality sub-criteria	Rating	Evidence
Overall comments on quality:			

6. Staff survey

The purpose of the staff survey was to reach out to a wider network of programme staff dealing with grant management. The survey was implemented towards the end of the main study phase so that it could be used as a means of testing and validating the findings that emerged from the in-depth assessments with a wider sample of staff.

The survey was distributed to staff in the MFA, Norad and embassies dealing with development cooperation. Identification of recipients was done with the assistance of EVAL staff.

The survey was sent out to 477 staff. A total of 157 people opened the survey to look at the questions. Of these, 126 continued to complete Part 1 (Question 4) and 97 continued to complete the remainder of the questions.

1. Identification

Welcome to the survey of planning and managing for results in grant management. The survey is in two parts. A first section designed to give a rapid overview of your opinions about arrangements for results in grant management. This will take less than five minutes and we urge all recipients of this survey to try and complete this. The second section goes into a little more detail about the three main stages of Preparation, Follow-up and Completion. Again, it is not very long.

This survey is your opportunity to contribute to our evaluation and to have your say about managing for results. The more staff who respond, the more confidently we will be able to use these findings in our analysis, and the better our report will reflect the experiences of all staff.

All replies are anonymous and cannot be traced back to individuals. Thank you for taking part.

1. Department and section

2. Organisation

☐ MFA☐ Norad☐ Embassy☐ Norfund

3. Position

Headquarters

Embassies

Please select one of the options from the appropriate menu:

2. Quick assessment on Measuring results during grant management

4. Please indicate the extent you agree or disagree with the following statements:

	Strongly agree	Agree	Disagree	Strongly disagree	Don't know/not applicable
1. The internal guidelines and rules provide me with comprehensive guidance on supporting partners in planning for and measuring results (guidelines and rules include Grant management manual, Results Management in Norwegian Development Cooperation, Grant Scheme Rules).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. There are a number of rules and guidelines for results measurement in grant management and it is not always clear to me which ones I should use.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. I regularly use guidelines and rules at all stages of the grant management process (preparation, follow up and completion).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Last time I received a grant recipient's progress or final report I reviewed and analysed it to identify what has been learned from the work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. I always take the time to discuss progress and final reports with partners to identify areas for improvement.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. The internal training that is available on planning for and measuring results is relevant to the realities of managing grants.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. I am confident in my ability to assess partners' results frameworks to identify areas for improvement (e.g. goal hierarchy, indicators and baseline).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. I have a clear sense of the minimum requirements that a partner's results framework should have.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Building my skills in how to plan for and measure results is a high priority to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Financial management is the most important aspect of grant management.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. There are clear incentives (such as praise or promotion) to encourage me to ensure high quality results measurement in the grants that I manage.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. There are clear sanctions for me if the grants that I manage are unable to demonstrate/describe results in a credible way.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13. Evaluation is highly valued within the MFA and Norad.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. The senior leadership within MFA and Norad are driving the focus on measuring results.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. I have adequate resources (time and money) to ensure the measurement of results is put into practice.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Please continue with a more detailed assessment below which will take another few minutes, or if you are short of time you can exit the survey here

- ☐ Continue with the survey
- ☐ Exit survey now

3. Preparing for a grant

Please answer with reference to the most recent application that you have processed for approval (not rejected)

6. Have you managed any new grant applications within the past 3 years?

☐ Yes

☐ No

7. How familiar are you with the policies and procedures for examining an application for approval?

☐ Very familiar

☐ Familiar

☐ Somewhat familiar

☐ Not familiar

8. Please indicate how useful you think the policies and procedures are in helping review grant applications in the following areas:

	Very useful	Useful	Somewhat useful	Not useful
a) Assessing baseline data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) Reviewing how measurable and achievable project objectives are at different levels (outputs - outcomes - impact)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) Reviewing the logic of the project design	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) Assessing the quality of indicators and data collection plans	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e) Reviewing financial management	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f) Reviewing management and design of mid-term reviews, final reviews and evaluations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g) Assessing the appropriateness of the budget for Monitoring and Evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h) Assessing risks	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. Please indicate how much attention the following aspects of a new grant application receive when deciding whether to provide funding?

	Very much	Much	Little	Very little
a) Results framework	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) Financial management systems	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) Who the grant applicant is (their reputation)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) Fit of grantee with Norad/MFA strategic priorities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e) Other	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you have chosen e) Other, please specify:

10. From where did you request technical support for planning results at the preparation phase, and indicate how helpful you found it?

	No technical support requested	Very helpful	Helpful	Somewhat helpful	Not helpful
a) Specialists in Norad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) AMOR (Avdeling for Metode og Resultater)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) External experts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) Other	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you have chosen d) Other, please specify:

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?

☐ Completely
 ☐ Mostly
 ☐ Somewhat
 ☐ Minimal

12. What improvements do you think could be made to ensure results measurement and evaluation are given more prominence when preparing for a grant?

Please specify:

4. Follow up

Please answer with reference to what you consider your main or most important project or programme grant at the present time.

13. What importance do you attach to the following actions during the follow up phase in grant management?

	Very Important	Important	Routine	Low Importance
a) Timely disbursement of funds	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) Receiving financial and audit reports	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) Receiving progress reports	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) Conducting reviews	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

14. What importance does your manager attach to the following actions during the follow up phase in grant management?

	Very Important	Important	Routine	Low Importance
a) Timely disbursement of funds	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) Receiving financial and audit reports	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) Receiving progress reports	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) Conducting reviews	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

15. To what extent are you satisfied that your partner's annual reports provide an acceptable update on outputs and outcomes?

☐ Very satisfied ☐ Satisfied ☐ Unsatisfied ☐ Very unsatisfied

16. From where did you request technical support about results at the follow-up phase, and indicate how helpful you found it?

	No technical support requested	Very helpful	Helpful	Somewhat helpful	Not helpful
a) Specialists in NORAD	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) AMOR (Avdeling for Metode og Resultater)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) External experts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) Other	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you have chosen d) Other, please specify:

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?

☐ Completely ☐ Mostly ☐ Somewhat ☐ Minimal

18. To what extent do you find the time during the implementation phase to discuss with partners how they are measuring results and how this can be improved?

☐ Plenty of time ☐ Adequate time ☐ Minimal time ☐ No time

19. Please add any additional remarks you may have about results at the implementation phase:

5. Completion

Please answer in relation to the most recent project or programme grant that has reached completion.

20. Are you satisfied that your partner's final report provides an acceptable final statement on results ?

☐ Very satisfied ☐ Satisfied ☐ Unsatisfied ☐ Very unsatisfied

21. In your experience, how well do partner's final reports document outcome and probable impact?

- ☐ Comprehensively ☐ Satisfactorily ☐ Minimally ☐ Not at all

If you have chosen 'minimally' or 'not at all', please explain:

22. To what extent do you find the time to discuss with partners project results, implications and lessons when the project has been completed?

- ☐ Ample time ☐ Adequate time ☐ Minimal time ☐ No time

23. Do you feel like you have the necessary skills and knowledge to make an informed judgement on whether a partner's final reports are of a satisfactory quality?

- ☐ Completely ☐ Mostly ☐ Somewhat ☐ Minimal

24. To what extent was the final report used for lesson learning and to inform the design of new grants?

- ☐ A lot ☐ Quite a bit ☐ A little ☐ Not at all

25. Please add any additional remarks you may have about results at the completion phase

6. Training

A number of courses have been offered during the last years that cover aspects of grant management, and planning for and managing results. These are listed in the Foreign Service Institute (Utenrikstjenestens kompetansesenter) Catalogue and the current and updated courses can be found on their website.

Have you participated on any of the listed courses, or in similar courses (please refer to their website if you are uncertain), in the following topics during the past three years?

For each course please indicate how effective you consider it was in improving your skills relevant for planning for and managing results in grant management.

26. How effective were these courses at building your skills in results measurement?


	Not attended	Very effective	Effective	Slightly effective	Ineffective
a) Økonomistyring i tiskuddsforvaltningen (3 timer) [Financial management of grants]	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) Mål, resultat- og risikostyring i tiskuddsforvaltning / Mål, resultat- og risikostyring for ledere i utenriks tjenesten (3 timer) [Objectives, results and risk management in grant management/ in the foreign service]	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) Mål, resultat- og risikostyring i tiskuddsforvaltning, fordypning (2 dager) [Objectives, results and risk management in grant management - specialization]	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) Gjennomgang og evalueringer i tiskuddsforvaltningen (1 dag) [Review and evaluation in grant management]	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e) Utviklingspolitikk og tiskuddsforvaltning for ambassadepersonell (5 dager) [Development policy and grant management for embassy staff]	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f) Other courses or training that have included some aspects of planning for and managing results in grant management. Specify:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Other (please specify)

End of survey

Thank you for participating in the survey.

7. Focus group discussion outline



Evaluation of the Norwegian Aid Administration's systems and practices to ensure evaluability of Norwegian grants

Focus Group Discussions

Date: 8th October

Results in Development

Purpose of today's discussion

Explore each of the three hypothesis related to grant management, by drawing on your individual and collective experience and knowledge

Results in Development

Structure for the discussion

- We have developed a series of statements related to the three hypothesis. These have been grouped into three areas:
 - Grant policies, procedures and systems
 - Training and technical support
 - Incentives and culture
- For each statement you will be asked to indicate the extent to which you agree/disagree with the statement using sticky dots. The scoring will be discussed in a group.

Results in Development

Agenda

10 min	Introduction
20 min	Grant policies, procedures and systems
20 min	Training and technical support
20 min	Culture and incentives
20 min	Overall reflections and wrap up

Results in Development

Grant policies, procedures and systems	<ul style="list-style-type: none"> Internal policies and procedures (e.g. grant management manual, grant scheme rules, results management in Norwegian Development Cooperation) on how to manage grants and appraise partners' results frameworks provide appropriate and comprehensive guidance. I am familiar with internal policies and guidelines and use them at all stages of the grant management process (preparation, follow up and completion) I view reviews/evaluations of grants as an important way of monitoring progress and outcome achievement
Training and technical support	<ul style="list-style-type: none"> The technical support from AMOR on reviewing partner's grant applications and results frameworks is timely, relevant and practical The training that is available on results measurement is relevant and practical I am confident in my ability to appraise partner's results frameworks and identify areas for improvement
Culture and incentives	<ul style="list-style-type: none"> The senior leadership within MFA and NORAD are key drivers of the results agenda There are clear incentives within MFA and NORAD to focus on working with partners to show results Because projects are designed by our partners it is sometimes difficult to ensure results are clearly measured /described

Results in Development

8. Quality checklist for EVAL report

Our proposed framework for reviewing the quality of evaluation reports applies four criteria. We assess to which degree the report:

- identifies a credible counterfactual and measures the results against this counterfactual;
- establishes the theory of causal links between the intervention and its expected results;
- makes use of data suitable for measuring the results;
- documents the reliability of the evaluation and validates the conclusions.

Within each of these quality criteria we have identified a number of sub-criteria and for each of these we have developed a statement which describes appropriate practice. Content has been taken from both the OECD-DAC evaluation standards relevant for our assignment and from the research literature on programme evaluations (see Norad 2008 <http://www.norad.no/no/evaluating/publikasjoner/publikasjon?key=109790>). Each evaluation report will be assessed against each of the quality criteria using a five point rating scale (see below). In order to average across sub-criteria we score each rating. The scoring evidence that justifies why a particular rating has been given will be provided alongside.

