



2014:13

Sida Decentralised Evaluation

Dorian LaGuardia
Sabine Becker-Thierry
Florence Bonoll

Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks

Final Report

Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks

**Final Report
January 2014**

**Dorian LaGuardia
Sabine Becker-Thierry
Florence Bonoll**

Authors: Dorian LaGuardia, Sabine Becker-Thierry and Florence Bonoll

The views and interpretations expressed in this report are the authors' and do not necessarily reflect those of the Swedish International Development Cooperation Agency, Sida.

Sida Decentralised Evaluation 2014:13

Commissioned by Sida, Department for Organisational Development

Copyright: Sida and the authors

Date of final report: January 2014

Published by Citat 2014

Art. no. Sida61715en

urn:nbn:se:sida-61715en

This publication can be downloaded from: <http://www.sida.se/publications>

SWEDISH INTERNATIONAL DEVELOPMENT COOPERATION AGENCY

Address: S-105 25 Stockholm, Sweden. Office: Valhallavägen 199, Stockholm

Telephone: +46 (0)8-698 50 00. Telefax: +46 (0)8-20 88 64

E-mail: info@sida.se. Homepage: <http://www.sida.se>

Table of Contents

Table of Contents	6
Abbreviations and Acronyms	8
Preface	9
Executive Summary	10
1 The Evaluated Intervention	15
2 Methodology	18
2.1 Sources	19
2.1.1 Assessment of Evaluation Reports	19
2.1.2 Qualitative Data	21
2.1.3 Quantitative Survey.....	21
2.1.4 Best practices & Leading Research	22
2.1.5 Review Team’s Expertise	22
2.2 Using Qualitative Data Graphs	23
3 Findings	25
3.1 Operationalisation (Evaluations).....	25
3.1.1 Number & Type of Evaluations	25
3.1.2 Evaluation Team Composition	29
3.1.3 Implementation.....	31
3.1.4 Guidance & Strategy	35
3.2 Quality (Evaluations)	39
3.2.1 Overall	39
3.2.2 Changes in Quality.....	51
3.2.3 Service Partner Development	58
3.3 The Model (Evaluations).....	65
3.3.1 Functionality (procurement, partnership & quality).....	65
3.3.2 Comparative Analysis	66
3.3.3 Market Implications	68
3.3.4 Knowledge Management & Institutional Learning	77

3.3.5	Costs & Efficiencies	79
3.4	Advisory Services on Results Frameworks (Operationalisation)	82
3.4.1	Overview	82
3.4.2	Number, Type & Costs of Advisory Services	82
3.4.3	Consultant	85
3.4.4	Sida Organisation	86
3.5	Advisory Services on Results Frameworks (Quality)	87
3.5.1	Relevance	88
3.5.2	Effectiveness & Impact.....	90
3.6	Advisory Services on Results Frameworks (Model)	94
4	Conclusions	96
5	Lessons Learned	100
6	Recommendations.....	103
7	Annex.....	106
7.1	Terms of reference	106
7.2	Inception Phase Report (as a separate Annex)	119
7.3	Data & Analysis Report (as a separate Annex)	119
7.4	Comparative Procurement Models	120
7.5	Reports Assessed as Part of this Review	125
7.6	List of Persons Interviewed During Field Phase	130
7.7	Evaluations Conducted under Framework Agreement	134
7.8	Qualitative Data Interview Protocols.....	145
7.9	Evaluation Questions.....	150
7.10	Bibliography.....	152

Abbreviations and Acronyms

Advisory Services	Advisory Services on Results Frameworks
ASEGE	Are Sida Evaluations Good Enough? An Assessment of 34 Evaluation reports.
CRM	Customer Relationship Management
EBRD	European Bank for Reconstruction & Development
EQs	Evaluation Questions
EU	European Union
Evaluations	Reviews, evaluations and related activities conducted under the Framework Agreement
FA; Framework Agreement	The Framework Agreement for Sida Reviews, Evaluation and Advisory Services on Results Frameworks
HQ	Sida Headquarters
IBRD	The International Bank for Reconstruction and Development
IEG	Independent Evaluation Group of the World Bank
Indevelop	Indevelop AB; the first ranked service provider within the Framework Agreement
IPR	Inception Phase Report
MT Reviews	Mid Term Reviews
NONIE	The Network of Networks for Impact Evaluations
OECD DAC	Organisation for Economic Co-operation and Development: Development Assistance Committee
Orgut	The Consortium ORGUT Consulting AB; the third ranked service provider within the Framework Agreement
Partner Organisations	Organisations that receive Sida support and who have been the subject of evaluation and advisory services under the framework agreement
PO	Sida Programme Officer
QA	Quality Assurance
RBM	Results Based Management
Review	Mid Term Review of The Framework Agreement for Sida Reviews, Evaluation and Advisory Services on Results Frameworks
SEK	Swedish Krona
Sida	Swedish International Cooperation Development Agency
Sida Commissioners	Any Sida or Embassy staff that has commissioned services under this framework agreement.
SIPU	The Consortium SIPU International; the second ranked service provider within the Framework Agreement
SP	Service Partners and/or Providers, mainly those who provide services under the Framework Agreement
ToR	Terms of Reference
TL	Team Leader
UPPIS	Sida's Procurement Unit
UTV	Sida's Monitoring & Evaluation Unit

Preface

The Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks provides a comprehensive review of the complexities associated with a framework agreement and how such agreements can influence the quality of evaluations, reviews and advisory services.

The evaluation team, Dorian LaGuardia, Sabine Becker-Thierry, Florence Bonoll, used different evaluation tools and methodologies to develop a rich evidence based appraisal. This included innovative qualitative evidence assessment, tailor made surveys for different cohorts, and a standardised matrix for evaluating over 80 evaluation reports.

Over 144 individuals were interviewed, at Sida, among partnership organisations, and among the framework's service providers, and others. This comprehensive approach contributes to solid evidence based conclusions and recommendations.

Transtec and the authors would like to thank all those who participated in the evaluation. We recognise that people working in international cooperation and development can be very busy and so we value the time they provided us. We hope that the result honours their contributions.

Executive Summary

The “Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks” (Framework Agreement; FA) serves to reduce procurement costs and to provide a more centralised view of reviews, mid-term reviews, evaluations, assessments, and strategic reviews (evaluations) being done by the Swedish International Cooperation Development Agency (Sida) and Swedish Embassies (Embassies). The Framework Agreement also means to increase the quality of services overall while providing a greater focus on evaluation best practices and standards.¹

The Framework Agreement provides significant costs-savings in the procurement process and is a source of statistical and strategic information about decentralised evaluations. **It is not, however, a sufficient mechanism in and of itself to improve the quality of evaluations** as is intended. This Review recognises the complexity of what constitutes ‘quality’ to different stakeholders. It includes everything from the detail and feasibility of a Terms of Reference, to engagement/utilisation levels during the process, to the report and other outputs, and to how Sida and Partner organisations use these. This Review considers these and other aspects of quality and establishes that there are significant deficiencies in key areas of evaluation, especially methodologies, data collection and analysis, and evidence-based conclusions and recommendations.

For instance, this Review shows that, when considering the overall quality of evaluations, 38% of respondents view quality positively while 42% view it negatively. (Section 4.2.1.1, Graph 11) The assessment of evaluation reports shows a similarly middling level of quality. On a 1 – 6 scale, methodologies overall are ranked at 4.28, data collection at 4.39, and data analysis at 4.18, amongst other metrics. The total component average amongst all metrics is 4.31. (Section 4.2.1.1, Table 4) While levels of engagement/utilisation are relatively positive, 27% of respondents still cite negative issues related to this. (Section 4.2.2.2, Graph 16) Furthermore, this Review demonstrates that the quality assurance process focuses on the drafting of the reports, by and large, and could be significantly strengthened through a focus on the process, methodologies, analysis and evidence based conclusion and recommendations.

¹ For more about the Framework Agreement, its purpose and structure, see Section 2 “The Evaluated Intervention.”

While there are some noted improvements in the Terms of References and how methodologies are described, as well as sufficient quality in engagement/utilisation and the overall drafting of the reports, these are not enough to overcome the risks and issues associated with inadequate methodologies, data analysis, and evidence based conclusions. These issues, central to evaluation, have the potential to provide ineffective conclusions and recommendations and to detract from the results of the evaluations and their utility by stakeholders.

The issue of quality in these key areas calls into question the Framework Agreement's ranked system for awarding contracts. This results in one service provider assuming over 80% of the evaluation assignments and nearly all advisory services. This may be justified if there are sufficiently high levels of quality for all services under the Framework Agreement. This does not appear to be so in the case of evaluations. The absence of high quality in essential elements of evaluation or substantive and widespread quality improvements overall not only has direct implications for partner organisations. It implies that Sida may be diminishing the quality of evaluation services by not leveraging a broader network of evaluation expertise within and beyond Sweden. With one Service Provider (SP) doing most of the work, and not providing the highest quality and innovation in key evaluation methodologies, the innovations and best practices for evaluation remain beyond Sida's reach. It has locked itself in with one SP and is therefore dependent on the quality they provide, with little to no reference to other evaluation practices.

This Review concludes that, **while the Framework Agreement met a need at the time of its implementation and provides ample lessons for moving forward, the ranked system is not an adequate procurement model for promoting quality evaluation services.** At the same time, this Review supports the need for a framework agreement. Direct and comparative analysis indicates that framework agreements are an effective procurement device for public sector organisations. Given this, Sida may be served best by a pre-qualified roster of suppliers who then compete in "mini-competitions" for individual assignments.

Mini competitions would enable Sida and Swedish Embassy staff commissioning evaluation services (Sida Commissioners) to make better informed decisions about evaluation teams and methodologies while still decreasing procurement costs. Mini competitions would support local and international markets for evaluation services, engaging a wider breadth of expertise and providing a more ample channel of intelligence about Sida's decentralised evaluations to this growing community of evaluators. Given rapidly changing standards, tools, and approaches in the field of evaluation, having a procurement mechanism that draws from more local and international service providers is important to Sida's long-term need for quality evaluations.

Sida should maintain elements contained within the Framework Agreement. This includes an insistence on a 'core team' of evaluation experts who can participate in and/or provide comprehensive quality assurance on any evaluation assignment. Sida

should continue to mandate the Framework Agreement's use by Sida Units and continue formal and informal communication amongst Swedish Embassies to ensure that all are aware of its purpose and utility.

Despite limited resources, Sida's Monitoring and Evaluation Unit (UTV) should continue to insist on high quality standards from Service providers and broaden this scope. UTV has a distinct influence on the first ranked Service Provider, as exhibited in relation to engagement/utilisation and methodologies. This is not surprising but should stand-out as a success. This influence would surely extend to the other Service Providers. UTV may thus assume a more prominent role in stressing other evaluation standards, e.g. data collection and analysis, and in advocating for a more detailed standardised approach to evaluation that can facilitate a comprehensive and strategic view of decentralised evaluations. This standardized approach can go beyond issue of Methodology and Terms of Reference and establish common standards for data collection methodologies and approaches, analytical tools, and how these are incorporated into reports.

In regard to advisory services on results frameworks (Advisory Services), this Review has found these to be relevant and effective service to Sida and Embassy staff and, most importantly, to the Partner Organisations that receive these services. **There is sufficient evidence to support the conclusion that Advisory Services have a direct and positive impact on Partner Organisation's work with results frameworks.** This is a significant achievement.

Having a single supplier provide these Advisory Services within a framework agreement is appropriate. The limited duration and discrete nature of these services, e.g. providing direct counsel on established standards about results frameworks, has a level of consistency and standardisation that is not readily possible in evaluation services, by comparison. In addition, the first ranked Service Provider has exhibited the requisite experience and talent for the facilitative, participatory approach that these services require. If not quite leading in evaluation methodologies and analytics, this Service Provider is leading here.