Rating	Score	Quality
Poor	0	Quality deficient in important respects
Sufficient	1	Sufficient for informing about results
Good	2	Fully satisfactory quality
Very good	3	Best-practice model, suitable for emulation

Key quality criteria	Quality sub-criteria	Rating	Evidence
Quality criterion 1 Identifies a credible counterfactual and measures the results against this counterfactual			
1.1 Explaining evaluation methodology	The report describes and explains the evaluation methodology and its application for identifying the results of the intervention. The methodology includes specification and justification of the design of the empirical verification approach and the techniques for data analysis and estimation.		
1.2 Identifies the output that stems from the intervention	The evaluation makes a clear distinction between the output that can be attributed to the intervention from that of other contributors.		
1.3 Identifies the counterfactual	Implicitly or explicitly the counterfactual is established (necessary for outcome and impact evaluation).		
1.4 Measures the difference between the actual situation and the counterfactual in a credible manner	The result of the intervention is measured against the counterfactual in a credible way. Distinguishes between intended and unintended results (necessary for outcome and impact evaluation).		
Overall rating for the quality criterion 1:			
Overall comments on criterion 1:			
Findings on the role of evaluators' competencies (technical skills, experience etc.) in the rating, particularly for outcome and impact assessments. Refer to the evaluation matrix, Hypothesis 2, key evaluation questions 2.1 and 2.2, and sub-questions, especially. To what extent have evaluators put forward a methodology appropriate to the evaluation questions and clearly focused on gathering outcome and impact level evidence? To what extent have the evaluators got the necessary technical skills and experience to deliver on their proposed methodology? How well have evaluators implemented their specified methodology?			
Quality criterion 2 Establishes the theory of causal links between the intervention and its expected results			
2.1 Articulation of intervention logic	The report presents how the implementers of the intervention thought that the intervention would lead to the desired results. Identifies the intervention's results chain, logic model, theory of change or similar. The intervention and its underlying model is clearly defined, described and explained.		
2.2 Assessment of intervention logic	The evaluation assesses the intervention logic or theory, including underlying assumptions and factors affecting the success of the intervention. Is it a plausible model for the intervention? Are there alternative models that should have been applied for the intervention?		
Overall score for the quality criterion 2:			

Overall comments on quality criterion 2:			
Findings on the evaluators competencies:			
Quality criterion 3. Makes use of data suitable for measuring the results			
3.1 Drawing on existing M&E data	The evaluation makes use of the existing M&E systems and data collected by the project intervention.		
3.2 Identifying and using available data suitable for the evaluation	The evaluator reviews the available existing data sources, assesses whether they are suited for the evaluation and records the implications (i.e. insufficient data for evaluation implies that additional data needs to be collected).		
3.3 Collecting own data	Data collection is designed appropriately for the evaluation questions to be answered. Sampling strategy is clear and appropriate for the evaluation questions under scrutiny. Limitations of the representativeness of the samples are identified.		
3.4 Data mapping	Based on 3.1-3-3 develop a map of data and assess to what extent the evaluation to a sufficient extent has managed the data requirements for the evaluation questions.		
3.5 Information sources	The report describes the data used (documents, respondents, administrative data, literature, etc.) in sufficient detail so that the adequacy of the information can be assessed.		
Overall score for the quality criterion 3:			
Overall comments on quality criterion 3:			
Findings on the evaluators competencies:			
Quality criterion 4 Documents the reliability of the evaluation and validates the conclusions			
4.1 Incorporation of stakeholders comments	Relevant stakeholders have been given the opportunity to comment on the draft report <i>and</i> the final evaluation report reflects these comments and acknowledges any substantive disagreements.		
4.2 Triangulation of information and methods	The evaluation triangulates data from different sources and critically assesses the validity and reliability of the information.		
4.3 Clarity of analysis	Findings flow logically from the data, showing a clear line of evidence to support the conclusions. Conclusions are substantiated by findings and analysis. Recommendations and any lessons follow logically from the conclusions. Any assumptions underlying the analysis are made explicit. The relative importance of findings is made clear in the report.		

4.4 Clear basis for judgement	The process for making judgements is clear and there is a clear explanation of the extent to which the evidence supports the judgements being made; gaps and limitations in the data are clearly explained.		
4.5 Reliability	All documentation and data used is either published can easily be made publicly available. If privacy /confidentiality conflicts with such transparency, a procedure for independent assessment of the data should be specified. The evaluation report discusses limitations of the study and its reliability.		
4.6 Replicability	The evaluations should document each step in the process sufficiently for others to be able to replicate their study.		
Overall score for the quality criterion 4:			
Overall comments on quality criterion 4:			
Findings on the evaluators competencies:			

OVERALL QUALITY SCORE		
Comments on the role of evaluators competencies in the rating:		
Overall comments on quality:		

9. Evaluator competencies survey

The purpose of the evaluator competencies survey was to generate data on evaluators' self-assessment of their skills. The survey used proxy questions about the extent to which consultants were actively engaged in evaluation-related activities (such as being a member of a professional association, attending a training course, publishing an article or book, or holding a formal evaluation post of employment), and other questions about their familiarity with a wide range of theories and methods or tools.

The survey was sent out to the consultancy firms or team leaders from the six evaluations reviewed for this study, plus those consultants for whom email addresses were available – 14 in total. Altogether, 42 consultants were identified as having worked on the studies but many of these had technical roles rather than being evaluation experts. The companies were invited to distribute to all team members but some chose only to send to the evaluators. A total of 16 responses were received.

Introduction

This is a short survey to make an assessment of the skills profile of evaluators who have worked for EVAL. It is estimated to take less than 10 minutes.
All responses are anonymous and cannot be traced back to individuals.

1. How many years have you been undertaking evaluation assignments such as design of evaluations or as leader or member of an evaluation team?

- ☐ Less than 3 years ☐ 4 to 5 years ☐ 6 to 10 years ☐ More than 10 years

2. Are you a member of

- | | Yes | No |
|--|-----------------------|-----------------------|
| a) a national evaluation society? (If yes, give details) | <input type="radio"/> | <input type="radio"/> |
| b) a regional evaluation society (such as the European Evaluation Society)? (If yes, give details) | <input type="radio"/> | <input type="radio"/> |

Details

3. Have you ever attended a formal training course in project/programme or policy evaluation? Please indicate the type and duration:

- | | Yes | No |
|---|--------------------------|--------------------------|
| a) University level long term (part of a Bachelor's or Master's, or PhD Degree) (If Yes give details) | <input type="checkbox"/> | <input type="checkbox"/> |
| b) University level short course (less than 3 months) Most recent (If Yes give details) | <input type="checkbox"/> | <input type="checkbox"/> |
| c) Any continuing professional development short course within past 5 years (from an evaluation society or research institute or government agency including by distance learning; less than 3 months)(If Yes give details) | <input type="checkbox"/> | <input type="checkbox"/> |

Details

4. What is your educational background? Please indicate level (Bachelor's, Master's, PhD; and relevance of subjects studied to evaluation)

5. Have you ever designed, taught or led a training course in evaluation? If Yes, give details of all major/relevant

- ☐ Yes
☐ No

Details

6. Have you ever held a formal professional appointment as an evaluator (e.g. a designated university appointment; or a post in an evaluation office of an NGO or aid agency; or a post in a consultancy firm)? If Yes, give details

☐ Yes

☐ No

Details

7. Have you ever published a book or a journal article dealing with any aspect of evaluation methodology or the findings from an evaluation? If Yes, give details

☐ Yes

☐ No

Details

8. Please indicate those evaluation theories and approaches which you have used within the past five years, those additional ones you have not used and are uncertain about their utility, and those not used but which you would like to explore further. There is space below for additional comments if you have any.

	Used within past 5 years	Not used, unsure about its usefulness/ application	Not used, would like to explore this further in the future
Theory based evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Utilization-focused	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Participatory evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Experimental & quasi-experimental approaches	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Realist Evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Results based evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mixed methods	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Additional comments (optional)

9. Please indicate those evaluation and data collection methods and tools which you have used within the past five years, those additional ones you have not used and are uncertain about their utility, and those not used but which you would like to explore further. There is space below for additional comments if you have any.

	Used within past 5 years	Not used, unsure about its usefulness/ application	Not used, would like to explore this further in the future
Theory of change	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Logic models	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Use of the logical framework	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Selection of indicators	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Design of data collection for baseline	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Simple and multistage sampling for surveys	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Multivariate statistical analysis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cost benefit analysis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Propensity score matching	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Document review & analysis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Individual key informant interviews	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interviews with small groups	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Focus group discussions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Participatory methods	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Qualitative methods	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Qualitative Comparative Analysis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Case study	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ratings for assessments	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Value for money analysis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Systems analysis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Contribution analysis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Additional comments (optional)

10. List the five main evaluation criteria commonly used when preparing an evaluation design

11. What is the technical term used to denote what was expected to happen if the project or intervention had not taken place?

12. About you

Gender - Male or Female

Affiliation - Company (name), or Institution (name), or freelance

10. Summary assessment of Grant Scheme Rules results' requirements

MFA Grant Scheme Rules	Quality assurance (general, legal, No)	Evaluation of the scheme	Expert guidance / appraisal (Mandatory, mandatory under certain conditions, optional)
Cultural cooperation with countries in the South	None	Concrete plans in place	Optional
Culture, public diplomacy and information	None	Optional	Optional
Foreign policy debate	None	Concrete plans in place	None
EEA funds	None	Vague	None
Peace and democracy project, Cyprus	None	Vague	None
Information about European cooperation	None	Concrete plans in place	None
Peace and state building in the West Balkans	None	Vague	None
Cooperative programme with Romania and Bulgaria	None	Vague	None
Grant for European political dialogue	None	Vague	None
International relations research and development in Europe	None	Concrete plans in place	None
Regional allocation	Mandatory (Legal – public bodies and multilaterals; and General – Oil for development)	Vague	Mandatory (>50mill or if 'oil for dev'); optional (15–50mill)
Technical cooperation	Mandatory (Legal)	Vague	Mandatory (>50mill or if 'oil for dev'); optional (15–50mill)
Transitional assistance (gap allocation)	Mandatory (Legal – public bodies and multilaterals; and legal – >15mill)	Vague	Mandatory (>50mill); optional (15–50mill)
Arctic cooperation	None	Vague	None
Action plan for nuclear activities	None	Vague	Optional
Barents 2020	None	Vague	None
Disarmament and anti-terror	None	Vague	None
Development and disarmament	None	Vague	None
Project aid to ODA-approved OSSE countries in Eurasia	None	Vague	None
Educational work for peace	None	Vague	None
Project aid to northern Caucasus	None	Vague	None
Project cooperation with Russia	None	Vague	None
Project aid for development and disarmament – civilian emergency management and security sector reform	None	Vague	None
Climate and environment	Optional (Legal – only if large and/or complex project)	Vague	Optional
Women and gender equality	Optional (Legal – only if large and/or complex project)	Vague	Optional; mandatory (>15mill)
The global campaign for the health-related millennium development goals	None	Vague	None

Peace, reconciliation and democracy	Optional (Legal – only if large and/or complex project)	Vague	Optional
Voluntary core grants to the UN system	None	Vague	None
Additional funds to the UN system	Mandatory (Legal)	Vague	Optional
Membership fees to the UN system	None	Vague	None
Debt relief, support for payment balance, and capacity building	None	Vague	None
Capital increase and fund replenishment to financial institutions	None	Vague	None
Co-financing via financial institutions	Mandatory	Vague	Optional (when complex funding)
Human rights	Optional (Legal – only if large and/or complex project)	Vague	Optional
Natural disasters and humanitarian aid	None	Vague	Optional

Norad Grant Scheme Rules	Quality assurance (General, legal, No)	Evaluation of the scheme	Expert guidance/appraisal (Mandatory, mandatory under certain conditions optional,
Rules for support to civil society actors – Chapter 160.70	Partner requires QA systems	Vague	Optional
Rules for cooperation about framework conditions for private sector development in the South	None	Vague	Mandatory (>15mill); mandatory (more extensive >50mill)
Rules for application based support to the private sector, Chapter 161, post 70, Private sector development	None	Vague	None
Grant scheme rules for research and research Dissemination, Chapter/Item 165.70	None	Vague	Optional
Rules for democracy support through political parties, Chapter 160.72	Partner requires QA systems	Vague	None
Rules for the information work of voluntary organisations, Chapter post 160.71	Partner requires QA systems	Vague	Optional
Rules for Department for Civil Society's part of Chapter post 166.71 International processes and conventions – split post	None	Vague	None
Rules for the Climate and Forest Funding to Civil Society, Chapter 166.73	None	Concrete plans	Optional
Rules – rights of religious minorities, Chapter 163.72	None	Vague	Optional
Grant scheme rules for support to International Organisations and Networks, Chapter post 160.75	Partner requires QA systems	Vague	None

11. Grant commitment tables

Table 11.1: Number and value of project and programme grants committed

Grant type		2009	2010	2011	2012
I – Project- and programme grants	Number	1698	1535	1309	1192
	Value NOK million	10,983	15,983	13,133	12,037
	Average value NOK million	6.47	10.41	10.03	10.10
	<i>Number</i>	<i>58%</i>	<i>58%</i>	<i>55%</i>	<i>51%</i>
	<i>Value NOK million</i>	<i>37%</i>	<i>40%</i>	<i>52%</i>	<i>51%</i>
Total all grants	Number	2948	2663	2380	2343
Total all grants	Value NOK million	29,877	40,282	25,411	23,695
	<i>Percentages</i>	<i>100</i>	<i>100</i>	<i>100</i>	<i>100</i>
	<i>Percentages</i>	<i>100</i>	<i>100</i>	<i>100</i>	<i>100</i>

Table 11.2: Number and value of project and programme grants committed by grant size (million NOK) by year

Year		1>	1–5	5–10	10–25	25<		Total
2009	Number	43%	35%	11%	9%	3%	100%	1698
	Value NOK million	2%	13%	11%	20%	54%	100%	10,983
	Average NOK million	0.34	2.35	6.75	14.67	134.16		6.47
2010	Number	35%	34%	13%	11%	7%	100%	1535
	Value NOK million	1%	8%	8%	16%	67%	100%	15,983
	Average NOK million	0.39	2.38	6.81	14.37	99.78		10.41
1	Number	37%	34%	12%	11%	6%	100%	1309
	Value NOK million	2%	8%	8%	16%	66%	100%	13,133
	Average NOK million	0.43	2.32	6.69	14.80	104.99		10.03

2012	Number	35%	32%	13%	12%	7%	100%	1192
	Value NOK million	1%	8%	9%	18%	64%	100%	12,037
	Average NOK million	0.42	2.36	6.76	15.01	94.86		10.10

Table 11.3: Number and value of grants committed by duration of grant (years)

Year		1>	1–2	2–3	3<		Total
2009	Number	46%	38%	10%	6%	100%	1,698
	Value NOK million	18%	26%	13%	43%	100%	10,983
	Average NOK million	2.55	4.41	8.53	47.74		6.47
2010	Number	42%	36%	13%	9%	100%	1,535
	Value NOK million	19%	21%	19%	41%	100%	15,983
	Average NOK million	4.72	6.17	14.99	45.77		10.41
2011	Number	41%	33%	13%	13%	100%	1,309
	Value NOK million	14%	21%	23%	42%	100%	13,133
	Average NOK million	3.37	6.34	18.16	33.20		10.03
2012	Number	34%	41%	15%	10%	100%	1,192
	Value NOK million	11%	33%	24%	32%	100%	12,037
	Average NOK million	3.29	7.99	16.63	32.75		10.10

Annex 5 Comparator agency study

This annex contains the following

1. The **findings** from our review of three comparator agencies
2. A list of **references** consulted for the study
3. A copy of the **assessment framework** that was used

1. Findings from our review of three comparator agencies

Introduction

The terms of reference for this assignment specify under Part 3 point 7, ‘*Assessment of similar kinds of standard documents (described under point 3)² 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.*’ This paper explores these documents and procedures for the World Bank, DFID and Danida. Some additional information is quoted about the Ministry for Foreign Affairs, Finland. The approach is to establish a yardstick, by summarising practices and procedures using standardised frameworks. The work is limited to a review of what organisations say they do (policies and systems) rather than their actual practices, which would be beyond the scope and resources of this study. However, some information is available about the effectiveness of other organisations’ systems and this is discussed in the report. The report explains the approach taken, sets out the findings on evaluability arrangements, discusses the way other organisations learn about results management and considers institutional culture and incentives. A final concluding section indicates what lessons Norway could draw from the practices of other agencies.

Objectives

The purpose of the review is to compare the regulations and practices of other aid institutions in supporting and encouraging evaluability and results documentation in order to identify possible alternative practices and lessons that could be of use to Norway.

Methodology

Data for the comparisons was collected through a combination of **documentary evidence** from comparator organisations such as policies, processes and structures which guide results management in the grant-making or lending process³ and the functioning of the evaluation units, and **telephone interviews** with key informants. Data from grant and project management, quality assurance and the evaluation function formed the basis of the assessment. Annex 3 contains details of the assessments. A rich text rating scale was developed by the team based around criteria for grant or project management. The assessments were summarised under a three-part scale as ‘Comprehensive’, where documents indicate the organisation follows the criterion closely; ‘Moderate’, where the general approach follows the criterion but with some shortfalls on coverage or details; and ‘Slight’, where this criterion is dealt with much less fully than the stated standard. A fourth category of ‘Absent’ was also considered, but no instances were found where the organisation did not address the criteria at all. All the ratings are supported by reference to documents with specific pages or paragraphs noted where relevant.

The analysis looks first at the arrangements in the three comparators against a checklist of good practice. The next section examines how those agencies’ procedures differ from those of Norway and

² Assessment of the degree to which relevant documents in the Norwegian aid administration such as guidelines, handbooks, standard contracts, grant scheme rules, instructions, etc., focus on – and set a standard for – evaluability and results measurement.

³ Because the nature of interventions, various among the comparators (loans, grants etc.) the word ‘project’ is used as a generic descriptor of all interventions.

draws some lessons. Following these comparisons, issues of lesson learning and results culture are explored.

Evaluability and results in the project cycle

Criteria were developed by the team based on a broad review of UN and OECD-DAC standards and norms for evaluation, to review how evaluation is embedded in intervention management, looking at policy and guidelines for stages in the programme cycle. The criteria used are noted in *italics* next to each stage.

- **Design and development of an intervention for appraisal or approval**
 - Logic model and indicators: *Clear guidance is given about the need to base designs on evidence and the use of results, and to construct a logic model or theory of change.*
 - Evaluability: *The question of evaluability is examined, with provisions for baseline data and consideration of planning for future evaluation.*
 - Appraisal and approval: *The organisation's procedures include review of the evidence to support the intervention design and arrangements for an evaluation plan.*
 - Quality Assurance: *Provision exists for arms-length QA process to review (i.a.) intervention logic/theory of change, evidence base, need for evaluation, and economic analysis.*
- **Follow-up**
 - Monitoring, reporting and reviews: *Arrangements include regular reporting against structured performance criteria and periodic independent reviews no less than one at mid-term.*
 - Completion: *Arrangements include provision of a completion report by self- or independent evaluation, to include assessment against OECD-DAC Evaluation Criteria and others as appropriate.*
- **Completion – review and evaluating the intervention**
 - Self and independent evaluation: *Arrangements include a coherent model either of independent evaluation or a structured combination of self-evaluation and independent validation; plus thematic or cross-cutting studies.*
 - Management follow-up: *Evaluation recommendations are published and require an explicit response by the governing authorities and management addressed by its recommendations.*

Table 1 presents a summary of ratings colour coded Comprehensive (green); Moderate (amber); and Slight (yellow).

Table 1 Evaluation practice and results focus in intervention management

Stage	Aspect	Norway	Danida	DFID	World Bank
Design and development of an intervention	Logic model and indicators	Grant Management Manual (GMM) Section V04 gives a guide to management results and risks. Sets out key results concepts with a very brief (1 page) overview of a goal hierarchy. Some discussion of actions during the project cycle but with more emphasis on risks	Specific text about developing an intervention logic. Checklist of document context required of project planned by partners includes reference to intervention logic and expected outputs and outcomes	Clear and consistent guidance covering theory of change, logframe and indicators through EvD materials and the Business Case	Uneven treatment. The Operational Manual OP/BP 10.00 has minimal guidance on results, no mention of evaluability and very basic notes on M&E and indicators
	Evaluability	No discussion in the GMM; only a small reference in the Results Management (RM) document	Evaluability is considered as part of preparing for an evaluation. The five OECD-DAC criteria are applied when designing new interventions	Evaluability and results dealt with under: • Strategic case: expected impact and outcome • Appraisal Case: assessing the strength of the evidence base; Theory of Change for the preferred option; what measures can be used to assess value for money for the intervention; • Management case: How will progress and results be monitored, measured and evaluated	Also uneven but some guidance advises that the project document includes a results framework with measurable indicators (with quantitative baselines and targets, whenever possible) for monitoring progress during implementation and evaluating outcomes on completion
	Appraisal and approval	Some treatment in GMM but with no illustrative examples and no cross referencing to other sources. Better treatment in RM different stages of the programme cycle	Some references are made to subsequent evaluation in the Guidelines for Programme Management	Checklist-based scrutiny closely linked to the QA process	Minimal treatment with results framework being finalised after appraisal
	Quality assurance	No explicit treatment in either GMM or RM. RM describes appraisal as 'quality control of the programme proposal'. QA is supposedly performed by advisers in Norad when projects are being considered for approval, but this is not a mandatory stage	Technical Advisory Service is responsible for quality assurance through appraisal of major programme support proposals	Specific QA using checklists covering results with arrangements by size of project	Historically a detailed system, with reviews for content on samples of projects; currently under redesign
Follow-up	Monitoring, reporting and reviews	The GMM references a reporting proforma (S51) with requirements for indicators of outputs and outcomes. But this is not mandatory and there are no references to evaluation criteria nor are there any assessments of performance	Little detail in the guidance and no use of evaluation criteria	Provision for published Annual Reviews; plus a How-To note: reviewing and scoring projects	Six-monthly Implementation Status Report includes ratings of implementation performance and progress towards development objectives

Stage	Aspect	Norway	Danida	DFID	World Bank
	Completion	All grants have to have a final report, with a proforma S61, but like the progress report this format is not mandatory and there are no references to evaluation criteria nor are there any assessments of performance	Project Completion Report (PCR) is based on the implementing partners' final reports. The PCR template includes provision for comments on effectiveness, efficiency and sustainability with regard to indicator performance monitoring	Provision for published Programme Completion Review on all projects; plus a How-To note: reviewing and scoring projects	All project have to have an Implementation Completion Report (ICR). Ratings do not follow OECD-DAC but are long established in the Bank to cover performance
Completion	Self and independent evaluation	There are no references to evaluation in the GMM. Regulations for Financial Management in Government state that all agencies shall ensure that evaluations are performed	Arrangements centre on the PCR. The Quality Assurance Department will make the initial quality assessment of the PCR after completion and the Evaluation Department will undertake cross-cutting analyses of selected issues in the PCR's for accountability and/or learning purposes as relevant	New arrangements since 2010 with an Independent Commission on Aid Impact and DFID's Evaluation Department strengthening decentralised evaluation in DFID	The ICR is a core building block in the model of Self-Evaluation validated by Independent Evaluation pioneered by the WB
	Management follow-up	Evaluations are published and there is provision in the 'instruction' for follow-up to recommendation decisions	The Quality Assurance Department is responsible for reporting to the Programme Committee concerning follow-up to evaluations completed within the past two years	A formal management response is required for all DFID evaluations. ICAI publishes a detailed follow-up of previous reports as an annex to its annual report	Formal procedures exist for management response and follow-up, scrutinised by a committee of the board

There is quite a bit of variation across the agencies in this assessment, with only DFID rated as having comprehensive provision across all the criteria (see Box 1). Note that the assessment is about the extent to which there is guidance on results and planning for evaluability, not on the extent to which this guidance is put into practice, nor the overall quality of project design and implementation. Many other important aspects such as dealing with risk and in treatment of cross-cutting issues are not the subject of our analysis here. It is important to note also that institutional arrangements and standards are constantly changing. Awareness of the need to analyse the planned intervention logic has changed considerably in the past 20 years, from a 'nominal' requirement for a logical framework with indicators, to a more comprehensive approach identifying risks and contextual factors, a theory of change diagram or narrative, standardised indicators and structured links between project and corporate reporting. For example, new systems are being introduced for quality assurance in the World Bank and once established our assessment might well change. Details of the documentation reviewed to support these ratings can be found in Annex 2 and the fully referenced assessments are in Annex 3.

Danida has the 'simplest' system of the three comparators, with the majority of guidance concentrated in two core documents: *Guidelines for Programme Management*, and *Danida Evaluation Guidelines*. This may reflect the relatively small aid programme and simpler organisational structure than DFID or the World Bank. There may be some advantages in being able to summarise complex procedures in few documents, especially if many staff use the procedures infrequently and are not technical or project management specialists. That could be an important lesson for the Norwegian Aid Administration. Both DFID and the World Bank have a much greater diversity of sources of guidance and variation in the extent to which these sources are mutually coherent. Thus for project preparation DFID draws on a range of UK government-wide materials and has formal Office Instructions supported by more specific technical advice or 'How-To' notes. The evaluation function has a separate comprehensive policy statement and there is a technical evaluation handbook.

Box 1 Why DFID is rated *Comprehensive* for 'Design and development of an intervention'

Five features of DFID's approach combine to justify ratings of Comprehensive:

- **Broad coverage in the documentation:** there is a continuity of message across guidance for preparing a project Business Case;⁴ quality assurance arrangements; evaluation policy; and evaluation training materials, with some cross referencing.
- Recognition that a clear logic model and results based on prior evidence strengthens the **quality of project design** rather than being a formality to complete a project proposal.
- **Evaluability is assessed from several perspectives:** expected impact and outcomes; strength of the evidence base; theory of change; and what arrangements are needed to measure, monitor and evaluate progress and results.
- **Documentation includes detailed descriptions**, training or self-briefing materials and examples for staff to follow.
- There is **consistency of message** across planning guidance, appraisal and approval, with a detailed checklist for **quality assurance**, questioning:
 - Is a suitable range of credible evidence sources quoted?
 - Is there a credible theory of change which explores and provides evidence for the critical assumptions and linkages from input to impact?
 - Is there a good analysis of behavioural change?
 - Does the Business Case set out a clear evaluation plan following the guidance in the How-To Note?
 - Is the logframe based on a clear results chain? (link to theory of change)?
 - How well does the logframe set out indicators, baselines and milestones, and data sources?

DFID's approach is the only one that emphasises the systematic treatment of evaluability through the planning stage; for training of staff; for strengthening the evidence base that underpins evaluations; and for requiring managers to make a formal response to evaluations. And that approach is backed up by detailed guidance for theories of change, the logical framework, indicators and impact evaluation.

The World Bank has the most extensive and diverse set of guidance. There is a large body of core material about Operational Policies and Bank Procedures, but much of this is quite dated. A

⁴ The Business Case is a UK government-wide approach to investment planning that considers an investment from five different perspectives to make the case to invest: strategic; appraisal; commercial; financial; and management.

substantial reform of the Operational Manual was started in 2011 and is still under way. More up to date guidance on project design, arrangements for M&E and quality assurance comes from the Operations Policy and Country Services (OPCS) Vice-Presidency. But some of the most relevant materials are very new, first published in 2013 in several instances, so, until now, there is no experience of how effectively these will influence working practice. In support of this central guidance, regional vice-presidencies have a high degree of autonomy in promoting quality and results focus. In addition, there are a number of Bank-wide programmes to promote impact evaluation such as the Development Impact Evaluation Initiative (DIME) and Strategic Impact Evaluation Fund (SIEF). The plethora of initiatives makes for a fertile and innovative set of approaches but from reviewing the documents it is not clear how well connected they are. For example, the OPCS 2013 note on Results Framework and M&E Guidance does not contain any cross references to either DIME or SIEF, nor to lessons learned from the Independent Evaluation Group.