This implies that **a ranked system for Advisory Services is appropriate for future framework agreements.** While market implications remain, especially in regard to strengthening incumbency, this is less of an issue when compared to evaluation services. The quality of these services, the impact on partner organisations, and the relatively lesser risk as compared to evaluations that provide discrete conclusions and recommendations, implies that these benefits outweigh any market or other concerns. Additionally, there does not seem to be any direct value in including these services under the same framework agreement as evaluation services.

While these conclusions indicate a different type of framework agreement going forward, it should not diminish the dedication and professionalism that all three service providers have applied to their respective assignments. As the first ranked Service

Provider, Indevelop AB (Indevelop) demonstrates commitment, both intellectual and commercial, to working with Sida. As Indevelop's Annual Reports, the synthesis of 71 reports, and its on-going engagement with Sida demonstrate, it is an intellectually driven organisation whose work provides value to Sida and Swedish embassies who commission their services.

Finally, this Review concludes that Sida itself should be diligent in how it makes decisions regarding any future framework agreements and procurement processes. This Framework Agreement was ambitious and innovative, attempting to incorporate leading evaluation theories and procurement practices in one model. While the intellectual effort can be lauded, practical experience may have been neglected. It attempted to combine multiple, complex commissions under one framework and to increase the quality of these services overall. The potential to achieve all of these would be exceptionally difficult for any organisation.

Sida may be better positioned to make related decisions based on practical questions and considerations. Do we have the internal resources to ensure that service providers meet acceptable quality standards in all of their work? Is there a sufficient priority placed on managing the framework to ensure institutional support? Is Sida/UTV organised in a way to maximise these benefits associated with any framework agreement? This is not meant to diminish the value of being ambitious. It implies simply that an organisation should be diligent in considering what it knows and doesn't know about such ambitions and recognise that the greater the ambition, the greater the risk.

Given these and other findings and conclusions contained in this Report, we propose the Recommendations below. Please note that these are brief summaries of this Review's Recommendations. They are detailed in the relevant parts of the "Findings" section and elaborated in the section on "Recommendations" toward the end of the Report.

Current Framework (UTV)

1. Stress the need for evaluation questions that follow OECD DAC evaluation levels to Sida Commissioners. UTV should also ensure that Service Providers follow these questions, developing an evaluation matrix that includes all questions, their data sources, limitations and constraints, and risks and opportunities. This Evaluation Matrix should be mandatory for all Inception Phase Reports. (Section 4.2.1.1)
2. Provide Service Providers with expected standards and practices in relation to data collection methodologies and approaches, especially as they apply to field visits. Ensure that Service Providers improve data collection methodologies and approaches. (Section 4.2.1.1)
3. Standardise what information, statistics, or materials are important for ongoing institutional learning about development evaluations in collaboration with Indevelop. (Section 4.3.4.2)

4. Develop guidance about how much evaluations should cost as based on scope, requirements, and proposed methodologies. (Section 4.3.5.1)

Current Framework (Service Providers)

5. Ensure that all Core Team members and team leaders recognise the importance of utilisation, especially during the inception phase, and that they are sufficiently versed in best practices for engagement/utilisation throughout the evaluation process. (Section 4.2.2.3)
6. The increasing importance of RBM and results reporting implies that organisations must not only instil these as part of standard business practices but to use these as tools for increased performance. Toward that, both Indevelop and Sida would be served well by a post-facto survey among Partner Organisations that have received Advisory Services. This survey could be designed to gauge how, in precise and practical terms, programmes have changed due to their increased use of RBM practices. (Section 4.6.1.2.)

Current Framework (Indevelop)

7. Update Toolkit to reflect best practices and common standards in data collection, analysis, and results. (Section 4.2.3.1)
8. Develop a common QA procedure for data collection and analysis during the field/date collection phase, amongst all other areas of the evaluation, and demonstrate how this contributes to increased results. Include these results in Service Partner's Annual Reports. (Section 4.2.3.1)

Post-Framework Agreement (All Sida)

9. Sida should insist on a small core team (Max 10) whose primary expertise and role is to serve as the evaluation lead on any assignment. If this is maintained across multiple Service Providers, as would be the case in “mini-competitions”, then a diversity of talent and perspective could be maintained. While this does present a limitation, in view of the diversity of consultants who may bring different perspectives and innovative tools, this is superseded by the need to have a more consolidated and standardized approach to how evaluations are conducted. (Section 4.1.3.1)
10. A system of ‘mini-competitions’ between pre-selected suppliers is a more appropriate procurement process given the need and the complexity of development evaluations. (Section 4.3.3.2)
11. Future framework agreements may include a separate contract for advisory services on results frameworks and apply a ‘winner take all’ approach. (Section 4.7.1.1)

1 The Evaluated Intervention

This Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks (Review) serves as an opportunity to identify improvements that may be necessary and/or possible for the second half of the Framework Agreement. It may also provide information and suggestions about how Sida may develop any future framework agreements. The detailed purpose of the Review is included in the Terms of Reference included in the *Annexes*. Specific Evaluation Questions for the Review are also included in the *Annexes*. The Inception Phase report is included as a separate *Annex*.

In brief, Sida recognises the links between evaluations, results based management (RBM), and learning that can contribute to increased performance and results. Evaluations, in particular, are “careful, systematic and retrospective assessments of the design, implementation and results of development activities.”²

Sida conducts both ‘strategic evaluations’, those that have an agency-wide interest and are used for institutional strategy and development, and ‘decentralised evaluations’ that serve the on-going needs of commissioning units and Embassies. These support decision-making, programme development, funding, and other operational issues. This Review concerns ‘decentralised evaluations’.

Sida expects that this Framework Agreement supports increased efficiency and costs savings while also increasing the quality of services overall. This was in reaction to a need for costs savings and recognition that there was no centralised repository about how many and what types of evaluations were being conducted within Sida and Swedish embassies. Prior to the FA, evaluations were commissioned through different thematic framework agreements and through procurement of individual assignments, defined by sector, at different departments and units and with no centralised approach.

² Terms of Reference for this Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks.” Reference No. 13/000428. Appendix 5, page 1.

There was also evidence, as reported in “Are Sida Evaluation Good Enough?”, that quality did not adhere to adequate standards for evaluation.³ This report critiqued the evaluations under assessment for not properly adhering to the OECD DAC standards for development evaluations.⁴ It also criticised them for a range of other issues. (Please see Section 5 “Conclusions” for a list of these and how they compare to this Review’s results.)

Given this context, Sida developed a Framework Agreement (FA) that came into effect on 11 April 2011 and that remains valid until 10 April 2015. This FA includes two primary components:

1. The design and implementation of strategic and decentralised evaluations of contributions funded by Sida, including pre-studies/ concept papers as input to the design of evaluations, with a contract value not exceeding 1,200,000 SEK.
2. Advice on results frameworks for contributions funded by Sida, with a contract value not exceeding 100,000 SEK.⁵

The FA is used when Sida Units procure development evaluations and advisory services while Embassies may choose but are not required to use the FA.

One of the defining characteristics of the FA is its “ranked” system of call-offs. In brief, three service providers were selected to be part of the FA. According to the procurement process, they were ranked as follows:

- 1st Indevelop AB (Indevelop)
- 2nd The Consortium SIPU International (SIPU)
- 3rd The Consortium ORGUT Consulting AB (ORGUT)

This ranked system implies that Sida shall always seek services from the first ranked provider who may opt to provide a proposal for the services or pass to the second ranked provider, who may do the same.

³ Kim Forss, Evert Vedung, Stein Erik Kruse, Agnes Mwaiselage, and Anna Nilsdotter, “Are Sida Evaluations Good Enough? An Assessment of 34 Evaluation Reports.” Sida, 2008.

⁴ See <http://www.oecd.org/dac/evaluation/>. For a good summary and critique of the OECD DAC criteria, see: Thomas Chianca, “The OECD/DAC Criteria for International Development Evaluations: An Assessment and Ideas for Improvement.” *Journal of MultiDisciplinary Evaluation*, Vol. 5. No. 9, March 2008.

⁵ Terms of Reference for this Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks.” Reference No. 13/000428. Appendix 5, page 2.

Other unique features of the FA include:

- Service providers have a fixed group of evaluation and subject matter experts (Core Team) of whom at least one should participate in every evaluation.
- Service providers should have a dedicated management and quality assurance to ensure the highest possible standards and to improve quality over time.
- Evaluations should be conducted in ways that stress engagement/utilisation. This is meant to improve evaluations' quality and to contribute to greater utility of evaluation results. Sida recognises utilisation and engagement as integral to quality.
- Expectation that a closer relationship with a primary and then secondary and tertiary suppliers would support a common understanding of what Sida overall expects from decentralised evaluations and how this and the features mentioned above can support quality.

Finally, the results framework component is a new service intended to raise the quality of partner organisation's results frameworks and thereby also the quality of Sida evaluations and Sida-supported programmes. Results framework services are meant to increase capacity and understanding on results based management, reporting, and techniques to support how organisations develop indicators, monitor results, and report on overall performance. These services are of limited duration (maximum 10 days). This provides an ample introduction or 'primer' for organisations with an expectation that they may continue to develop their results frameworks. Given its novelty, this Review presents an opportunity to gauge these services' successes and shortcomings and how they may be adapted going forward.

The primary stakeholders and audience for this Review include Sida, and Sida's Monitoring and Evaluation Unit (UTV) and Procurement Unit (UPPIS) in particular, as well as the Service Providers that are party to the Framework Agreement. Sida Programme Officers and Embassy staff who procure services from the FA (Sida Commissioners) may be interested in the Review's results as well. Given the Framework Agreement's market implications, service providers in and beyond Sweden may also have an interest in this Review's results. Sida may decide how best to engage these audiences.

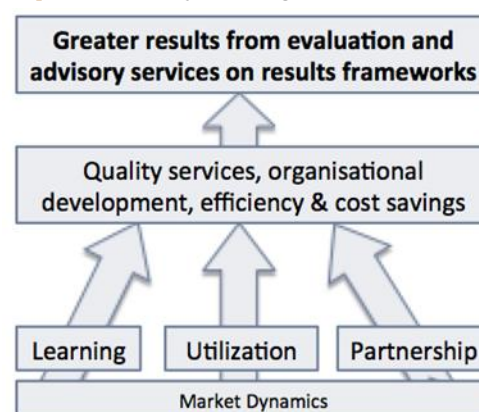
For more on the purpose, logic, and structure of this Review, please refer to the Terms of Reference and to this Review's Inception Phase Report included as a separate Annex.

2 Methodology

This Review’s methodology is based on an analysis of how the FA contributes to results. This includes a contribution analysis approach to the primary theory presented during the Inception Phase, mainly that the FA supports quality services, organisational development, and efficiency and costs savings (outputs) that will contribute to “greater results from evaluation and advisory services on results frameworks.” (Please see this Review’s Inception Phase Report for more details about the choice of this methodology and its limitations.)

While this Review’s Terms of Reference implies that a common pool of ranked consultancies may contribute to greater utilisation in evaluation and advisory services, we believe that the theory of change is broader than a focus on utilization. (See Graph 1 to the right.) It includes increased institutional learning, partnership based on common interests, utilization (precisely how services are implemented), and the market dynamics that underpin such a procurement model. These are all related to the quality of services and to the sustainability of this quality over time.

Graph 1: FA Theory of Change



Given this, the Review included the development of a detailed Analytical Framework that includes agreed upon questions for each area of interest, mainly the model, its implementation, and the quality of services so rendered. (See the Analytical Framework included as part of the Inception Phase and the Evaluation Questions included in the Annexes.) These questions provide the basis for analysis and each question (Evaluation Question) is answered in the “Findings” section of this Report.