Grant and project management – a comparison

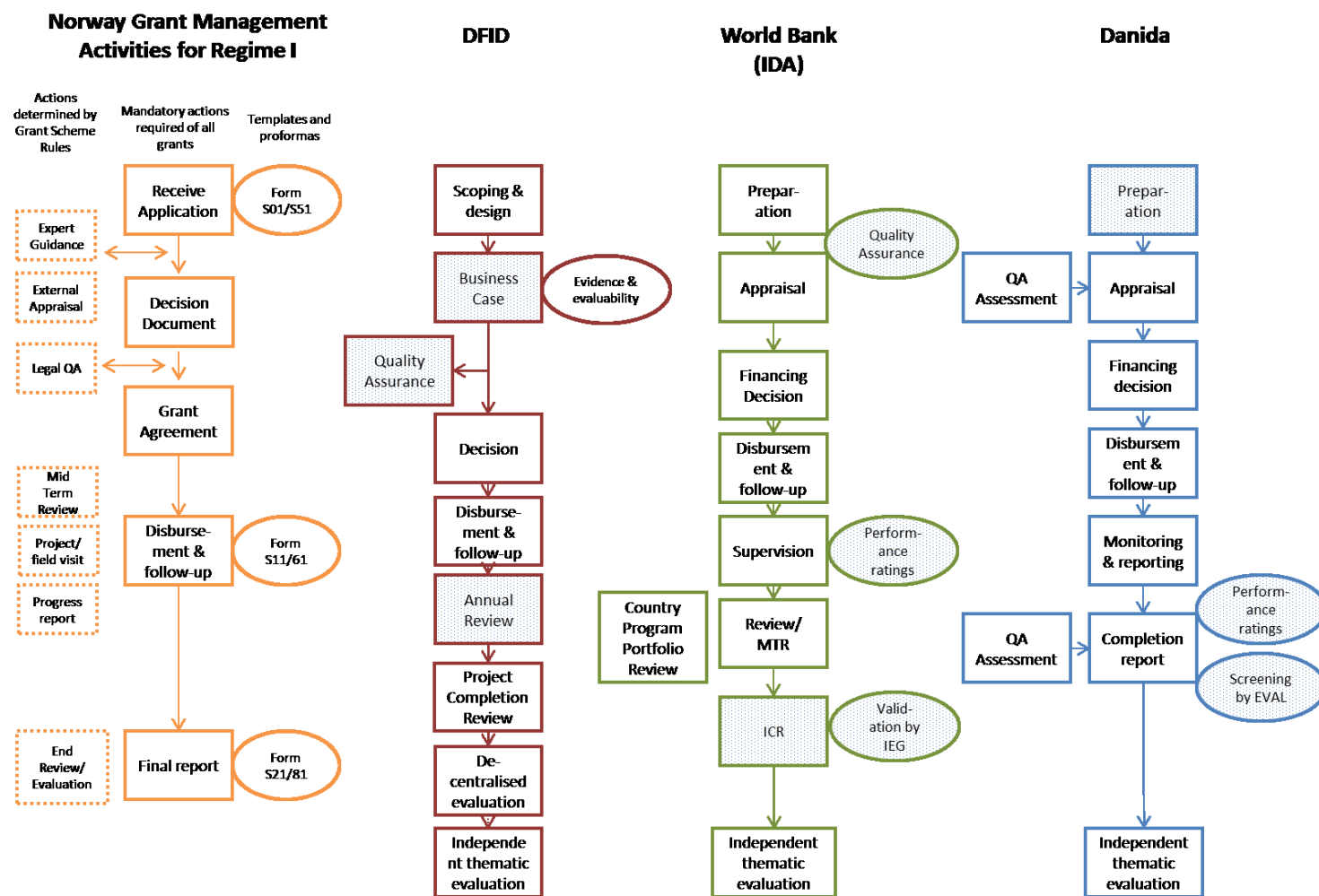
Figure 1 presents a schematic comparison of the grant/project cycle processes of Norway, DFID, World Bank (IDA) and Danida. The stages do not show in full detail every aspect of each agency's systems as these are complex with numerous detailed variations according to type and size of loan or grant. However, the stages shown have been chosen to highlight those fundamental aspects of the process where comparisons about a focus on results and evaluability can be drawn between the approach adopted by Norway's MFA and Norad, and those of the comparator agencies.

The shaded boxes indicate process stages where a comparator has an approach that has some additional benefits over the Norwegian system. While there is, unsurprisingly, much common ground in the processing stages among all four organisations, key differences about the way evaluability is handled are discussed under five headings:

- *The granting framework*
- *Grant application and the planning stage*
- *Quality assurance*
- *Follow-up and monitoring*
- *Completion*

The analysis focuses on how other organisations' systems deal with evaluability in a different way from Norway.

Figure 1 Project cycle stages and comparisons



The granting framework

A fundamental difference between the approach of Norway and all the other comparators is that staff of the MFA and Norad do not prepare projects, but receive applications from potential grant recipients and negotiate the objectives, work plan and financing of the project. While the comparator agencies all ground their projects in the national policy objectives and wishes of their development partners, responsibility for grant/project 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. Thus for 18 of 45 grant schemes, the grant manager is required to seek expert guidance when reviewing an application; and for 5 of 45 grant schemes an external appraisal of the application has to be carried out.⁵

Since 2010, Norway has a standard proforma (form S01 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 on-line or used as a checklist against a grant applicant's own documentation.

Grant application and the planning stage

If the provisions of form S51 are followed correctly, then a grant applicant will set out a hierarchy of objectives together with planned indicators for follow-up. The guidance for completing S51 contains a very brief statement about goals and indicators. Some supporting information can be found in section V04 of the Grant Management Manual. Both are brief with no detailed explanatory text or examples. All the comparators have some variations in project processing according to size or purpose of grant or loan but none of the others treats technical scrutiny or independent appraisal as an optional process. All three comparators specify requirements about results and evaluability in more detail.

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 **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 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.

Quality assurance

Norad's Quality Assurance Department (Metode og resultater – Avdeling for Metode og Resultater (AMOR)) has programmes to help develop staff capacity and conducts grant management reviews at embassy level. These reviews do not review the *quality* of the grant application, progress report and final report form, but rather the *embassies' assessment* of the said documents. In other words the quality assurance deals with quality of processes rather than content. All three comparators make use of quality assurance of content to test project design prior to a financing decision.

DFID has a multi-stranded approach to quality assurance, 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 quality assurance checklists include detailed questions about results and evaluability. For example:

- Q17. Is there a credible theory of change which explores and provides evidence for the critical assumptions and linkages from input to impact?
- Q18. Is there a good analysis of behavioural change?

⁵ In fact, the guidance for these grant schemes is only mandatory in three cases according to the size of the grant in NOK, and for all 'oil for development' grants. In all other cases the guidance specifies 'if necessary' or 'if relevant'.

- Q30. Does the Business Case set out a clear evaluation plan following the guidance in the 'How-To Note' (para. 130)?
- Q31. Is the logframe based on a clear results chain? (link to theory of change)
- Q32. How well does the logframe set out indicators, baselines and milestones, and data sources?

Arrangements for quality assurance 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 revealed a decline in ratings of project outcomes which has been linked to problems with quality of design and readiness for implementation; weaknesses in monitoring and evaluation; complexity; and weaknesses with proactive supervision.

The new quality assurance arrangements have not yet firmed up but will involve a review of all proposed projects and include optional quality enhancement reviews during preparation.

The quality assurance system acts as a way of setting standards and assessing performance. In both DFID and the World Bank quality assurance is a key feature, used to set standards and assess how well projects meet those standards. Coherence between technical guidance, training for staff and quality assurance is essential for the system to be effective, as explained later in Box 6. Figure 1 illustrates this layered system.

Danida has provision for quality assurance 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.

Follow-up and monitoring

Norway's requirements for progress reports and reviews vary according to the grant scheme and are not mandatory for all grants. A new proforma, S11/S61 provides a structure for the information that is required, which includes some provision for reporting against indicators for outcome and outputs. Completing the form is not mandatory and it can be used as a checklist.

Both DFID and the World Bank go further in their monitoring reports, with the use of rating systems to document performance.

DFID has a formal annual review and project completion review. The annual review scores projects against actual achievement of expected results alongside an assessment (but not a score) of the outcome. All Annual Reviews and PCR 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 a former system in which large projects of long duration were rated for delivery of outcomes.

The **World Bank** has had a long-established six-monthly progress report, now called the Implementation Status and Results Report. This includes ratings for progress towards development objectives and implementation progress. The development objectives rating has to be supported by reference to indicators from the project logframe.

The **Danida** minimum demand from each programme support component is one annual plan and budget, as well as one annual progress and financial report. The day-to-day monitoring 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.

Completion

All Norway's grants have to have a final report. A new proforma, S21/S81 provides a structure for the information that is required, which includes some provision for reporting against indicators for outcome

and outputs. The form is not mandatory and can be used as a checklist. All three comparators go further in their attention to results.

DFID's Project Completion Review uses the same scoring as the Annual Review but now 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 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 Project Completion Report 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 (development and immediate objectives).

Overview of comparisons

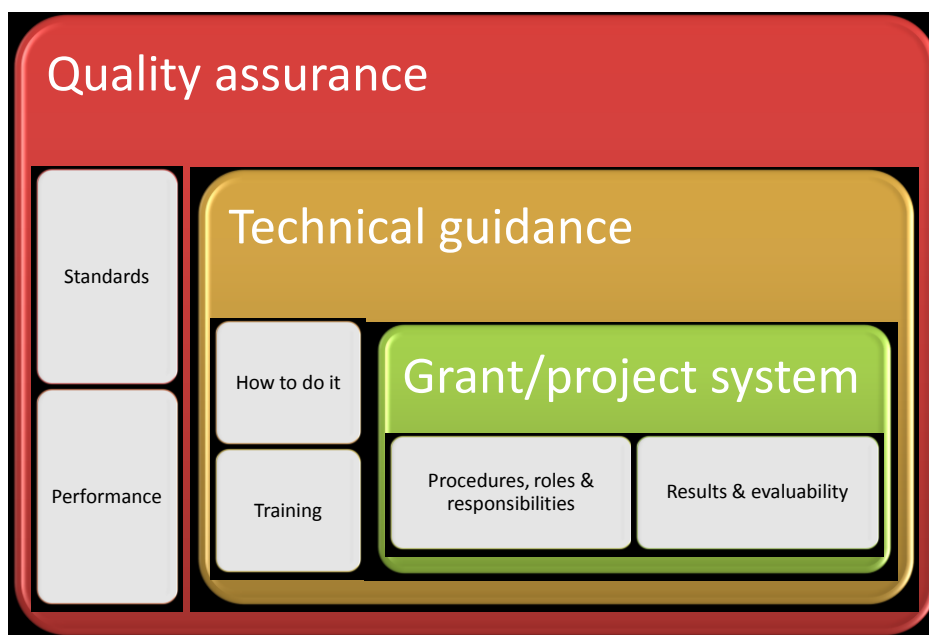
Table 2 summarises the differences among the systems reviewed here.

Table 2 Summary of key differences in the comparators' systems

Grant stage	Norway	DFID	World Bank	Danida
Grant framework	Receives applications to consider for approval	Prepares or commissions preparation	Appraises prepared project proposals	Prepares and appraises some and reviews applications as well
Grant application and the planning stage	Requires a goal statement, indicators and baseline	Requires evidence to support the intervention, a theory of change, logframe and its indicators, and an evaluation plan	Requires an abbreviated version of the logframe and arrangements for monitoring and evaluation	Specific requirement for intervention logic and theory of change
Quality assurance	QA in the form of non-mandatory advice by technical experts in Norad during approval, then independent QA of compliance with grant management rules	Independent or self-administered QA of each grant including theory of change, results chain, indicators, baselines and milestones, and data sources and evaluation plan	System under change; previously a mix of quality enhancement reviews and independent quality at entry reviews of a sample of projects	Design stage QA by technical advisory services then separate QA Department follow-up during implementation
Follow-up and monitoring	Progress reports and reviews vary according to the grant scheme and are not mandatory for all grants; guidance includes provision for reporting against indicators for outcome and outputs	Mandatory published annual review score against actual achievement of expected results alongside an assessment of the outcome	Six-monthly Implementation Status and Results (ISR) report includes ratings for progress towards development objectives and implementation progress	No detailed requirements but reporting is supposed to involve SMART indicators and progress towards expected outputs and outcomes
Completion	All Norway's grants have to have a final report. A proforma provides a structure which includes some provision for reporting against indicators for outcome and outputs	Project Completion Review (PCR) uses the same scoring as the annual review but includes a rating for the achievement of outcome as well as outputs. All projects must have a PCR	Implementation Completion Report (ICR) for all projects with ratings of performance. All ICRs are subject to validation by the Independent Evaluation Group	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

Staff need access to technical support that explains how to implement the grant or project procedures and training material for their own professional development. Some of the DFID materials such as 'How-to notes' represent good practice here. The examples from the World Bank are very recent and have not been in use for long.

Figure 1



These three components need to be mutually supporting for a high-quality system to function. Thus the requirements to plan expected results and ensure that a project can be evaluated necessitates that staff and grant applicants receive training and access to background technical guidance. The desired approach then needs to be reflected in the criteria set out for quality assurance and reported as part of quality assurance reviews so that lessons can be fed back to project planners and reviewers.

This review has focused on what are largely technical criteria about how projects are planned, implemented and evaluated. But are these technical procedures sufficient? Studies and events in the comparators suggest that other factors may be as or more important. The next two sections look at how the organisations learn about results and evaluation, and the role of institutional culture and incentives.

Summary of key findings

- ❖ Because Norway neither prepares nor appraises grants, the scope for interaction at the planning stage is much less than for the other donors.
- ❖ Norway's requirements for objectives and indicators at the planning stage are less detailed and do not include description of the intervention logic or an evaluation plan.
- ❖ DFID's requirement for evidence to support an intervention provides a logical foundation for planning the type and detail of evaluation to be planned.
- ❖ The three comparators all have more detailed quality assurance of individual grants at or before appraisal.
- ❖ Norway requires basic reporting but frequency and content varies by type of grant. Two of the three comparators make use of ratings to summarise performance.
- ❖ All Norway's projects have to have a final report (not an evaluation) which should include indicators for outputs and outcomes. The three comparators require a completion report (more analytical than the report specified by Norway) and require a rating assessment of performance as well.

Learning about results and evaluation

Although the grant or project management procedures and their QA systems are generally designed to produce regular flows of information about performance, often issues are only identified as a result of a more in-depth study. Several examples are quoted here by way of illustration. They are not exhaustive.

One of the best known examples is the so-called Wapenhans Report into portfolio management that highlighted failings in projects and gave rise to a new business model in the World Bank, with greater attention to quality of implementation, or 'results on the ground'. See Box 2.

Box 2 Periodic self-assessment at the World Bank

The Bank has a long-established practice of taking stock of its own performance. A landmark 1992 report of the Task Force into portfolio management, known as the Wapenhans Report after its principal author, identified that over a third of World Bank projects completed in 1991 were judged failures by the Bank's own staff, a dramatic 150% rise in failures over the previous 10 years. The report laid out in forthright terms the long-standing worries about the World Bank's 'lending culture', which tends to reward operational staff for the volume of their lending, with (it is argued) too little weight given to the quality of lending or on knowledge-related products or services. That report was the genesis of a series of initiatives. A shift in emphasis by the Operations Evaluation Department (OED) to examine Bank policies and practices as well as project performance; occasional studies into monitoring and evaluation;⁶ greater investment in performance such as the Annual Review of Portfolio Performance and subsequently creation of the Quality Assurance Group which instigated quality at entry and quality of supervision reviews.

Arrangements have continued to evolve. Major initiatives in recent years has been the establishment of an annual Results and Performance Report⁷ by the Independent Evaluation Group (IEG, formerly OED); proposals to reform the Bank's Operational Policy Manual (2011); adoption of a new lending instrument, Program-for-Results; developing the IDA results measurement system; and the Corporate Scorecard.

Ten years after Wapenhans, concerns about development results were still prevalent, and work by the IEG in the Bank led to a greater acceptance that successful project outcomes are more often associated with good designs for M&E. Not just because results are measured, but through the contribution that M&E analysis brings to the intervention logic of a project design and ability to evaluate performance. See Box 3.

Box 3 The key role of M&E design at the World Bank⁸

In the mid-2000s, there was increased international emphasis on the importance of development results. This culminated in several internal initiatives, including: (i) the creation of the Results Steering Group within the Bank in June 2006; (ii) a review of operational policies and guidelines; and (iii) greater attention to the results agenda. The IEG also started to put more emphasis on results and delivery of outcomes of Bank projects and required strong evidence in ICR to demonstrate the achievement of outcomes.

a) In July 2006, the Bank and IEG agreed to harmonise their rating scales and the Bank replaced its own internal four-point scale with IEG's six-point scale that had been in place since 1995.

b) In the same year, IEG also started putting more emphasis on the evidence provided in ICRs. They added the following reminder in their evaluation form: 'When insufficient information is provided by the Bank for IEG to arrive at a clear rating, IEG will downgrade the relevant ratings as warranted beginning July 1, 2006.' OPCS issued new ICR guidelines in the same year.

c) In response to the increased emphasis on results and efficiency, IEG began a systematic rating of the M&E quality of projects in July 2006. A separate rating for efficiency was also introduced by IEG around the same time.

⁶ See for example: *An Overview of Monitoring and Evaluation in the World Bank*. June 30, 1994. Report No. 13247; *Monitoring and Evaluation Plans in Staff Appraisal Reports Issued in Fiscal Year 1995*. 29 December 1995. Report No. 15222

⁷ See for example: IEG (Independent Evaluation Group) 2011. *IEG Annual Report 2011: Results and Performance of the World Bank Group*. Washington, DC: IEG, the World Bank Group

⁸ See: OPCS – Bank quality assurance framework July 2013.pptx; Results Framework and M&E Guidance Note. OPCS April 2013; Operational Policy Manual Issues & Prospects for Reform Approach Paper 2011

A review of the portfolio showed convincing evidence that project design failings were more prominent among projects that received less than satisfactory ratings. The supporting analysis identified three main contributing factors (i) M&E design; (ii) other unspecified design issues; and (iii) over ambitiousness. Projects with better M&E systems had better outcome ratings.

Similar examples can be seen at DFID. A series of studies were made between 2000 and 2010 into the logical framework, its influence on the quality of objectives and indicators, and the quality of periodic project reviews.⁹

The sorts of problems that emerged included weak indicators, with only 18 per cent of performance indicators measuring outcomes and 64 per cent measuring 'inputs' and 'processes'. Indicators were neither specific nor time bound. Measurement and the evidence base were not well planned and assumptions/risks were not monitorable. The studies indicated that clear targets were associated with better performance.

Attention shifted to the quality of evaluation reports (Box 4) and, like the Wapenhans Report 17 years earlier, a technical review into programme quality ended up raising questions about management and culture.

Box 4 Quality of DFID's Evaluation Reports¹⁰

The Independent Advisory Committee on Development Impact¹¹ commissioned a review to assess the quality of DFID's evaluation reports and assurance systems in early 2009. Noting that there is no agreed way of assessing evaluation quality, the consultants began by developing a methodology for the review. In their review the assessment of quality is viewed as a delicate balancing act of different factors whose importance changes from evaluation to evaluation.

The review drew attention to both strengths and weaknesses in evaluation quality. Some weaknesses could be addressed through action taken by DFID's Evaluation Department, but many were held to be systemic in nature and needed to be addressed by DFID's top management, requiring a significant change in culture. A key overarching problem identified was an unduly defensive attitude to evaluation leading to an overarching recommendation that DFID top management needs to address this head-on.

Many of the detailed recommendations pre-date changes associated with the Business Case approach, calling for evaluability issues to be considered at the planning stage; for training of staff; for strengthening the evidence base that underpins evaluations; and for requiring managers to make a formal response to evaluations.

A theme running through so many of the reflective or analytical documents reviewed for this benchmarking is that while high technical standards and systems are a necessary constituent of effective evaluation and development results, what makes those systems and procedures effective is the organisational culture. That is examined in the next section.

Institutional culture and incentives

Writers on results-based management (RBM) literature emphasise that it is a management strategy rather than a set of technical tools. For RBM to be successful, organisations need to develop and nurture a culture of results where enquiry, evidence and learning are valued as essential to good management.¹²

⁹ See for example: Poate, Derek and Christopher Barnett (2003) Measuring value for money? Evaluation report EV645; Aguilhas (2007) Assessing the quality of DFID's project reviews; Drew, Roger and Rachel Albane (2008) Baseline audit of the state of monitoring and evaluation in DFID

¹⁰ Roger C Riddell (2009) The Quality of DFID's Evaluation Reports and Assurance Systems. IACDI (The Independent Advisory Committee on Development Impact)

¹¹ The Independent Advisory Committee on Development Impact was a short-lived initiative to bring independent scrutiny to DFID's evaluation programme. It was replaced by the creation of ICAI noted above

¹² See for example: Mayne, John (2008) Building an evaluative culture for effective evaluation and results management. Institutional Learning and Change (ILAC) Initiative Brief 20. CGIAR 4p; OECD DAC (2006) Managing for Development Results Sourcebook, Paris

An example of how a comprehensive technical approach to RBM failed to deliver a working system was seen in a review commissioned by the Ministry for Foreign Affairs, Finland (Box 5).

Box 5 Findings and lessons from a review of results-based management in Finland's aid programme¹³

The study summarised its findings under eight conclusions and three broad lessons:

- The ministry did not have a well-functioning RBM system. Tools and procedures, mostly drawn from the European Commission approach to Project Cycle Management are comprehensive and well established but have not ensured that results frameworks and results monitoring are in place.
- There is no overarching strategic results framework and its absence means that the unit of analysis for performance remains at the level of individual projects and programmes.
- Despite clear and comprehensive guidelines, good standards of project design are not consistently applied although good practice examples can be found.
- A Quality Assurance Board is overworked and ineffective at setting and maintaining standards.
- A low priority appears to have been given by managers to monitoring, reporting and evaluation. Most monitoring reports were activity-based or financial and there was little reporting against logframes.
- Comprehensive reports are available for parliament and in the public domain. But the most prominent one, the annual report, does not utilise findings from monitoring and evaluation and does not provide sufficient information for the public to assess the nature and achievements of the contribution being made by Finland.
- Managing for results depends not only on technical methodology, but also on the way the development cooperation programme is organised and managed. Senior managers (directors and above) are clear in their views about the importance of information but that is not followed through into the way people work. The approach is characterised as being risk-averse and there are few examples of results from past experience being used to inform future policy.
- Impending changes in human resources management and information management were both expected to give greater prominence to managing for results and improve provision for incentives and rewards.

Three lessons about RBM arose from that evaluation:

- It is not sufficient to take only a technical approach to measuring results.
- Arrangements need to include how staff are managed in terms of the incentive and accountability systems in place.
- A way forward for development cooperation needs to take into account the multiple roles and agenda of the Ministry for Foreign Affairs.

A key issue is the question of what incentives are most important for staff and managers. Here, recent reviews appear to revisit the findings of Wapenhans, 20 years ago. Box 6 draws on work by OPCS and interviews under this benchmarking study to summarise the experience with quality assurance at the World Bank. Box 7 reports an ongoing study at DFID about the effectiveness of the new Business Case approach. Key findings are highlighted in bold and italic script.

Box 6 Experience with Quality Assurance at the World Bank¹⁴

The Wapenhans Report (see Box 2) led to calls for better design, more focus on supervision, and closer overall attention from management. In 1996 the creation of the Quality Assurance Group (QAG) increased accountability for quality and outcomes, and the following year the introduction of the matrix organisation assigned responsibility for specific aspects of quality to sector departments, country departments and networks. The early 2000s saw the strengthening of regional quality units and learning from the QAG experience. Today the Bank has quality assurance arrangements at project, regional and corporate levels:

In FY09, management undertook a review of QAG to determine whether it was continuing to meet the Bank's needs. The review noted the positive contributions from QAG: it fostered a culture of upstream attention to quality; it had a systematic, robust and transparent approach; it provided senior management with arms-length input on quality assurance developments; and it had wide enough coverage to generate lessons across regions and sectors. Nonetheless, the review concluded that for a variety of reasons, QAG was no longer a good fit for the Bank's purposes. For example, its reports could not be used for accountability or resource

¹³ Poate, Derek and Ann Bartholomew (2011) *Evaluation of the Results-Based Approach and M&E of Finnish Development Cooperation Interventions*. Evaluation report 2011:1 Ministry for Foreign Affairs of Finland, Kopijyv Oy, Jyv skyl 

¹⁴ OPCS – Bank quality assurance framework July 2013.pptx

allocation; and it **was perceived as encouraging risk aversion**, at a time when risk aversion had begun to be seen as a cause of churning and reduced institutional effectiveness.

The Internal Audit Department noted that the Bank had not created any mechanism to take QAG's place – and specifically, that it had not developed indicators for investment lending quality or instituted arms-length secondary review to mitigate problems of realism and candour in self-assessments. It expressed a concern that technical quality reviews were not always as useful as they could be because of the choice of peer reviewers and the lack of a learning loop.

The matrix system, introduced in the 1997 reorganisation, was expected to simplify review processes through clear responsibilities for quality. However, a recent IEG evaluation of the Bank's matrix organisation concluded that the proliferation of quality control layers for lending has increased neither efficiency nor effectiveness in ensuring technical quality. The review found that available tools for quality enhancement, such as peer reviews and quality enhancement reviews, are not being used effectively by sector units. Moreover, the evaluation reported feedback from staff and managers that the Bank's **quality control systems are more concerned with reputational risk than technical quality**. The evaluation found that, in the regions, the responsibility for fiduciary and safeguard functions is vested on the same units in charge of overall operational quality control. The reported result was that fiduciary and safeguards tend to receive much more attention than other aspects of quality. The report recommended strengthening the peer review system and quality enhancement reviews, ensuring appropriate management accountability for quality, and establishing an arms-length quality assurance mechanism for monitoring and reporting to senior management.