With a detailed Analytical Framework in place, our approach was based on the collection of data from multiple sources (documentary, interviews, surveys, and comparative analysis) to establish evidentiary trends. For instance, if a majority of respondents express the same position and this is confirmed through surveys, supported by an analysis of the Evaluation reports, and is aligned with best practices and leading research, then one can make a fair conclusion. If all these sources “contribute” to the conclusion, it can be considered sound.

This includes ensuring that each data set is as complete as possible, e.g. that the sample size is appropriate, that there are no significant gaps, contradictions, or other peculiarities in specific data sets, and that any corresponding evidentiary trends contribute to answering specific Evaluation Questions. In this sense, the data collected during the Review proved valid. (See “Data & Analysis Report.”) The majority of evidentiary sources were consistent and conclusive. In the end, conclusions are based on this evidence and analysis and recommendations flow directly from these.

There were limited anomalies and inconsistencies in the qualitative evidence but this does not alter overall trends or conclusions. Experiences within and among cohorts will vary and this will be reflected in corresponding qualitative data sets. Yet, to ensure that individual respondents do not unduly influence the analysis, our findings are based on significant trends. This does not imply that we do not consider issues and problems that individual cohorts raise. These are considered in view of the Review Team’s expertise, leading research, and other references that can substantiate such issues.

The methodology and this Report do not follow strict OECD DAC evaluation levels (relevance, effectiveness, impact, efficiency and sustainability) as the structure (operationalisation, model & quality) and the breadth of evaluation questions are aligned with Sida’s needs for this Review.

2.1 SOURCES

In addition to best practices, leading research and other comparative data, the following sources contributed to findings and conclusions.⁶ These have been organised and analysed in a separate “Data & Analysis Report” that is referenced throughout this Report.

2.1.1 Assessment of Evaluation Reports

The Review included an **assessment of 80 evaluation reports** out of the 98 completed evaluations. This does not include 11 reports that were not appropriate given the assessment approach, e.g. desk studies, and others that were in languages beyond the Review Team’s competence. The complete list is included in the Annexes.

⁶ By best practices and leading research, we refer to OECD DAC, UNEG, and the World Bank IEG along with other research. These are cited as appropriate throughout the findings section.

The assessment was done according to standardized metrics that are based on those used in the “Are Sida Evaluations Good Enough?” (ASEGE) evaluation of 2008.⁷ This includes the ranking of 50 different metrics with supporting notes and references to the reports under review. This enables the Review to make reference to the conclusions from that report and to gauge, with distinct limitations, any changes from that time. All of these assessments are included in this Review’s “Data & Analysis Report.”

Calibrating ASEGE Metrics

This Review calibrated the use of the ASEGE metrics to ensure that there could be suitable comparisons. For instance, the ASEGE approach includes a ranking of ‘1’ any time an element was missing while this review assesses the importance of the element and its mention in the Inception Phase report, thus resulting in most instances classified as “ND” or “not done”. To rectify this, we removed relevant rankings of “1” so that there is a more accurate comparison. We have also considered the standard deviation from the mean between the ASEGE rankings and those conducted as part of the Review. Anything within that standard deviation, as indicated whenever cited, is not considered significant. Finally, we discussed the metrics with one of the report’s authors, Mr. Kim Forss, to ensure that we were interpreting these metrics similarly.

This adjustment in rankings, using a standard deviation calculation between the Review’s and ASEGE’s, and ensuring a similar interpretation of the metrics, enables a fair comparative data set. This is especially so given the number of reports under review in both ASEGE (34) and this Review (80). This is a large enough sample to establish conclusive findings, e.g. the average ranking for any quality criterion.

Table 1: Interviews per Cohort	Totals
Sida (UTV; UPPIS)	10
Sida Commissioners (Advisory Services)	15
Sida Commissioners (Evaluations)	22
Sida Commissioners (Evaluations & Advisory Services)	13
Partner Organisations (Advisory Services)	24
Partner Organisations (Evaluations)	21
Partner Organisations (Evaluations & Advisory Services)	2
Service Providers (Management)	8
Service Providers (Coalition Partners)	6
Service Providers (Core Team)	17
Service Providers (Consultants)	4
Non Framework Agreement Consultants	2
TOTAL:	144

⁷ “Are Sida Evaluations Good Enough? An Assessment of 34 Evaluation Reports,” Kim Forss, Evert Vedung, Stein Erik Kruse, Agnes Mwaiselage, and Anna Nilsson. Sida, 2008.

Reasonable people could disagree with some of the rankings and conclusions. However, the purpose of the assessment is not to firmly establish the quality of individual reports but to ascertain whether there are reasonable trends in different areas across all reports assessed. Toward that, we are confident that such trends are evident.

2.1.2 Qualitative Data

This includes **144 interviews** that used standard protocols, as included in the Annexes. These interviews constitute the source for most qualitative data. This includes, in total, 2099 qualitative data points. All of this data and initial analysis is included in the “Data & Analysis Report.”

Additional qualitative data was collected during the field phase. This includes interviews with consultants in Sweden not directly associated with the FA and a few others. Qualitative data from these were not included as they served mainly as a control against other data sets.

Interviews were spread across cohorts as depicted in Table 1. This met the targets established by the Review Team and agreed upon in the Inception Phase for each cohort established during the Inception Phase, exceeding those for Partner Organisations. A full list of people interviewed is included in the Annexes.

2.1.3 Quantitative Survey

A survey was launched to different cohorts at the end of the field phase. The primary purpose of the survey was to support other evidence and to provide an opportunity to triangulate key data, especially from qualitative data sets. The response rates per cohort are as follows:

Table 2: Response Rates to Quantitative Survey			
Cohort	Target	Completed	%
Sida – Operations (UTV; UPPIS)	12	3	25%
Sida Commissioners (Evaluations)	83	28	34%
Sida Commissioners (Advisory Services)	50	16	32%
Partner Organisations (Evaluations)	24	12	50%
Partner Organisations (Advisory Services)	25	15	60%
Service Providers (Management & Coalition Partners)	15	8	53%
Service Providers (Core Team)	28	17	61%

The survey included only partner organisations that were interviewed. Others would not have enough information about this Review to make an appropriate contribution.

2.1.4 Best practices & Leading Research

Evidence is always assessed according to best practices and leading research. We draw on these to support claims and to buttress or contradict other sources. For instance, respondents may have a relatively uniform opinion about a subject that is in contradiction to established theories or best practices. In these cases, we will point out this contradiction and then draw appropriate conclusions.

In other cases, there may be no direct evidence from different cohorts related to specific evaluation questions. In these cases, we draw on established theories or leading articles. Some theories or practices can be widely debated or contentious. Our view in these is established through the citations we choose. Reasonable people could draw from different research.

2.1.5 Review Team's Expertise

Finally, our own expertise and perspectives are brought to bear on the subject. This includes how we interpret different data and the way we craft findings and conclusions. Yet, we try to minimise this and to instead draw on other sources and solid evidentiary trends.

Our independence is not only forged from our separation from the subject but also by the keen awareness we have of human bias. This is why we are so diligent in establishing solid evidentiary trends and comparative evidence. We expect to be surprised by the data—it should tell a story that we could not see or did not expect. Gratefully, this was often the case in this Review. We were surprised by the positive changes in the ToR and in the relatively high levels of engagement and utilisation, amongst other things.

More importantly, the major findings and conclusions are discussed given our unique expertise and perspectives. We actively challenge one another and bring in different examples and evidence that may support uncertain conclusions. This should lend itself to the credibility and strength of our conclusions and enable stakeholders to successfully move forward with this Review's results.

For other aspects of the Methodology, including relevant limitations and constraints, please see this Review's Inception Report included as a separate Annex.

2.2 USING QUALITATIVE DATA GRAPHS

This Report uses qualitative data graphs throughout. Qualitative analysis graphs demonstrate trends by categorising responses according to a set scale and organising them according to specific Evaluation Questions and cohorts. Qualitative data is inherently difficult to analyse but can provide strong evidence for trends in current perceptions.⁸

The Qualitative Data graphs used throughout this Report are based on the findings of the Qualitative Data and Quantitative Survey described above. All collected data and related analysis is included in the “Data & Analysis Report” that is included as a separate Annex.

Each Qualitative Analysis Graph includes colour/grey scale coded and numbered boxes for each piece of qualitative data. Thus, if there were 35 responses, as in the example below, there will be 35 corresponding boxes, with box “1” corresponding to statement “1,” box “2” to statement “2,” etc. It is expected that each question has a different number of responses—qualitative data analysis generates different numbers of responses for each subject.

Each respondent statement (data point) has been evaluated according to the following scale:

☐ - Positive; achieved expected results ☐ - Neutral; Mixed results ☐ - Negative; did not achieve expected results

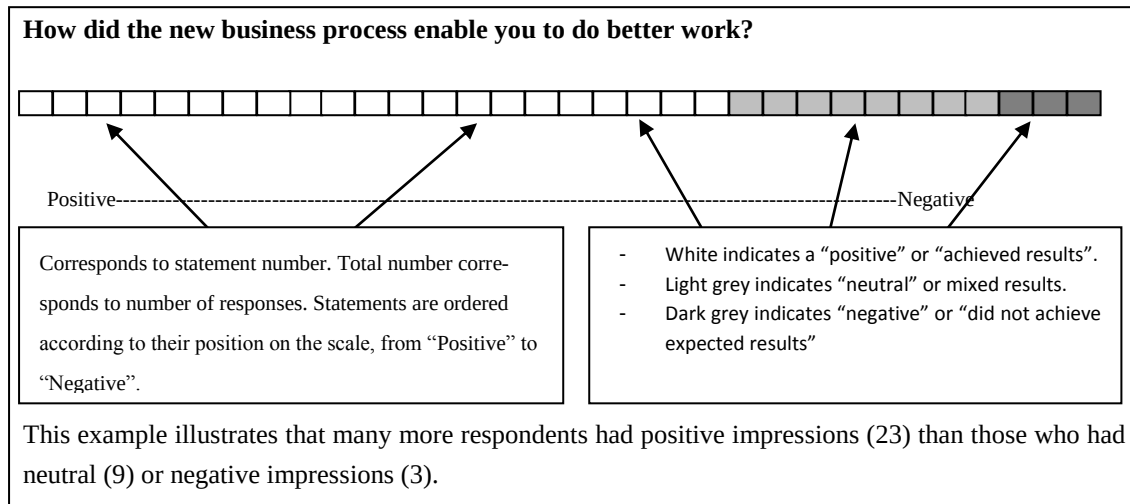
These rankings are subjective. They are based on the statement as confirmed by respondents and on their overall intent. Others may reasonably arrive at different conclusions. Each qualitative data point includes identifiers as to stakeholder. Some qualitative data points are repeated as different cohorts often say precisely the same things and it is best to establish this linguistic trend. Every attempt has been made to preserve respondent’s anonymity although confidentiality cannot be assured.

This Review presents high-level trends across qualitative data sets. While deeper analysis, including the precise number of response types, comment verbiage, and other qualitative issues is possible, it is not within the resources and/or remit of this Review to conduct this level of analysis.

⁸ For a review of evidence based evaluations and the use of qualitative data, see “Evidence-based Evaluation of Development Cooperation: Possible? Feasible? Desirable?,” Kim Forss and Sara Bandstein. *Network of Networks on Impact Evaluation (NONIE)*, World Bank, January 2008.

All data is available in the “Data & Analysis Report” included as a separate *Annex*.

EXAMPLE:



3 Findings

This section is organised according to the three primary areas of concern for both evaluations and advisory services (operationalisation, model, and quality). This includes all Evaluation Questions agreed upon as part of the Inception Phase. (See “Evaluation Questions” in the Annexes.)

These findings are based on evidence collected during the field phase and on best practices, as determined through Sida’s Evaluation Manual and related documents, OECD DAC standards, the World Bank’s Independent Evaluation Group (IEG) standards, the United Nation’s Evaluation Group (UNEG) standards, leading research, and the Review Team’s experience in evaluations, procurement, and related fields. At the same time, there are no formally accepted standards to dictate all aspects of evaluations and related services either for Sida or internationally. Thus, some may reasonably disagree. Every attempt is made to draw on this evidence and indicate when it is not wholly conclusive. (See “Section 3: Methodology” above.)