Box 7 DFID's review of Business Cases and the programme cycle¹⁵

Since adopting the Business Case approach to project design, DFID has conducted regular reviews through a Quality Assurance Unit. For example, the annual report for 2011/12 noted Business Case processes 'going well' and pointed out some areas of weaknesses: in the Appraisal Case (which includes options analysis), and in the Theory of Change, and Evaluation in the Management Case, that needed attention.

In 2013 the Secretary of State commissioned an end-to-end review of the programme management cycle to establish whether the changes introduced have made DFID more rigorous and effective at delivery, or less able to adapt to the complex and changing environments in which we work. Early findings identified issues with incentive structures and culture; with capabilities and with processes. Key points were:

- Evidence of an emerging culture of **risk aversion**, treating guidance as rules
- Incentives geared towards design rather than delivery and **enhancing personal reputation**
- Programme management skills not valued and in short supply
- Programme management guidance has mushroomed in recent years. It is seen as too long, difficult to find and hard to follow.

This ongoing study is looking at how programming culture and behaviour can be reshaped to improve performance.

Both examples reveal a fundamental tension that has a direct bearing on the scope to introduce a results-based approach to management. Staff are inherently risk-averse and pressure to introduce complex systems tend to be managed to minimise risk to personal reputation. Successful reforms need to find ways of overcoming these entrenched positions.

Conclusion

This benchmarking review leads to several conclusions of direct relevance to Norad. First, that experiences reviewed here, especially from the World Bank but also to an extent from DFID, show that organisational procedures and systems should be seen as dynamic and not static. In other words, there is no single solution but rather systems need to be introduced, used, tested, reviewed and then updated in a rolling cycle.

Second, there are clear examples of good practice that reflect current thinking about managing for results and planning for evaluability that Norad can draw on from the three comparator organisations. All would need to be adapted to Norway's own policies and procedures. At the core are definitions and descriptions of technical standards which need to be supported by guidance material and training. The

¹⁵ See: Quality Assurance Unit Annual Report 2011/12; Better Programme Management: Update on the End-to-End Review of the Programme Cycle May 2013

best examples seen approach the challenges of evaluability and evaluation with a holistic view of how good project design is a first step towards successful outcomes and evaluation.

Third, guidance alone is not enough. Project design, approval and evaluation should be supported by quality assurance that sets standards and assesses performance towards those standards. The criteria for QA need to be coherent with good practice planning for results and evaluability.

Fourth, good technical specifications and effective QA may not be enough. The extent to which they are implemented effectively will be determined by the culture and incentives in the organisation. There is a danger that too many reviews and too much QA can lead to risk-averse behaviour and a concern to protect personal reputation rather than manage for results. How to manage this risk is clearly a topical issue at both DFID and the World Bank.

2. List of references consulted for the study

Danida

Checklist for appraisal.docx (projects less than DKK 5 million)
Concept Note Guidelines for Presentations to the Programme Committee Danida Grant Committee and Council for Development Policy Kapitel.pdf
Danida Evaluation Guidelines EVAL-guidelines.pdf (updated 2012)
Evaluation Programme 2013–2014
Freedom from Poverty – Freedom to Change 2010
General Guidelines for Multilateral Organisations Aug 2007.pdf
General Guidelines Grant Administration Private Organisations.pdf
General Guidelines National NGOs, Aug 2007.pdf
General Guidelines Government, Aug 2007.pdf
Guidance note on exiting final uden annex.docx
Guidelines for agreements on development cooperation.doc
Guidelines for presentation of appropriations and strategy documents final1.pdf
Guidelines for presentation of appropriations and strategy documents REV 2122013.pdf
Guidelines for programme management, Sept 2011.pdf
LC-guidelines 3rd edition October 2007.doc
LGA Danish embassies revision Oct 2009 final incl. multilateral LGA-pdf.pdf
Major changes, 2011.doc
Major changes, October 2009.pdf
Major changes, September 2011.doc
NGO HIV-Aids efforts Guidelines, August 2011.doc
PCRguidelinesJuly20112 red17082011.pdf
PCRtemplateaboveDDK5million03102011.doc
PCRtemplatelessDDK5million2red170811.doc
Technical Note Final 50 Programme Support.pdf (2009)
Template for Process Action Plan for programme support preparation and formulation.pdf

DFID

Annual Review Quality – Quality Assurance Unit Pilot Study Nov 2012
Better Programme Management: Update on the End-to-End Review of the Programme Cycle May 2013
Business Case Approval Flow Chart
Collaboration with HoPs and CPOs on formal Quality Assurance (n.d.)

DFID Programme Management Competence Framework (n.d.)
 DFID/UKES Training Workshop
 Evaluation Handbook 2012
 H M Treasury Green Book. Appraisal and Evaluation in Central Government (2003 updated 2011)
 HM Treasury Magenta Book + supplements
 How-To note: beneficiary participation in monitoring (August 2012)
 How-To note: reviewing and scoring projects (November 2011)
 How-To note: assessing the strength of research evidence (January 2013)
 How-To note: writing a Business Case (December 2012)
 ICAI Workplan 2013–14
 Inputs To The Quality Assurance Unit
 International Development Evaluation Policy 2013
 Mandatory Evaluation Quality Assurance in DFID
 Operational Plan 2011–2015. Evaluation Department, Research & Evidence Division. Refreshed April 2013
 Principles of Evaluation January 2013
 Quality Assurance Unit Annual Report 2011/12
 Quality Assurance Unit Annual Report 2011/12
 Quality Assurance Unit Procedures and Timeframe:
 Quality Assurance: Template for Entry Level
 Quality Assurance: Template for Exit Level
 Reviewers Template for Quality Assurance of Business Cases that Require Formal QA Before Approval
 Revised QAU Reviewers' Checklist-9
 Technical Competency Framework for Evaluation (2011)
 Training course 'Monitoring and reviewing in project cycle management'

World Bank

BP 10.00 Investment Project Financing April 2013
 BP 2.11 CAS June 2005
 CAS Guidelines – OPCS April 2012
 Development Impact Evaluation Initiative
 Draft 2014–2016 IEG Work Programme *R2013–0087.pdf*
 Draft PPAR guidelines (outline) with Annexes_Nov 29 2011.pdf
 Evidence to Policy, a monthly note series on learning – health, education, social protection, water and sanitation and labour
 Guide to Country Programme Evaluation Methodology Annex.pdf
 Guidelines for Global and Regional Programme Reviews (GRPRs).pdf
 How We Measure Results
<http://web.worldbank.org/WBSITE/EXTERNAL/PROJECTS/0..contentMDK:22453640~menuPK:5122355~pagePK:41367~piPK:51533~theSitePK:40941,00.html>
 ICR Evaluator Manual draft_May 10 2012.pdf
 IEG Approach Paper Completeness and Quality Checklist (2012)
 IEG Mandate http://ieg.worldbankgroup.org/content/dam/ieg/aboutIEG/dge_mandate_tor.pdf
 IEG Quality Assurance Process for Approach Papers (2012)
 IEG Quality Guidelines and Checklist for Developing Good Recommendations in Evaluation Reports

IEG Quality Standards for Approach Papers (2012)

Implementation Completion Report

Implementation Status and Results (ISR) Report

Investment Project Financing – Preparing the Project Appraisal Document (PAD) OPCS April 2013

LIST OF CORE SECTOR INDICATORS, July 2012 (updated July 2013 to 26 sectors and themes)

OP 13.60 M&E (note no BP) June 2007 (Revised March 2012 & April 2013)

OP/BP 10.00 Investment Lending, June 1994

OP/BP 10.00 Investment Lending, June 1994 (revised up to April 2013)

OP/BP 13.05 Project supervision, July 2001 (revised August 2004)

OP/BP 13.16 CPPR Sept 1994 (revised March 2012)

OP/BP 13.55 ICR July 1999 – Dec 2011 (revised April 2012)

OP/BP 8.60 Development Policy Lending, February 2012 (revised April 2013)

OP/BP 9.00 Program-for-Results Financing, February 2012

Operational Policy Manual Issues and Prospects for Reform Approach Paper 2011

PowerPoint: Bank Quality Assurance Framework July 2013.pptx

Results Framework and M&E Guidance Note. OPCS April 2013

Strategic Impact Evaluation Fund

Terms of Reference for the Committee on Development Effectiveness of the World Bank Board of Executive Directors, which oversees IEG

3. Comparator agency assessment framework

Evaluation Practice in Intervention Management	Approach developed by Itad/CMI drawing on UNEG-DAC material	Comprehensive practice	Moderate practice	Slight practice	Absent	Comments
Design and development of an intervention	<i>Theory of change and indicators</i>	Clear guidance is given about the need to base designs on evidence and use of results, and to construct a logic model or theory of change	Guidance for planners makes reference to some sort of logic model and the selection of indicators	Minimal reference to results		
	<i>Evaluability</i>	The question of evaluability is examined, with provisions for baseline data and consideration of planning for future evaluation	Guidance makes some reference to baseline and evaluation, but not a binding requirement	Minimal reference to evaluability		
	<i>Appraisal and approval</i>	The organisations' procedures include review of the evidence to support the intervention design and arrangements for an evaluation plan	Some consideration is given to assessing evaluation arrangements during appraisal	Minimal guidance about evidence and evaluation		
	<i>Quality assurance</i>	Provision exists for arms-length QA process to review (inter alia) theory of change, evidence base, need for evaluation and economic analysis	QA in place but less complete coverage of issues	Little or no arrangements for QA		
Follow-up	<i>Monitoring, reporting and reviews</i>	Arrangements include regular reporting against structured performance criteria and periodic independent reviews no less than one at mid-term	Regular reporting including some reference to evaluation criteria	Regular reporting but without systematic use of performance criteria		
	<i>Completion</i>	Arrangements include provision of a completion report by self- or independent evaluation, to include assessment against OECD-DAC evaluation criteria and others as appropriate	Requirement for an end of implementation report, but without clear evaluation criteria assessment	Inconsistent requirement for completion reporting		
Evaluation	<i>Self and independent evaluation</i>	Arrangements include a coherent model either of independent evaluation or a structured combination of self-evaluation and independent validation; plus thematic or cross-cutting studies	Structured approach with some reference to implementation reporting	Unstructured, ad hoc plans for evaluation of interventions		

Annex 6 Methodology and analytical framework

This annex describes the research strategy and methods, with references to supporting material in other annexes. First, 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.

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 terms of reference) 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.

The methodology for the study has three main components, described here in turn:¹⁶

- Assessment of grant management processes (1)
- Assessment of EVAL (2)
- A desk review of systems and procedures in three comparator agencies (3)

1. Assessment of grant management processes

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

- **Step 1** – Map grant management processes and systems
- **Step 2** – Assess the quality of grant management processes and systems for ensuring evaluability
- **Step 3** – Assess the quality of results-based management training for staff
- **Step 4** – Assess the practices of evaluability and results measurement across a sample of 20 grant-funded interventions
- **Step 5** – Survey a wider sample of staff to test emerging theories
- **Step 6** – Validate the findings with key stakeholders

Step 1: Mapping of grant management systems and processes

Grant management in both Norad and the Ministry of Foreign Affairs (MFA) is governed by a multiple set of processes and guidance, some of which have been redesigned and re-issued during the period covered by this study. We started during the Inception Phase by collecting and sorting all documents related to grant management and the approach to results measurement in grant management. We then developed process maps to illustrate key steps. The maps are reproduced as Section 1 in Annex 3.

¹⁶ Full details can be found in the Inception Report for this evaluation

Step 2: Assess the quality of grant management processes and systems for ensuring evaluability

After developing the process maps we analysed their comprehensiveness and quality for ensuring results measurement to identify any gaps. We developed a simple checklist which was informed by the findings from the review of comparator agencies (described below) and reflects current practice within other aid organisations. A summary of the review appears as Section 2 in Annex 4

Following the mapping and evaluability process assessment, we interviewed staff from the Grant Management Unit in the MFA (Tilskuddsforvaltningsenheden rediger seksjons) and staff in AMOR to check our understanding of processes and its content and to clarify points of detail. We examined all processes and no sampling was involved.

Step 3: Assessing the quality of results-based management training for staff

The approach to staff training on results measurement was examined in four ways. The courses held by the Foreign Service Institute (UKS) were reviewed, and statistics were obtained of attendance and participants' immediate post-course assessments of the training. Next, materials used for the training were reviewed for scope and content with a particular focus on to what extent it provided support in key areas of evaluability such as: developing theories of change, formulating indicators, conducting evidence reviews to inform programme design and appraising results measurement systems. Clearly, quality of delivery is important. We interviewed one former EVAL staff member who had attended a course prior to an overseas posting to get an impression of delivery, and the emphasis being given to different aspects of grant management. Third, we interviewed AMOR staff who are responsible for the content of the results and risks management parts of training courses, and often do the delivery themselves. Last, we questioned a sample of staff about the training they had received and their perceptions about their competency in dealing with results measurement.

Step 4: Assess the practices of evaluability and results measurement across a sample of 20 grant-funded interventions

We drew a random sample of 20 project and programme grants that were started and completed during the period 2008–12. Guidance on the sample size was given in the terms of reference for the assignment, and was based on the resource implications and practicability of securing necessary documents. A description of the sampling procedure is in Box 2. The grants were reviewed using a composite checklist that combined: **policy compliance**, used to assess the extent to which key mandated actions have been followed in a grant-making process; **evaluability**, looking at programme design and arrangements for monitoring and evaluation; and a **quality assurance** checklist, which was applied to those projects for which there was a mid-term review or completion review available. A copy of the checklist is included in Section 5 in Annex 4.

Following these desk reviews, telephone interviews were held with the responsible staff from Norad and the MFA, and with representatives from the grant-receiving partner, where it was possible to make contact. Interviews were conducted with four partners.

Box 1: Grant assessment sampling procedure

Working with a database of grant disbursements provided by the Statistics Unit in the Department for Quality Assurance Section, a population of eligible grants was created by selecting all agreements/projects completed between 2008 and 2012 and removing any other agreements/projects, including events with DAC main sector codes 910, 930 and 998/20, from the parent population.

The new parent population was then grouped according to the top six sectors listed in the Norad database.

A total of 26 agreements/projects were drawn proportional to the volume of lending in each group from the six sector sub-groups (i.e. five from the 'economic development and trade' sub-group, four from 'emergency

assistance', etc.). In practical terms this was accomplished by randomly listing the agreements/projects by sector sub-group with the 'Rand' formula in Excel. Subsequently, the first five randomly listed agreements/projects from the 'economic development and trade' sub-group were selected, followed by the first four randomly listed agreements/projects from 'emergency assistance', etc. As each selection was made, all disbursements were reviewed to ensure that the total disbursements were above the cut-off of NOK 10 million.¹⁷

The distribution of grants across regions was also managed proportionally (eight agreements/projects not geographically allocated; six Africa; two from Asia, etc.) by inspection as the grants were sampled. This process was repeated until all contingents by region and categories of disbursement units were filled.

Last, the sample was checked to make sure it included grants managed by departments in the MFA, departments in Norad and the embassies.

Six projects were treated as spares in case practical problems arose about availability of data. This did occur and the resulting sample comprised 20 grants (see Annex 3 Section 3 for a list of the sampled grants).

Step 5: Surveying a wider sample of staff

To supplement the data collected through the in-depth assessments of grant management processes we conducted a **staff survey**. This enabled us to reach out to a wider network of programme staff dealing with grant management. The survey was implemented towards the end of the main study phase so as that it could be used as a means of testing and validating the findings that emerged from the in-depth assessments with a wider sample of staff.

The survey was implemented using the proprietary software 'Survey Monkey'. A copy of the survey instrument can be found in Section 5 in Annex 3. The survey was distributed to staff in the MFA, Norad and those embassies dealing with development cooperation. Identification of recipients was done with the assistance of EVAL staff. Job titles do not necessarily provide an accurate guide to the extent to which staff actually deal with grants, so the sample was drawn on a more inclusive basis, realising that some recipients would not work on grant management. The covering email asked staff who did not deal with grants to ignore the survey. Table 1 below summarises the distribution and response statistics. In view of the inclusive target population, we believe the percentage response underestimates the true proportion of grant managers who answered the survey.

The survey was designed in two parts. In the first, a short, Likert-scale opinion survey was designed to be fast and easy to complete to encourage a wide take-up.¹⁸ The second part examined results measurement and grant management in more detail. Fewer recipients continued to complete the second part. A total of 157 people opened the survey to look at the questions. Of these, 126 continued to complete Part 1 (Question 4) and 97 continued to complete the remainder of the questions. The extent of selection bias in the results is discussed later in this chapter.

¹⁷ During the Inception Phase, the question was raised about our decision to adopt a cut-off only looking at grants above NOK 10 million, arguing that most grants were below this size. Looking at grant statistics for 2012, it is true that 70 percent of project and programme grants are sized at NOK 10 million or below. But these account for only 18 per cent of the monetary value. In consideration of the need for accountability of a results focus, we consider that the larger grants are of greater importance. Moreover, staff argue that larger grants receive more attention to their results, therefore lessons about how they are treated should convey more useful information about actual practice and potential for improving a focus on results.

¹⁸ The approach we adopted was a bipolar scale, measuring either a positive or negative response to a statement following a 'forced choice' structure, but with a 'Don't know/Not applicable' option.

Table 1: Staff survey data

Recipient	No. of staff to whom the survey was sent (100%)	Opened survey	Completed Part 1	Completed Part 2
MFA headquarters	210	46 (21.9%)	33 (15.7%)	23 (11%)
Norad headquarters	144	50 (34.7%)	43 (29.9%)	34 (23.6%)
Norwegian embassies	123	61 (49.6%)	50 (40.7%)	40 (32.5%)
Total	477	157 (32.9%)	126 (26.4%)	97 (20.3%)

The survey allowed for self-selection, which inevitably introduces the problem of selection bias in the results. The main validator we have for the survey are the findings from a set of focus groups conducted as part of our validation process (see Step 6). For most questions, the findings from the focus groups align with those from the survey. Where differences have emerged, these are discussed in the text of the report.

Step 6: Validation

The original plan was to validate the findings of the grant assessments after they were completed. Delays in obtaining grant documents during the July holiday period and the need to extract some documents from archives held at embassies meant that the validation meetings took place when about half the grants had been reviewed.

We held three focus group meetings in Oslo, separately for clusters of staff from different departments in the MFA and Norad. The feedback presentation used for those meetings is in Section 6 of Annex 3, together with summaries of an exercise used in the discussion, designed to test questions for the staff survey, and comments from the focus groups and individual interviews held to follow-up grant assessments.

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 to what extent 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. Participation in the staff survey was voluntary because the survey was transmitted directly to individuals by email.

In both cases, forms of selection bias are likely to have occurred, compared with a random sample. Such a bias may result in the subjects in the sample being unrepresentative of the population of interest. Common types of selection bias include volunteer or referral bias, and non-respondent bias. Volunteer or referral bias occurs because people who volunteer to participate in a study (or who are referred to it) are often different from non-volunteers/non-referrals. This bias usually, but not always, favours a greater awareness of the issues because volunteers tend to be more motivated and concerned about the topic. Non-respondent bias occurs when those who do not respond to a survey differ in important ways from those who respond or participate. This bias can also work in either direction.

Likert scales may be subject to distortion from several causes. Respondents may avoid using extreme response categories (central tendency bias); agree with statements as presented (acquiescence bias); or try to portray themselves or their organisation in a more favourable light (social desirability bias).

2. Assessment of EVAL: methodology and data

2.1 Overview

Here we describe the **methodology and data** used for assessing whether EVAL puts sufficient emphasis on results measurement in the planning, commissioning and quality assurance of evaluations (Hypothesis 1) and if the evaluators commissioned by EVAL have the appropriate competencies for results measurement (Hypothesis 2). A key reason for putting forward these hypotheses was 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'* (ToR p.1). We therefore paid particular attention to this level of results (outcomes and impact) although a complete testing of the hypotheses necessitated the inclusion of results at the output level as well.

The evidence used in the assessment was drawn from six main sources: (1) a quality assessment of six EVAL commissioned evaluation reports; (2) an assessment of the process associated with the evaluation reports; (3) interviews with the EVAL staff involved in the reviewed evaluations; (4) a review of 26 EVAL Evaluation Report terms of references; (5) a comparative analysis of the evaluation functions from other development agencies; (6) a self-assessment of the competencies of the evaluation consultants.

From our theory of change for the EVAL assessment we postulated that a sufficient focus on results measurement in planning, commissioning and quality assurance together with hiring qualified consultants would lead to high-quality evaluations documenting results in accordance with international best practice (as agreed in the DAC Standards for Development Evaluation).¹⁹ A key element in our methodology has therefore been to assess the quality of the EVAL Evaluation Reports. This approach was supplemented with an assessment of whether EVAL has been sufficiently results focused in planning and managing the evaluations since high-quality evaluations could still be produced even without any emphasis on results measurement from EVAL's side (for example if good evaluators ensure high-quality results evaluation irrespective of EVAL's inputs and actions). The Evaluation Reports constitute a key output from the department with regards to results measurement and have therefore been central in the assessment.

We applied **backward induction** to reveal the critical factors that have led to the quality of each of the final EVAL Evaluation Reports. Backward induction is a method where reasoning backward in time from the final outcome (or output) to the previous step all the way to initiation can reveal the sequence of actions that led to the result of interest.²⁰ Based on an initial assessment of the process²¹ from evaluation planning to the final evaluation report, the backward induction was started from the key outputs and followed four steps:

Step 1 was to **select six EVAL Evaluation Reports with a results focus** out of the 37 reports completed in the past three years based on a pre-determined set of criteria (see subsection 2.2 below) and assess to what degree each of them have been conducted in accordance with the key elements of the OECD/DAC Quality Standards for Development Evaluation and indicators derived from our evaluation questions and sub-questions (see subsection 2.4 below, and the quality assessment checklist for evaluations, **Annex 4**). Each study was rated as either: very good, good, sufficient or poor.

Step 2 was to investigate what factors have contributed to the rating during the process from when the consultants start working on the evaluation, to the final round of commenting, revising and completing the evaluation report. The data used at Step 2 was drawn from a number of sources: documentary evidence such as evaluation decision documents, correspondence between EVAL staff and the evaluation team, comments on the inception and draft evaluation report etc. and phone and

¹⁹ The full theory of change can be found in the Technical Proposal submitted for this study

²⁰ 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

²¹ Simplified version based on "The instructions for evaluation activities in the Norwegian Aid Administration", see <http://www.norad.no/en/evaluation/attachment/393700?ts=138d70a555a&download=true>

face-to-face interviews with EVAL staff involved in the evaluations and information from members of the commissioned evaluation teams. The interviews were used to explore stakeholders' different experiences of the process and their perspectives on challenges and what contributed to the quality of the report.

Step 3 was to assess how the specification of the ToR influenced the quality of the Evaluation Report. The main steps leading from the specifications in the ToRs to the final report was assessed, including the competencies of the evaluators in order to pin down the specific reasons for the rating. In particular, we assessed whether the identification of results was clearly spelt out as a main task of the evaluation, the use of definitions around results, the specification of appropriate methodologies, the requirements for methodological competencies and the process of developing the ToR.

Step 4 was to link the findings from the previous steps with an assessment of the EVAL management of the evaluations and to assess some key features of the ToRs for the Evaluation Reports more generally and conduct an overall assessment of design of evaluations. For this purpose we combined in-depth analysis of the six ToRs of the selected Evaluation Reports with a review of all the 26 ToRs in our eligible sample (see subsection 2.2) and assessed some key design features that were likely to influence evaluation quality. For this assessment, we relied on all the collected evidence and the adjoining analysis, in particular interviews with involved EVAL staff and a process review of the archived documentation of the main correspondence. At this step we also assessed the evaluator's competencies based on a competencies framework (see Annex 4) and an assessment of how the ToR's specified the competencies required for conducting the evaluation. Details of this are elaborated in subsection 2.5 below.

The EVAL assessment also involved a review of the handbooks and the documentation relevant for evaluation work of EVAL, in particular the OECD/DAC documents listed in the "Instruction for evaluation activities in the Norwegian Aid Administration" and related documents. (DAC Quality Standards for Development Evaluation, 2006, DAC Principles for evaluation of development assistance, 1991, DAC Glossary and key terms in evaluation and results-based management, Guide to conducting evaluations, Ministry of Finance, 2005, White Paper 35, 2003–04, Ministry of Foreign Affairs, Rules for financial management, Ministry of Finance, etc.)²²

2.2 Sampling strategy for selecting EVAL evaluation reports

In selecting the Evaluation Reports we used a purposive sampling method where we chose evaluations that cover both a mixture of sectors and types of results to be documented (outputs, outcomes and impacts). The reason for covering several sectors is that results measurement can be more challenging in certain sectors than in others and require evaluators to have different levels of evaluation competencies. Similarly, documenting outputs may be much more straightforward than documenting outcomes and impacts, again imposing different requirements (see inset text box). It was useful therefore to explore the extent to which these factors affect the quality of the final evaluation reports, the quality of the design and management provided by EVAL and the ability of the commissioned evaluation teams to design and implement an appropriate methodology to measure results.

The full population of evaluations for our study consists of all 37 reports in the Norad EVAL Evaluation Report series since 2010 as published on the Norad web site accessed in April 2013. Our concrete selection criteria were:

- Evaluation Reports that are not intended to discuss results are excluded (Reports 9/11 and 7/12)
- Evaluation Reports where Itad, CMI or any of the team members have been involved are excluded to avoid potential conflict of interest (Reports 2/10, 12/10, 13/10, 18/10, 5/11, 6/11, 2/12 and 5/12)
- EVAL Reports that are compilation of other EVAL Reports published in the same series (Report 3/10)
- Three reports will cover outputs, and three reports will cover outcomes and/or impacts
- Each selected report will cover a different sector

²² See documentation at <http://www.norad.no/no/evaluering/h%C3%A5ndb%C3%B8ker-og-referansedokumenter> (in Norwegian).