While primary conclusions are included in the relevant section toward the end of the Report, some conclusions, questions, or prescriptive remarks are included from time to time. These are meant to provide greater clarity about the findings and/or to draw attention to issues that may arise from the findings but that are not definitive enough to include as final conclusions.

Recommendations are included as relevant to different findings and repeated in the “Recommendations” section toward the end of this Report.

3.1 OPERATIONALISATION (EVALUATIONS)

3.1.1 Number & Type of Evaluations

3.1.1.1 *How many evaluations have been conducted within the Framework Agreement (total and per consultant) and how many have they turned down? (A1.1.1)*

The total number of evaluations, reviews, assessments and other related activities (evaluations), completed and on-going, is as follows:

Table 3: Number of Evaluations Assigned under Framework Agreement						
	Evaluations	Reviews	MT Reviews	Assessments	Other	TOTALS
Indevelop	45	35	12	2	9	103
SIPU	11	4	3		1	19
Orgut	2	2				4
TOTALS	58	41	15	2	10	126

This demonstrates that 82% of the assignments have been assigned to Indevelop.

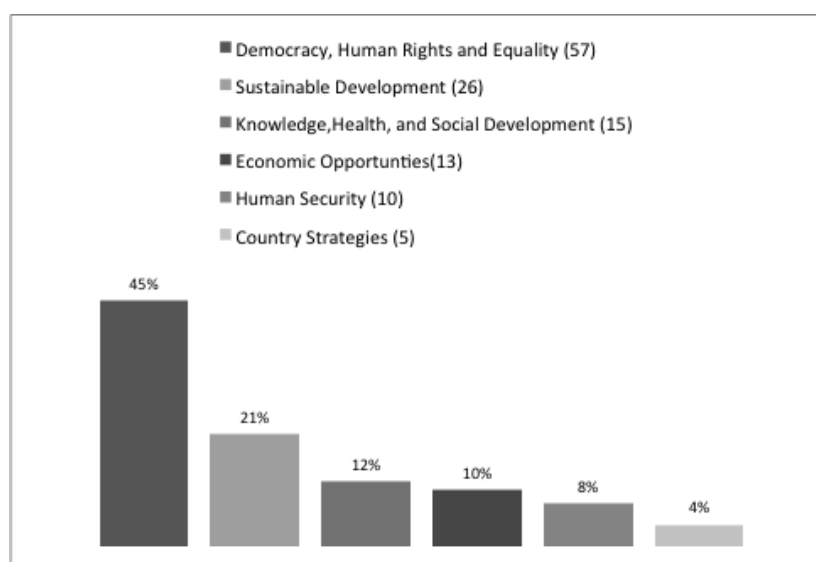
While the FA accounts for a wide range of evaluation activities, including summative, formative, theory-driven, impact assessments, lessons learning, rapid assessments, real time, meta- and synthesis evaluations, most of the evaluations assessed fall into the broad categories noted above. In fact, there is little differentiation in form between reports labelled ‘evaluations’ or ‘reviews’: there are varying approaches and analytical tools used in both and toward widely different evaluative results.

3.1.1.2 What kinds of evaluations have been provided? (A1.1.2)

The type of evaluations has been spread along the project cycle (mid-term, periodic, and final) with over half occurring at a programme’s end.⁹

These have been spread across thematic areas with the majority in “Democracy, Human Rights and Equality. (See Graph 2 below.)

GRAPH 2: Evaluations by Thematic Area (126 in total)

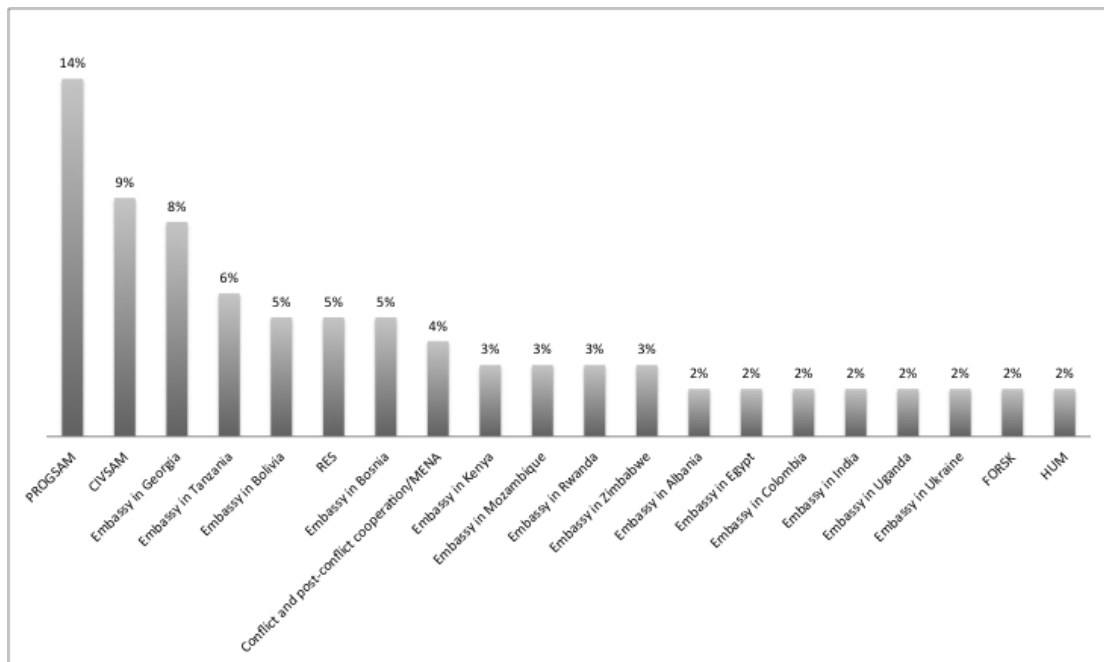


⁹ This is based on a review of evaluations from Orgut and SIPU and Indevelop’s most recent annual report, “Final Annual Report April 2012 – April 2013 (year 2),” 5 June 2013; p. 7.

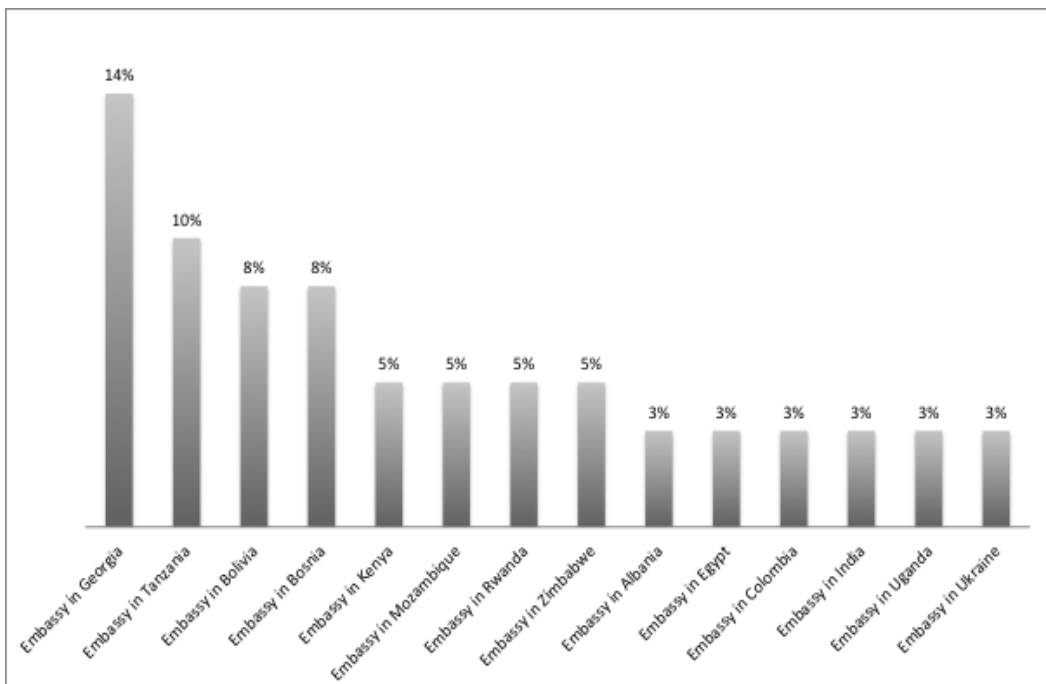
3.1.1.3 To what extent has Framework Agreement been used for commissioning evaluations by Sida Units/Embassies? (A1.1.3)

Approximately 40% of all evaluation assignments have been commissioned by Sida Units with the remaining 60% at embassies. The graphs below show the distribution amongst those that have commissioned more than one service. (For actual numbers, please refer to “Evaluations Conducted under Framework Agreement” in the *Annexes*.

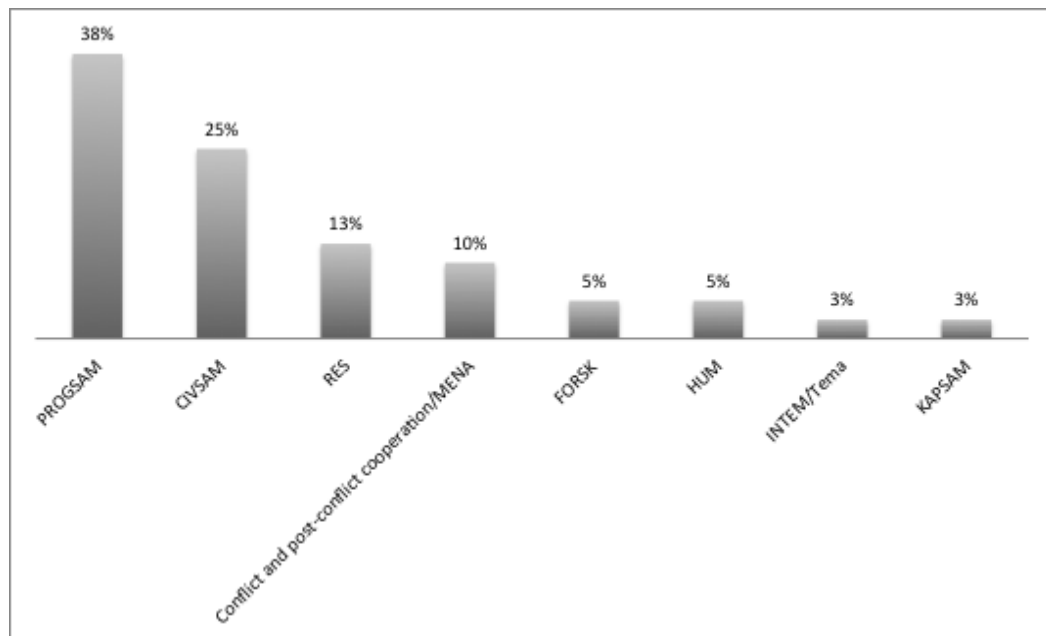
GRAPH 3: Evaluation Assignments across Sida and Embassies



GRAPH 4: Evaluation Assignments across Embassies



GRAPH 5: Evaluation Assignments across Sida Units



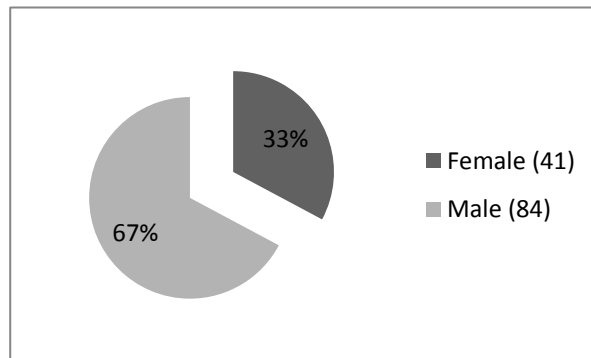
While there are some omissions and a concentration of activity in a few areas, this does not seem abnormal. Sida may wish to ensure that all Embassies are informed of the FA and its purpose.

3.1.2 Evaluation Team Composition

There has been a relative spread in demographics and competencies among consultancy teams. This is based primarily on Indevelop's reporting on this as well as a review of the other Service Provider's teams for individual assignments.¹⁰

There could be a better gender balance among Team Leaders (TL), as Graph 6 (to the right) illustrates. This shows that men make up 67% of TL positions.

GRAPH 6: Team Leader Composition by Gender (All past and current assignments, 121)



While this review cannot comment upon the different competencies and experience that go into each team, there has been a concerted effort to ensure that all teams have one person with evaluation expertise. This is as intended in the FA and evidence suggests that this has been the case in all evaluations, even when conducted by one person who has a mix of skills including evaluation experience. This is supported by qualitative evidence from Sida commissioners and Partner Organisations. (See Section 4.2.1.) While a proportion do raise the competencies of teams as an issue, these responses are a minority overall.

3.1.2.1 *What have teams looked like in terms of gender, Swedish/national/international team members? (A1.1.6)*

The primary source for this is Indevelop's most recent Annual Report. It shows that 67% of individual evaluators have been from outside Sweden and that 55% of teams have been male.¹¹ The gender composition of these teams shows an improvement from 2007 when Sida reported that evaluations conducted that year were composed of 40% women (as compared to 45% now).¹² This same report highlights that 54% of the evaluations included team members from partner countries or regions.

¹⁰ Please see: "Final Annual Report 2012 – 2013," Indevelop. June 2013.

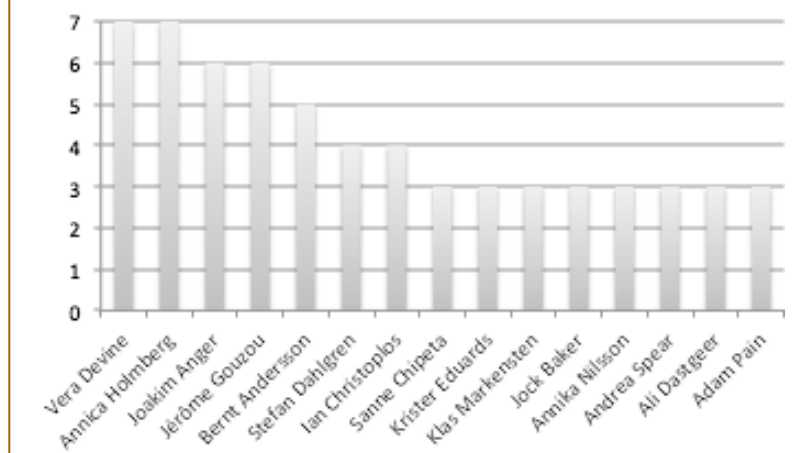
¹¹ IBID. Pages 24 – 25.

¹² Lennart Peck & Verena Knippel, "Sida Evaluations and Audits 2007: A Synthesis of Findings and Conclusions." UTV Working Paper 2008:1.

This Review determines that while the overall gender of evaluation teams is balanced the role of Team Leader (TL) is not. In reviewing the TLs for all assignments, only 33% are women. (See Section 4.1.2.)

We also reviewed how often different TLs have been used. For Indevelop, there is a considerable spread amongst

GRAPH 7: Team Leader Assignments at Indevelop (Those TLs who have done more than 2 assignments)



Team Leaders with no TL doing more than 7% of all evaluation assignments. (Graph 7.) Again, there is not enough evidence to draw suitable conclusions from The Consortium SIPU International (SIPU) or The Consortium ORGUT Consulting AB (Orgut).

3.1.2.2 *Are consultant competencies (experience, expertise, aptitude, attitude) appropriate and adequate for specific evaluations? (A1.1.7)*

Qualitative data suggests that, overall, consultants' competence was acceptable to Sida Commissioners and Partner Organisations. In responding to questions of quality, these cohorts seldom mentioned issues of competence or other inadequacies with the team. Or, more precisely, only a minority of respondents mentioned issues with the team: that they had the wrong competencies, that there were potential conflicts of interest, or that there was not sufficient depth of knowledge about an organisation or subject.

Out of 185 qualitative data points from Sida Commissioners and Partner Organisations regarding overall quality, there were 25 'negative' comments and 7 'neutral' comments about team competencies, or approximately 17% of the total. (See "Data & Analysis Report", pages 258 – 265.) This does not necessarily imply that 17% of assignments had teams with the wrong competencies. It simply implies that less than 20% of respondents saw this as an issue.

While there are always opportunities to improve, this does not seem to be an unreasonable finding given the complexity of evaluations and the subjects they treat, the varied expectations of Sida Commissioners and Partner Organisations, and the notion that one who has evaluation expertise is as important as one who has relevant subject matter expertise (both are important). Upon a general review, over a third of these negative remarks relate to this last misconception.

The use of a team member, and most often a Core Team member, who has evaluation expertise is prevalent among all evaluations under review. This is as intended in the FA: that evaluation expertise assumes prominence in every assignment. As shown elsewhere, while these competencies are prevalent, there is little evidence that this has lent to acceptable levels of quality in evaluations, especially in regard to data collection and analysis. (See Section 4.2.1.1.)

At the same time, Core Team members, especially within Indevelop, have improved their evaluation competence both on assignment and through on-going support. Indevelop has been particularly active in sharing best practices and holding semi-annual retreats where different methodologies are presented and discussed. (See Section 4.2.3.2.) For instance, this Review noticed an increase in the use of outcome mapping after this was presented in one of Indevelop's retreats. This indicates that the Core Team is improving their evaluation skills and that this may result in improved evaluation competence.

3.1.3 Implementation

3.1.3.1 *What are the primary attributes of the collaboration between Sida and the Service Providers and how has this impacted the way the FA was implemented and how quality services are assured? (A1.1.4)*

As the following graph shows, responses from both from Sida and SPs about collaboration are mixed. This graph shows qualitative data from different cohorts and whether individual comments are positive (no shade), neutral (light shade), or negative (dark shade). The total number for each establishes the trend.¹³

GRAPH 8: What are the primary attributes of the collaboration between Sida and the Service Partners and how has this impacted the way the FA was implemented and how quality services are assured? (A1.1.4; Combined cohorts: Sida operations, SP management, & coalition partners)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61																		

¹³ For reading this and similar graphs, please refer to "Using Qualitative Graphs" in "Methodology". Please see also the "Data & Analysis Report" included as a separate Annex for all qualitative data.

Primary issues raised by respondents include:

- The number and competencies of the core team;
- Different roles/approaches by UPPIS and UTV, especially UPPIS tightening of what is allowed under the contractual terms;
- Consultant competencies, e.g. the requirement of a Masters degree or higher.

Positively, all respondents state that the on-going relationship promotes a level of accountability that may not be possible in “one-off” contracts and an open, frank and productive communication with UTV throughout. (For all related qualitative evidence, please see “Data & Analysis Report,” pages 246 – 248.)

As noted, issues have been raised by UTV and SP Management cohorts regarding the changing role of UPPIS. This includes UPPIS’s insistence that there is not an opportunity to discuss the ToR, its implications and feasibility, until after a proposal is presented. Given discussions with UPPIS, this is based on the legal terms of the agreement and cannot be changed. Some state that this limits the potential value of the FA by limiting discussions with the first ranked about approaches and teams, amongst other issues. In fact, this Review indicates that these discussions have simply shifted to lengthier and more detailed Inception Phases. (See Section 4.2.2.2.) In this way, this contractual issue has made the FA work like most procurement processes, e.g. suppliers present their best bid in the hope to win and with the expectation that the details will be settled in the Inception Phase.

Of the issues noted above, the number and competencies for the Core Team has the most direct impact on how the FA was implemented and how quality services are assured. While not specifically stated in the original Terms of Reference, Sida expressed that Core Teams were meant to be small (around 5 members). This would ensure a dedicated cadre who would be involved in most if not all assignments and thus quickly develop expertise and efficient ways of working that would lend to higher quality. The Service Providers argued that it would not be possible to cover all thematic areas with only 5 core team members and, in their proposals and subsequent conversations this was expanded to 20.

While difficult to substantiate the causal links, the mixed quality across evaluation criteria in reports and the relatively un-even progress in improvement over the first 2 ½ years may be related to this compromise. The main areas of concern related to quality, mainly methodologies, data collection and analysis, and evidence based conclusions, includes the inconsistencies in these across the evaluations under review. Sometimes these and other evaluation approaches are done well. Approximately half the time, they are not. (See Sections 4.2.1 & 4.2.2.)

If there were a smaller cadre of evaluation experts involved in every assignment, not only might quality, in terms of methodologies, data collection and analysis, be higher but there would be a greater consistency in relevant approaches as well. The increased attention to methodologies, data and analysis, and related evaluation methodologies would, as best practices and leading research dictate, ensure even greater results. It

would also give UTV the opportunity to better manage the evaluations, e.g. maintain quality overall among Service Providers and in analysis about all commissioned evaluations.

Recommendation 1 (UTV): If a similar Framework Agreement is used in the future, Sida should insist on a small core team (max 10) whose primary expertise and role is to serve as the evaluation lead on any assignment. If this is maintained across multiple Service Providers, as would be the case in ‘mini-competitions’, then diversity of talent and perspective could be maintained. This will lend to higher levels of quality overall by having a smaller group accountable to the highest possible evaluation standards. At the same time, if consultants adhere to common standards and best practices in evaluation and rigorous QA is done for data collection and analysis, as recommended, the result should be higher quality of evaluation reports no matter the consultant. While this does present a limitation, in view of the diversity of consultants who may bring different perspectives and innovative tools, this is superseded by the need to have a more consolidated and standardized approach to how evaluations are conducted.

3.1.3.2 How has Sida organized itself to manage and utilize the Framework Agreement, including legal and administrative matters, support and communication? Has it been adequate and appropriate? (A1.1.8)

While an in-depth review of Sida structures and organisation is not part of this Review, it does seem appropriate that there is a division between the contractual and legal issues related to the FA (UPPIS) and the technical, strategic and partnership issues (UTV).¹⁴ This conforms to best practices and common standards.¹⁵ While the nature of this arrangement may fluctuate in precise terms, UPPIS does serve to ensure that the terms meet legal standards and that they are followed. This is obviously critical for any organisation let alone one trusted with public funds. Thus, while respondents have raised issues about UPPIS role and how it has interpreted the FA and its contractual terms at different times, this Review has found nothing that is out of the ordinary or that does not conform to common standards for such procurement arrangements.

¹⁴ For a lively argument regarding governance and quality, see Anders Hanberger, Framework for Exploring the Interplay of Governance and Evaluation,” Scandinavian Journal of Public Administration. Vol. 16, No. 3, 2012.

¹⁵ Price Waterhouse Coopers (PWC), “Public Procurement in Europe: Cost and Effectiveness.” European Commission, March 2011.

As noted in Section 4.1.3.1, several key respondents state that they expected there to be greater flexibility regarding the negotiation and transparency between all ranked Service Partners when a request for proposal is issued. For instance, perhaps all three could present a proposal. This would have enabled a more diversified choice for Sida Commissioners and thus greater competition overall. In any case, UPPIS deemed this legally impossible given contract terms. They also prevented any negotiation or discussion with the first ranked supplier prior to a written proposal. This is meant to ensure that the second and third ranked have an equal opportunity to present valid proposals. A discussion between the first ranked supplier and the Sida Commissioner could constitute an unfair advantage.