- Each selected report will cover a different aid modality

Applying these criteria produced a list of 26 reports that were candidates for review. **Table A6-1** presents the overview of the application of the initial selection criteria on the full population.

Table A6-1: Initial sample classification

Report	By Itad/CMI	Results	Aid type	Sector	Selection
1.10	No	Output	Bilateral	Governance	
2.10	CMI				Excluded
3.10	No	Output, outcome and impact	Bilateral	Private sector development	Excluded
4.10	No	Output, outcome and impact	Bilateral	Private sector development	
5.10	No	Output, outcome and impact	Bilateral	Private sector development	
6.10	No	Output, outcome and impact	Bilateral	Private sector development	
7.10	No	Output	Humanitarian aid	Peace, reconciliation and democracy	
8.10	No	Output	Aid to civil society organisation	Corruption	
9.10	No	Output, outcome and impact	Bilateral	Health	
10.10	No	Output	Multilateral	Democratic development	
11.10	No	Output	Multilateral	Human trafficking	
12.10	CMI			Climate	Excluded
13.10	CMI			Climate	Excluded
14.10	No	Output Impact	Bilateral	Climate	
15.10	No	Output Impact	Bilateral	Climate	
16.10	No	Output Impact	Bilateral	Climate	
17.10	No	Output Impact	Bilateral	Climate	
18.10	CMI			Climate	Excluded
1.11	No	Output Outcomes	NGOs	Health, education, capacity development	
2.11	No	Process Output	Support to research in Norway	Research on Norwegian development assistance	
3.11	No	Output	Bilateral	Culture and sports	
4.11	No	Output	Anti-corruption assistance	Corruption	
5.11	CMI				Excluded
6.11	Itad				Excluded
7.11	No	Output	Bilateral	Human rights	
8.11	No	Output	Multilateral	Trade	
9.11	No	Not an evaluation			Excluded
10.11	No	Output Outcome Impact	Bilateral	Health	
1.12	No	Outputs Outcomes ("to the extent possible" impacts)	All aid modalities	Aid to disabled people	

Report	By Itad/CMI	Results	Aid type	Sector	Selection
		mentioned among objectives)			
2.12	CMI				Excluded
3.12	No	Impacts	Bilateral Multilateral Multi-bi NGOs	Public sector capacity building, Civil Society Organisations strengthening	
4.12	No	Output	Multilateral	Health	
5.12	CMI				Excluded
6.12	No	Outputs Outcomes	Bilateral Multilateral CSOs Other donors	Public sector development (petroleum management: institution building, framework development)	
7.12	No	No	Civil society support	M&E	Excluded
8.12	No	Outputs Outcomes	Aid administration	Administration	
9.12	No	Output Outcome Impacts	Bilateral	Agriculture	

2.3 Evaluation reports selected for assessment

From the 15 evaluations with a ToR for identifying outcome/impacts, we randomly selected the following evaluation reports:

- 6.10 – Evaluation of Norwegian Business Related Assistance: Uganda Case study (Sector: Economic development and Trade – Private sector development)
- 16.10 – Real Time Evaluation of Norway's International Climate and Forest Initiative. Country report: Indonesia (Sector: Environment and energy)
- 10.11 – Evaluation of Norwegian Health Sector Support to Botswana (Sector: Health and Social Services).

These cover three of the six sectors with the largest disbursements in NOK in the period under scrutiny (see Table 5, Annex 6). Once an evaluation was selected, the other evaluations with the same sectoral focus were excluded in order to get the spread across sectors.

The same selection process applied to the evaluations with a ToR for identifying outputs gave the following result:

- 7.10 – Evaluation of Norwegian Development Cooperation with Western Balkans (Sector: Emergency assistance)
- 4.11 – Contextual Choices in Fighting Corruption: Lessons Learned (Sector: Governance)
- 4.12 – Evaluation of the Health Results Innovation Trust Fund (Sector: Health – multilateral aid)

These cover three of the seven sectors with the largest disbursements in NOK in the period under scrutiny.²³

2.4 Methodology for assessing the quality of evaluations

A development evaluation of sufficient quality assesses the intervention – its design, implementation and results (outputs, outcomes or impacts) – in a systematic and objective manner. Based on the most relevant (for our study) criteria from the DAC Quality Standards for Development Evaluation

²³ For details, see the Inception Report table 5 in Annex 6.

together with basic principles of sound programme evaluation,²⁴ we operationalized this through four main criteria that indicate whether an evaluation report is of sufficient quality. Our methodology for assessing the quality of the six EVAL reports consists of assessments based on four criteria:

1. Application of a credible **counterfactual**
2. Identification of the **logic** between the intervention and its aims (i.e. expected results)
3. Application of **suitable data** to test the theory and measure results
4. Assessment of the **reliability** of their evaluation and **validate** the results.

The details of these four criteria are elaborated below. To systematise the assessment across evaluation reports we used a **quality assessment checklist** containing the four criteria and the relevant sub-criteria (see Annex 4). Each report was assessed along the same parameters and scored based on a scale of: poor, sufficient, good, very good.

Application of a credible counterfactual – The main criterion for a high-quality evaluation of results is that it identifies a credible counterfactual, which could be of varying degrees of rigor, and measures the results against this counterfactual. The counterfactual is the situation or condition that hypothetically would have prevailed if there was no intervention. Since this situation or condition is unobservable empirically – there is never a situation where an intervention was implemented and one where it was not implemented on the same population simultaneously – a main task for the evaluators is to create a credible counterfactual which the intervention can be measured against. This is a necessary requirement for an evaluation of a sufficient quality that assesses outcomes and impacts. For output evaluations it can be sufficient to define outputs as the products (or services or capital goods) that have been produced by the intervention – although the principle of identifying the attribution of the intervention also applies in principle when it comes to outputs as results.²⁵ An evaluation of sufficient quality describes and explains the empirical evaluation methodology and its application in assessing outcomes and impacts, attribution and/or contribution to results are explained.

Identification of the logic between an intervention and its aims – A second criterion entails to reveal the theory, or theories, of the causal links between the intervention and its expected results. All development interventions include implicitly or explicitly a theory that explains how the financed activities are supposed to lead to the outcomes or impacts of interest. An evaluation report of sufficient quality identifies such theories, either in the form of log-frames, result chains,²⁶ intervention logic or in hierarchies of goals. Identifying underlying assumptions and factors affecting the results should also be part of this assessment. Good evaluations derive testable predictions from the theory and assess to what extent the theory holds in practice. Given that EVAL commissioned evaluations assess specific sectoral, thematic and/or policy issues and are not focused on individual interventions, it could be argued that those evaluations are unlikely to examine a specific theory of change for any one intervention. Even so, an evaluation should still take into account the theory of change of the underlying interventions and present a clear logical framework for structuring the data collection and analysis and identifying key evaluation questions.

Application of suitable data to test the theory and measure results – An evaluation report of sufficient quality makes use of data that is suited to the task of measuring/assessing the results. Moreover, such a report describes the sources of information used (e.g. survey data, documents, respondents, administrative data, literature) in sufficient detail so that the adequacy of the information can be assessed. It is necessary to explain and justify the selection of data (e.g. qualitative and/or quantitative data, case studies). Limitations of the representativeness of the samples must be identified and implications discussed.

²⁴ See Ravallion (2001)

²⁵ To exemplify: Assume the intervention is to build a primary school. If it is built in an area where the Government was going to build a primary school anyway, then there can be two ways of assessing the output: (1) the output is the school since that was what was produced, and (2) the intervention did not produce any results since the school would have been built anyway (counterfactual analysis).

²⁶ We use the OECD-DAC definition: "Results Chain: The causal sequence for a development intervention that stipulates the necessary sequence to achieve desired objectives beginning with inputs, moving through activities and outputs, and culminating in outcomes, impacts, and feedback. In some agencies, reach is part of the results chain." <http://www.oecd.org/development/peer-reviews/2754804.pdf>

Assessment of the reliability of their evaluation and validate the results – The final criterion concerns the validation and reliability of the evaluation. A sufficiently good report acknowledges any constraints encountered and how these have affected the evaluation, including the independence and impartiality of the evaluation, it justifies methodological choices and explains limitations and shortcomings. Moreover, it explains how information sources have been cross-checked and critically assesses the validity and reliability of data.

This assessment based on the four criteria also provided data for informing both our discussions about the competencies of the commissioned evaluators and the effectiveness of EVAL's quality assurance.

2.5 Methodology for assessing evaluator competencies

The in-depth assessment of the selected evaluations provided some indications of the evaluators' competencies. In order to triangulate these findings, and to provide a broader picture than what emerged from the six evaluations, we carried out a self-assessment with the evaluators involved in the selected sample (i.e. the six evaluations selected in section 2.3). Our assessment of the evaluator's competencies was triangulated with a self-assessment skills profile we asked evaluators to conduct.²⁷ The skills profile that formed the basis of the self-assessment is presented in Annex 4.

The purpose of the evaluator competencies survey was to generate data on evaluator's self-assessment of their skills. The survey used proxy questions about the extent to which consultants were actively engaged in evaluation-related activities (such as being a member of a professional association, attending a training course, publishing an article or book, or holding a formal evaluation post of employment), and other questions about their familiarity with a wide range of theories and methods or tools.

The survey was sent out to the consultancy firms or team leaders from the six evaluations reviewed for this study, plus those consultants for whom email addresses were available, 14 in total. All together 42 consultants were identified as having worked on the studies but many of these had technical roles rather than being evaluation experts. The companies were invited to distribute to all team members but some chose only to send to the evaluators. A total of 16 responses were received.

2.6 Validating the findings

The findings from the assessment of EVAL were validated in two steps. First through a workshop with EVAL staff to review the assessments of each of the six evaluation processes and inquire into the reasons for the results, and then second through an extensive review process of the Draft Final Report. As well as checking for inaccuracies, the workshop and review process was used to clarify and elaborate on some of the issues and critically assess the findings, conclusions and recommendations of the report.

2.7 Document review

A range of the most relevant documents for the assessment have been reviewed including the handbooks and the documentation relevant for evaluation work of EVAL, in particular the OECD-DAC documents listed in the *Instruction for Evaluation Activities in the Norwegian Aid Administration*, the instruction itself and the documents supporting or detailing the implementation of the instruction. References to the reviewed documentation are found throughout the report and not repeated here.

3. Study of comparator agencies

By a combination of guidance in the terms of reference 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. In addition, some information was reviewed from an assessment of results-based management in the Finnish Ministry for Foreign Affairs.

²⁷ During the inception phase, consideration was given to competency self-assessment but this approach was abandoned following advice from the reference group about the difficulties of avoiding biased assessments. Instead, a skills profile was used.

Data for the comparisons was 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.

Analysis was undertaken against a framework drawn up for this study, building on the approach to Norms and Standards for evaluation in the UN and DAC. Criteria were developed to assess how evaluation is embedded in intervention management, looking at policy and guidelines for the following stages in the programme cycle:

- **Design and development of an intervention**
 - Appraisal and approval
 - Quality assurance
- **Implementation**
- **Review and evaluation of the intervention**
 - Completion
 - Management follow-up

For ease of comparison between agencies, a scoring scale was developed for each dimension of the framework that defines three levels of quality: **comprehensive**, where documents indicate the organisation follows the criterion closely; **moderate**, where the general approach follows the criterion but with some shortfalls on coverage or details; and **slight**, where this criterion is dealt with much less fully than the stated standard. Details of the documentation reviewed for the three comparators, and assessments of Norway together with the three comparators are included in Annex 4, together with the review frameworks.

4. Analytical framework

One challenge with this assignment is how to draw rigorous conclusions from a small number of enquiries based around mainly qualitative information. The approach we have taken mirrors, in some ways, the orientation advocated by realist evaluators. Realist evaluation tests a theory of change in the context of institutional structures, cultural norms and practices and external influences that affect the intervention. It 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 taken to mean the steps and procedures adopted by the MFA and Norad to ensure the quality of grants being approved for financial support. Thus, the intervention logic can be represented by the development cycle, as it is presented in Chapter 2 of the current version of the MFA's Grant Management Manual. 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.

Our analysis does not follow a realist synthesis in a pure form, but draws on key characteristics that distinguish it from traditional systematic reviews:²⁸

- The appraisal of evidence is purposive and theoretically driven; our articulation of 'theory' is based on published good practice²⁹ about the need for grants to have an intervention logic with a stated hierarchy of goals, indicators to measure progress and information about risks and assumptions.
- Multiple types of information can be included; we draw on findings from surveys that can be quantified, together with reviews of document quality; and information gathered from interviews.

²⁸ See, for example Rycroft-Malone *et al.* Implementation Science 2012, 7:33

<http://www.implementationscience.com/content/7/1/33>

²⁹ See for example: European Commission (2004) Aid Delivery Methods. Volume 1, Project Cycle Management Guidelines. Brussels; IFRC (2010) Pproject/programme planning Guidance Manual. Geneva; UNDP 2009 Handbook on planning, monitoring and evaluating for development results. New York

- The findings from our analysis focus on explanations about why (or why not) the desired approach to good quality leads to grants that can be evaluated.

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