The potential and real tensions between UPPIS and other stakeholders is, perhaps unfortunately, fairly typical. There are on-going technical needs, beyond actual contractual terms, that can best be negotiated and managed by personnel who understand both commercial interests and technical requirements.¹⁶ This seems to assume that, in Sida's case, every UTV officer would have the same level of legal expertise as evaluation expertise, and vice versa with UPPIS. This implies an impractical concentration of technical and contractual skills in one person or unit.

This is compounded by the change of personnel. As noted in this review, UTV and UPPIS have seen a considerable turnover in its staff during the FA term. This implies not only may there not be people who are equipped for both the contractual and technical nature of the FA but who also lack sufficient institutional memory and knowledge of the original rationales for decisions that affected how the FA was implemented. Given the normalcy of these organisational dynamics, this Review has not found any other significant legal or administrative issues that have been or could be detrimental to the FA.

¹⁶ For a convincing case for adapting private sector commercial skills in managing strategic procurement initiatives, see: Filip Roodhooft & Alexandra Van den Abbeele, "Public Procurement of Consulting Services: Evidence and Comparison with Private Companies." *International Journal of Public Sector Management*, Vol. 19, 2006.

3.1.4 Guidance & Strategy

3.1.4.1 How has Sida organised itself to provide guidance and support for using the Framework Agreement? (A1.1.10)

GRAPH 9: How has Sida organised itself to provide guidance and support for using the Framework Agreement? (A1.1.10; combined cohorts: Sida operations, Sida commissioners, SP management, and SP core team)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
100	101	102	103	104	105														

UTV is organised to provide direct and ad-hoc counsel to Sida Commissioners as warranted and to play the role of primary interlocutor with FA Service Providers. Qualitative evidence is mixed on this point:¹⁷

While cohort responses cover a range of issues, some of which are not directly related to how Sida organises itself, some trends emerge:

- Sida Commissioners appreciate the opportunity to go to UTV when issues arise and have sought out support for ToRs and other matters;
- Sida Commissioners have used UTV templates and other guidelines made available on-line;
- Some Sida Commissioners state that there were no opportunities or time to seek out support from UTV;
- Some found UTV guidance on ToRs include issues that go beyond their specific needs;
- Service Partners management commented on some of the inconsistencies between UTV and UPPIS;
- Service Partner core team members commented on the importance of the Sida commissioning officer and how they could be either very supportive or have

¹⁷ For reading this and similar graphs, please refer to “Using Qualitative Graphs” in “Methodology.”

expectations that are not aligned with the evaluator's expertise and perspectives.¹⁸

There is not enough qualitative or corroborating evidence to confirm the spread or validity of these positive or negative issues. Instead, we note that there are indications of considerable support and guidance despite the limited human resources within UTV. 7 – 8 staff would seem hardly sufficient to manage the quality of over 100 evaluations in 2 and ½ years and to provide active support and guidance for all of related Sida Commissioners. In fact, there have only been 1 -2 people dedicated to evaluations and their time has been mostly committed to strategic evaluation and process issues. The fact that there are as many “positive” responses as there are speaks to the way UTV staff is able to provide guidance despite resource constraints.

It should be noted that the FA is primarily a tool and, mostly, a procurement tool. It is limited in how it can be used, *prima facie*, as a tool to increase quality evaluations. Sida's ability to organise itself to support quality evaluations must therefore be seen as one of available resources and priorities. As this Review states, what can be done has to do more with ensuring that any service provider use appropriate methodologies for data collection and analysis, that these methodologies are aligned with an evaluation's scope and questions, and that other issue related to evaluation quality are adhered to. If this is done, Sida will have greater confidence that evaluations are achieving common quality standards for evaluations and that subsequent findings, conclusions and recommendations are the best possible for Partner Organisations. It will also ensure that, as resources and priorities change, Sida is positioned to draw from the implicit and explicit knowledge that exists in these evaluations and thus use this for its own institutional learning, knowledge management, and other strategic purposes. This is UTV's role and, as noted, it has done so while there is always room for improvement.

3.1.4.2 How has Sida organised itself to develop strategic analysis and trends as based on how the Framework Agreement is used? (A1.1.11)

Not enough evidence exists to suggest that Sida has organised itself to develop strategic analysis and trends on how the FA is used. This has been largely left to Indevlop who provides this level of analysis in Annual Reports and who has provided a synthesis of major issues in a recent report.¹⁹ This lack of organisation for strategic analysis may be due to limited and changing human resources within UTV. (See Section 4.1.4.1.)

¹⁸ See “Data & Analysis Report,” pages 243 – 246.

¹⁹ Ian Christoplos, Anna Liljelund Hedqvist, and Jessica Rothman, “Swedish Development Cooperation in Transition: Lessons and Reflections from 71 Sida Decentralized Evaluations (April 2011 – April 2013).”

This may constitute a missed opportunity as the centralised aspect of the FA enables Sida, and UTV in particular, to have greater insight into the types of evaluations being conducted, by whom, and at what stage of any programme's cycle.

At the same time, when UTV has taken an active stance in relation to FA issues, e.g. toward improving the methodology sections in reports, the stress on utilisation, and even in the ad-hoc communication with Sida Units and Swedish Embassies, there are tangible results. (See Sections 4.2.1 and 4.2.2.2, among others.)

This implies that UTV may choose, given resources and needs, specific issues on which to focus. For instance, UTV may decide to focus on data collection and analysis as a critical area of improvement and thus stress this with the Service Providers. It has a proven track record in its ability to steer Service Providers and to keep them accountable.

3.1.4.3 *Has the Framework Agreement been communicated in ways across Sida to support a common understanding of its purpose, services, and functionality?*
(A1.1.13)

While an assessment of communicative patterns and channels in any organisation's context requires a level of analysis not possible in this Review, we did seek indications of how the FA was communicated and if there are signs that this has led to greater utilisation. While all relevant cohorts were asked this question, very few had substantive responses.

GRAPH 10: Has the Framework Agreement been communicated in ways across Sida to support a common understanding of its purpose, services, and functionality? (A1.1.13; combined cohorts: Sida operations and Sida commissioners)

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

Qualitative evidence on this is mixed as most of UTV and UPPIS staff related to the FA only cites informal communication mechanisms. (See Graph 10 and the "Data & Analysis Report," page 247.)

In general, this graph includes comments that highlight adequate or inadequate levels of communication. The qualitative evidence lacks specificity, in most cases, about what worked and what didn't. As noted here and in other data, Sida Commissioners who have used the FA state that they were aware of it from meetings with UTV staff and notices on the Intranet and that they welcomed the FA as an efficient procurement method.

There was a concerted effort to communicate the purpose and processes of the FA at its launch. All evaluation managers had responsibility for communicating on a regular basis with around 6-8 commissioning units each. Meetings were called with all units

to discuss the FA. In addition, each unit had a focal point for evaluation that was also briefed about the FA. There were two promotional events at Sida headquarters at the time of the launch during which all embassies had staff in Stockholm (in the country teams). There was additional communication from UTV staff, both formal and informal. Other mechanisms for communicating the FA have been through Sida's intra-net and information fliers and in the regular meetings between UTV staff and Sida Units and Swedish embassies. As based on the number of evaluations conducted across units and embassies, many do know about the FA.

At the same time, one wonders about Swedish embassies who have not used the FA. Is this due to the lack of need, the choice not to use the FA, or simply because they are not aware of it? Unfortunately, this Review could not answer this. However, qualitative evidence does suggest that this is an "easy" message to receive in that it enables commissioners to procure services quickly and with a pre-qualified group of providers. More effort may be made to ensure that all embassies are aware of the FA.

Another aspect worth considering is that Sida has launched a number of framework agreements at the same time as the FA. No evidence was discovered of a communication plan or strategy that could accompany the launch of multiple framework agreements, their purpose, utility or the rationale for this broad change in procurement procedures. While beyond this Review's remit, this may be considered as a missed opportunity. A concerted communication plan/strategy may have not only increased utility but also stressed the importance of evaluation and the other related framework subjects.

Another avenue of communication concerns guidance on how best to work with and use the FA. While UTV has been available to Sida Commissioners at any stage of the evaluation process, there is no evidence that there has been, since the FA's launch, an active and coordinated outreach to these commissioners other than at preparatory briefings for staff to be deployed to Embassies, as Sida commissioners shared in interviews. This may be due to the UTV's reduced advisory role and/or to staff turnover. This may be beyond UTV's priorities and resources but it may have implications.

While this finding may warrant a recommendation for a more elaborate communication strategy, we have not offered one as it may not be an appropriate priority for Sida and the details of such a communication plan or strategy fall beyond the remit of this Review. Communication issues related to this and other frameworks have been raised in meetings and other informal channels as part of this Review.

3.2 QUALITY (EVALUATIONS)

3.2.1 Overall

3.2.1.1 *What is the overall quality of evaluation services under the Framework Agreement? (A2.1)*

The outcome of quality evaluations may be best summarised by the ability to get relevant stakeholders to think and possibly act differently in relation to the subjects being treated.²⁰ Quality, in this sense, is as much about impact as anything else: impact on the Sida officers commissioning the service, impact on the Partner Organisation, as well as on the consultants themselves who may broaden and deepen their evaluation expertise. Quality resides in all of these areas and others.

Evaluations should stand as more than ‘reality tests’ of aid effort and strategies, as stated in the “Are Sida Evaluations Good Enough?” report.²¹ Instead, they provide a basis for improved performance, greater relevance, effectiveness, impact, efficiency and sustainability to use the OECD DAC level criteria.²² Evaluation’s importance is growing and will play a role in everything from demonstrating results and assessing institutional performance to the development of international cooperation and development strategies.²³ Sida basically says this itself in this review’s ToR when it says: “Evaluation is an important prerequisite for results based management and learning.”²⁴

Unfortunately, there is not much of a common, or commonly understood, standard for what constitutes quality within or beyond Sida. For example, Sida’s Evaluation Man-

²⁰ For a fair treatise on issues related to quality in development evaluations, see Anneke von Raggamby, Frieder Rubik, “Sustainable Development, Evaluation And Policy-Making: Theory, Practise and Quality Assurance.” Institute for Ecological Economy Research (IÖW), Germany, 2012.

²¹ Kim Forss, Evert Vedung, Stein Erik Kruse, Agnes Mwaiselage, and Anna Nilsson, “Are Sida Evaluations Good Enough? An Assessment of 34 Evaluation Reports.” Sida, 2008. P. 5.

²² OECD DAC has greatly developed the standards for development evaluations beyond mere independence and impartiality and these standard evaluation levels. It clearly recognizes the role of evaluations to provide “credible independent evidence about the relevance, effectiveness, efficiency, impact and sustainability of development activities.” P. 17, “Evaluating Development Activities: 12 Lessons from the OECD DAC”, OECD, 2013.

²³ These issues and others are treated in Ray C. Rist, Marie-Helene Boily, & Federic Martin, Influencing Change: Building Evaluation Capacity to Strengthen Governance. The International Bank for Reconstruction and Development (IBRD)/The World Bank, 2011.

²⁴ Terms of Reference, “Mid Term review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks.” Reference Number 13/000428.

ual and other support documentation provide ‘high level’ standards and/or “principles,” including utility, credibility, impartiality, and independence, amongst other issues. These are included as the basis for work by other international bodies and groups on evaluation, e.g. OECD DAC, the World Bank’s IEG, UNEG, et. al. Yet, these principles are sufficiently vague to allow for a whole range of approaches, from those that simply rely on the opinion of subject matter experts to those that use big-data analytics to assess diverse data sets. The gap between principles and actual practices is especially apparent in relation to ‘evidence based’ conclusions. There is little agreement, at least among international evaluation groups, about which data collection and analysis methodologies are best for different types of development activities and/or evaluation questions.

This can lead to confusion and frustration for those who commission evaluations but do not come from the field. They can be either subject to the credibility and reliability of the consultants at hand, a subjective and complicated level of assessment, and/or to their own experience and common sense about what they need and expect from an evaluation. This diversity of opinions about what constitutes quality is seen in the qualitative evidence collected as part of this Review. (See Graph 11 below.) Different cohorts, or more particularly, different individuals, assess quality differently.²⁵

GRAPH 11: Overall Quality of Services (Evaluations) ALL COHORTS (A2)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139
140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179
180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199
200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219
220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
240																			

The only discernable trend in this qualitative data relates to team composition. If the team is not right, many state, then quality suffers. While there may be some truth to

²⁵ A similar mixed trend emerges when qualitative data from Sida Commissioners and Partner Organizations is analysed. See page 13, “Data & Analysis Report,” Mid Term review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks, November 2013.

this, the randomness associated with individual capacity, let alone the dynamics between people on any evaluation team, make this very hard to judge. Is it always just about the team? Or, should some common standards apply?

The primary issues raised by respondents regarding these issues include:

- The most commonly cited issue regards the teams, including their composition and competencies. This includes both ‘positive’ and ‘negative’ comments. Some Sida Commissioners and Partner Organisations noted the exceptional competence of teams while others noted that some consultants did not have the requisite skills in the field/subject under evaluation, that there were issues of credibility and independence, and that some team members needed to be changed either at the insistence of the Sida Commissioner or for other reasons. None of these issues were common enough across cohorts to signal an issue that could be linked to the FA, the Service Providers, or the Sida Commissioners. Instead, they tend to highlight the complexity of both team composition and the subjects being evaluated.
- As elsewhere, a significant number of respondents state that the ease and frequency of communication with the Service Providers was a ‘positive’ aspect of quality.
- Partner Organisations, Sida Commissioners, and Service Providers state that the length of the assignment was not ‘ideal’ for the subject and that this resulted in sense of ‘being rushed’ or not having sufficient time to adapt and change given findings or other issues. As noted in Section 4.2.2.2, this may be related to enhanced engagement/utilisation levels, especially during the inception phase.
- It should be noted that, while this Review raises issues of quality related to methodologies, data collection and analysis, and other critical evaluation techniques, respondents except in a few instances did not raise these. This is not surprising. One can hardly judge something in absentia nor on techniques and approaches that are not from one’s specific field.

It is not surprising that a broad cohort would have mixed views on quality: it speaks to different expectations and no commonly understood standard. The qualitative evidence above demonstrates that most Sida Commissioners and Partner Organisations judge quality on output (the report, as described below) and on the team (its subject matter expertise and independence). Service Providers largely fault any deficiency in quality as due to limited resources or unrealistic expectations. They do not judge quality, by and large, according to evaluation standards and common techniques.

This lack of a common understanding of evaluation standards amongst relevant cohorts implies that one must turn to each element in turn (process, output, and outcome) to gauge evaluations’ quality. This Review contends that adherence to these standards is the best way to ensure quality and to move beyond the sometimes-subjective views of people involved in evaluations, as exhibited in the evidence above. This includes the distinct elements of evaluation (process, outcome, and out-

puts) and the interplay between them. All are important and all include common approaches, as established in the standards mentioned above, that can ensure evaluation quality, especially as regard to data collection, analysis and subsequent evidence based conclusions and recommendations.

In regard to **process**, utilisation and engagement are a useful way to increase learning and stakeholder acceptance of an evaluation's results.²⁶ As noted in Section 4.2.2.2, there is evidence that there are significant levels of utilisation and engagement in the evaluation services under review. Further analysis indicates that this is most likely due to UTV and Service Partner's stressing the importance of utilisation/engagement. However, utilisation is not sufficient in and of itself to guaranty quality. An evaluation team may discuss and engage stakeholders at various points but this can lead as much to personality-driven outcomes as ones based on sound evaluation processes. If people 'like' one another, they may be more prone to forgive some deficiencies or to ignore areas of contention or controversy. Evaluators need to engage stakeholders to support a common understanding of the process and expected outcomes but they should do so alongside a solid evidence-based, methodology. The evidence should dictate the subject of the discussions. If the evidence is uncertain, or if a confirmed methodology is not followed methodically, then discussions may be based on what people believe rather than what they know.

This Review also treats the **outcome**: mainly how Sida Commissioners and Partner Organisations use evaluation results. Toward this, there is evidence that both Sida Commissioners and Partner Organisations use evaluations to improve the performance. (See Section 4.2.2.3.) Sida Commissioners use the evaluations as tools to work with the Partner Organisations, to facilitate increased programme performance. Partner Organisations, use conclusions and recommendations to make practical changes. This is positive and may be related to the noted utilisation and engagement levels. Yet, the overall lack of evidence-based approaches may mean that subsequent conclusions and recommendations miss key issues, that they reiterate what is already known and what is therefore generally acceptable to all stakeholders.

This leaves the reports themselves, or the primary **outputs** from evaluation services. This Review included an assessment of 80 reports. (Please see the "Data & Analysis Report" for more regarding this assessment and the Annexes of this document for a

²⁶ Michael Quinn Patton, *Utilization-Focused Evaluations*, SAGE Publications, 2008. (4th edition) Of course, this is an off-shoot of 'Practical/Participatory' theories of evaluation and should be compared to values-based and emergent-based theories. Indeed, this review falls more in the emergent-based field of evaluation theories. For a good description of the mechanics of all three, see Mark Hansen, Marvin C. Alkin, and Tanner LeBaron Wallace, "Depicting the Logic of Three Evaluation Theories." *Evaluation and Program Planning*, Elsevier, March 2012.

list of the reports assessed. The metrics and their calibration with those in the “Are Sida Evaluations Good Enough” report are described in the section on “Methodology” above.) This included 50 metrics organised according to 7 areas, as included in the table below. This shows that most areas fall between a “4” and a “5” on a 6-point scale since the beginning of the FA:

Table 4: Average Rankings per Assessment level for ALL Reports under Review			
ALL REPORTS	Average (TOTAL)	Average (Last 12 Months)	Rankings
	ALL (80 Reports)	Aug '12 - Jul '13 (41 Reports)	
Terms of Reference	3.96	4.08	6 – Excellent
Evaluation Questions	4.19	4.38	5 – Adequate
Methodologies	4.28	4.48	4 – Minimally adequate
Data Collection	4.39	4.45	3 – not quite adequate
Data Analysis/Analytical Content	4.18	4.26	2 – significant problems
Conclusions & Recommendations	4.48	4.57	1 – very poor
Reporting	4.75	4.80	
Data Analysis/Analytical Content, Conclusions & Recommendations, and Reporting Average	4.41	4.54	
Total Components Average	4.31	4.41	

A similar pattern emerges for only evaluations under assessment:

Table 5: Average Rankings per Assessment level for Evaluation Reports under Review			
EVALUATIONS ONLY	Average (TOTAL)	Average (Last 12 Months)	Rankings
	ALL (41 Reports)	Aug '12 - Jul '13 (25 Reports)	
Terms of Reference	3.73	3.80	6 – Excellent
Evaluation Questions	4.37	4.58	5 – Adequate
Methodologies	4.43	4.54	4 – Minimally adequate
Data Collection	4.40	4.44	3 – not quite adequate
Data Analysis/Analytical Content	4.30	4.29	2 – significant problems
Conclusions & Recommendations	4.62	4.65	1 – very poor
Reporting	4.88	4.90	
Data Analysis/Analytical Content, Conclusions & Recommendations, and Reporting Average	4.60	4.61	
Total Components Average	4.39	4.40	

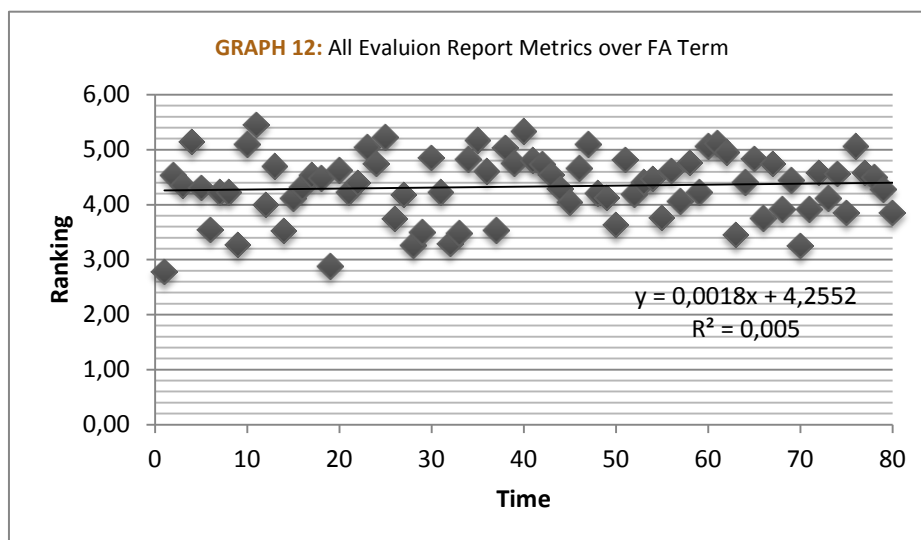
This indicates that, while there may be acceptable levels of utility, of the outcomes, the reports demonstrate a broad breadth of quality, both positive and negative. Each of these areas are discussed below.

Overall

In considering the rankings across all metrics, nearly all measures fall below 5.00 or ‘adequate’. (<4.50). Graph 12 shows the overall ranking of all reports from the beginning of the FA to July 2013. This shows a slight increase overall ($R^2 = +0.00497$, or a .4% increase²⁷) although there is little clustering with as many reports following below the mean as above it.²⁸ A similar spread is seen if only the reporting elements are included. (Graph 13.)

This analysis implies, that some evaluation services have been good (13 >5.0) while a similar number can be judged as having problems (24 <4.0). Some evaluations (13) demonstrate a ‘good’ level of quality across metrics while approximately 24 do not. This leaves the majority (43) somewhere in the middle.

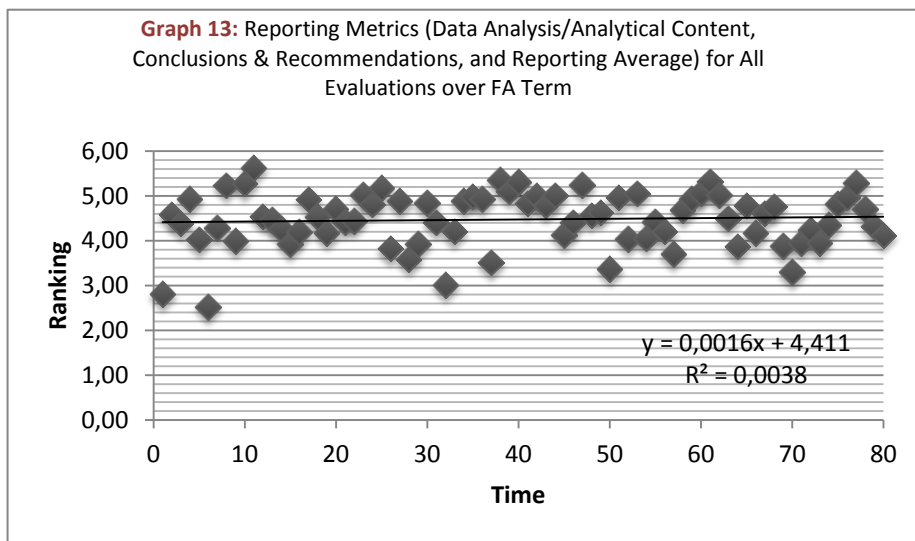
Similar trends occur when each evaluation category is analysed and compared to rankings from “Are Sida Evaluations Good Enough.” (ASEGE)²⁹



²⁷ R2 refers to the calculation to determine the average increase over time given the spread of ratings related to a common mean.

²⁸ This review has calibrated the metrics between this assessment and the “Are Sida Evaluations Good Enough” metrics. For a description of this, see “Calibrating ASEGE Metrics” in the “Methodology” section above.

²⁹ “Are Sida Evaluations Good Enough? An Assessment of 34 Evaluation Reports,” Kim Forss, Evert Vedung, Stein Erik Kruse, Agnes Mwaiselage, and Anna Nilsson. Sida, 2008. We have taken the rankings and produced averages for each category to establish a fair comparison with the rankings used in this review. However, the “Are Sida Evaluations Good Enough” approach included a ranking of ‘1’ any time an element was missing while this review assesses the importance of the element and its mention in the Inception Phase report, thus resulting in most instances classified as “ND” or “not done”. To rectify this, we have removed the relevant rankings of “1” so that there is a more accurate comparison.



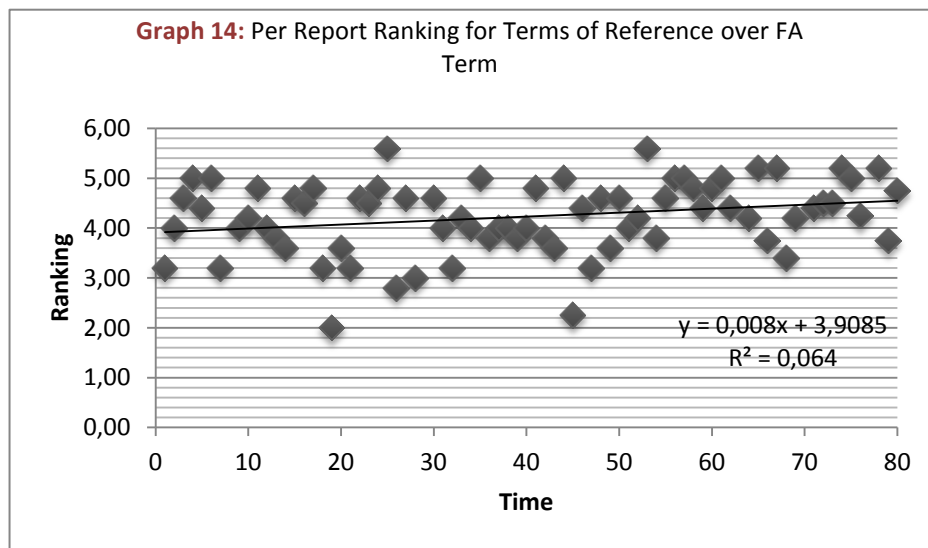
Terms of Reference

In relation to the **ToR**, this Review’s assessment ranked these overall at 3.96. ASEGE ranked the ToRs’ under its consideration at 4.40/ σ 0.19.³⁰ This implies, when considering the standard deviation within the ASEGE rankings and aligning that with this Review’s rankings, that the current ToRs are of slightly lower quality overall.

The ToRs, however, have also shown the greatest degree of improvement since the FA’s inception. (See Section 4.2.2.1.) As some SP managers argue, the quality of the ToR has a direct bearing on the quality of the evaluation itself. If the ToR is vague or overly ambitious, their argument goes, then an evaluation will tend to be correspondingly vague or ambitious. For related qualitative evidence, see the “Data & Analysis Report”, page 269 – 279.)

This Review could find no research that demonstrates the link between the ToR and an evaluation’s eventual quality. Common practices dictate that issues related to a ToR are addressed in a proposal or during the Inception Phase. The nature of the FA means that the discussion of what is feasible and not, what limitations and constraints exist, how evaluation questions may be adapted and changed, can be highlighted in a proposal and then discussed and agreed in the Inception Phase. This is how it is largely done in any procurement arrangement.

³⁰ σ 0.19 refers to the standard deviation from the mean associated with the ASEGE data sets. In other words, anything above or below 0.19, in this case, can be considered statistically relevant. Thus a 3.96 overall ranking is below this ± 0.19 statistical threshold. For more on how these metrics are calibrated, see “Calibrating ASEGE Metrics” in “Methodology” above.



At the same time, this assessment of the ToRs indicate that the most common “weakness” is that they missed certain elements, e.g. evaluation questions of OECD DAC levels, or that they were sufficiently vague to allow for just about any type of approach. Finally, some were unrealistically ambitious, seeking a level of analysis about impact or effectiveness that may not have been possible given available resources.³¹

Evaluation Questions

Evaluation questions from the ToRs, Inception reports and Final Reports were reviewed with those included in the Final Report assessed according to their specificity, feasibility, and relation to their organisation to the OECD DAC evaluation levels (relevance, effectiveness, impact, efficiency and sustainability.) Overall, these were ranked at 4.19 with a significant increase for evaluations only (4.37).³² While the evidence is not conclusive, this may also be due to an increased stress of this by UTV, up-dated templates for the ToRs, and Sida Commissioners increased recognition of the importance of evaluation questions that abide with OECD DAC evaluation levels and standards. However, this is limited to only 25 reports out of 80 (Table 5) and still represents a middling ranking overall.

³¹ We stress the conditional here as a greater attention to methodologies, data collection and analysis, and other analytical tools could allow a Service Provider to develop processes and approaches that become less resource intense over time. They establish and then continuously improve the methodologies so that they do become more possible even with limited resources.

³² In fact, most of this increase occurred in the last 12 months. As Table 5 above shows, the average ranking for evaluation questions for evaluations only is 4.58 making it second only to Reporting in ranking.

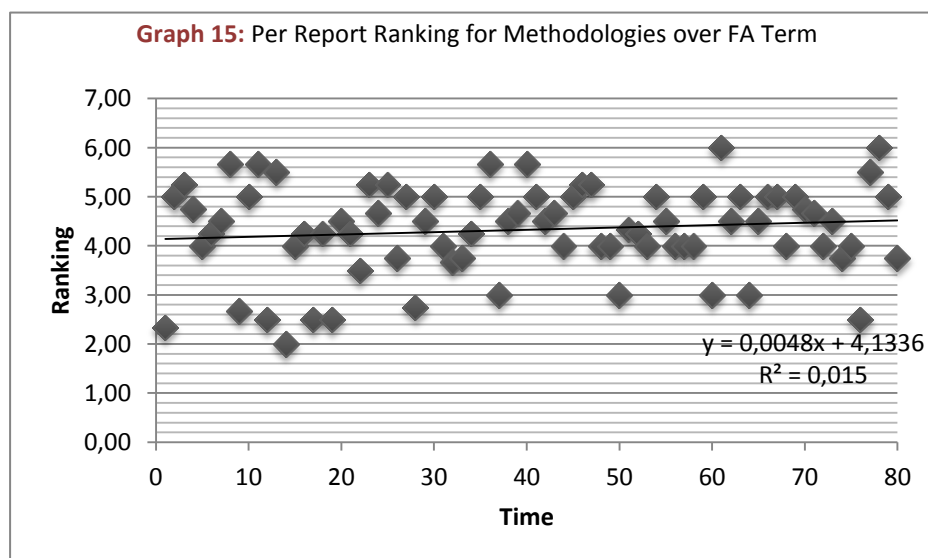
While this represents a relatively low overall ranking for Evaluation Questions, we have weighed this against a review of the Inception Phase reports where different levels are often changed or adapted, or removed entirely as with ‘impact’ in many cases. We have also compared this with the qualitative evidence that suggests significant levels of engagement/utilisation and where the purpose of the evaluation, its process and outputs, can be assumed to be discussed and agreed upon between stakeholders.

In general, this is an area where improvements could be made. Evaluation questions serve as the basis for all else in an effective evaluation. The breadth for the questions, their implications regarding evidence, constraints and limitations, and how they relate to practical considerations going forward, constitute an essential evaluation element. These were missing in many reports, not followed from the Inception Phase to the final report, and/or unevenly treated in over a third of the reports assessed.

Recommendation 2 (UTV): UTV should stress the need for evaluation questions that follow OECD DAC evaluations levels to Sida Commissioners. While there may be reasons to divert from or change the OECD DAC levels, a set of comprehensive and valid questions should be included in every ToR. UTV should also ensure that Service Providers follow these questions, developing an evaluation matrix that includes all questions, their data sources, limitations and constraints, and risks and opportunities. This Evaluation Matrix should be mandatory for all Inception Phase Reports. (Section 4.2.1.1)

Methodologies

Methodologies do present a significant change from the ASEGE rankings. This Review ranked overall methodology quality at 4.28 in comparison to 3.92/ ± 0.15 in the ASEGE assessment.



Our assessment indicates that methodological sections explained different approaches and tools, their utility and limitations, and how these related to the evaluation's purpose. These, as with our overall assessment, remain inconsistent, with some far better than others. However, there is a demonstrable increase in methodological quality over time, as described in Section 4.2.2.1.

This assessment is based on the Methodologies as written, not whether they are then used. As noted in the sections below, there is a noticeable gap between how these Methodologies are written and the actual data, collection, analysis and evidence for conclusions and recommendations.

Data Collection

Both the ToRs and Methodologies have indirect links to actual evaluation reports. For instance, Data Collection and Analysis and the Analytical Content overall are directly related to how evaluators consider and analyse the data in relation to evaluation questions and form evidence based conclusions and recommendations.

This Review ranked Data Collection at 4.39 overall and 4.40 for evaluations only. In fact, very few evaluations actually described or included data collection methods or results. Our analysis is based on the few instances when this did occur. Less than 10 reports included specific benchmarks, indicators, or protocols for data collection or analysis as related to the evaluation.³³ These were sometimes included in Inception Phase Reports although here to, the methodologies were not explained in relation to specific issues and/or evaluation questions and not subsequently referenced in the final reports. (See “Data & Analysis Report” from page 45.) Overall, the presentation of data, methodologies for its collection, and their relevant information are treated with a bare minimum and not enough to inform a reader about the nature of this data, its constraints and limitations, and why certain data collection methodologies were chosen over others.

This is of particular concern as nearly all evaluations included field visits. If interviews, in particular, do not include standard protocols and some record, however minimal, of trends across this qualitative evidence then they will tend to play into natural human biases.³⁴ As this and other literature suggest, conducting interviews is a high-

³³ Best practices would dictate that these be aligned with a project/programme's performance indicators. However, these are often inadequate or too far removed from the subject of an evaluation or specific evaluation questions. This implies that evaluations need to develop specific indicators for how they will establish evidentiary trends or other issues related to specific evaluation questions.

³⁴ The literature on this point is vast and growing. For a few good examples of the current state of the

ly problematic source of data as it tends to amplify individual preferences and biases. We humans tend to ‘hear’ what we already ‘understand’ and ‘believe’ from people for whom we have some personal affinity. The first aspect of this is problematic because one may not collect, or ‘hear’, a great deal of relevant data while the second part can lead to collusion, obviously, but also to being dismissive of anyone who may be significantly different, a highly troubling issue in the context of international development. This is why so many standards and best practices are emerging for how one collects and analyses qualitative data. This includes the use of standard protocols, asking the same question in the same way to a discrete sample group and ensuring that no one respondent, whatever their position or particular insight, can influence findings and conclusions—there need to be evidentiary trends across qualitative data for any such finding.

This can be frustrating to those commissioning evaluations because this more standardised approach to qualitative data can often reveal only the broadest trends. It seldom can access, by design in fact, the particular nuance or insights of one or two particularly articulate and/or informed respondents. Their comments shouldn’t be ignored but they also can’t be held as definitive evidence. It is one person’s perspective and it is up to the evaluator to then judge this, especially if it is in contradiction to evidentiary trends, as based on best practices, research, and their own expertise. It cannot, however, be held up as ‘qualitative evidence’.

These and other issues related to the complicity of data, be it qualitative or quantitative, implies that there needs to be a thorough and validated description of how evaluators will collect and analyse data and the constraints, limitations and opportunities associated with those choices and how they relate to specific evaluation questions. This methodological discussion, for the issues cited above and given the often disparate understandings of what constitutes quality, makes this methodological discussion of paramount importance.

Recommendation 3 (UTV & Service Providers): Provide Service Providers with expected standards and practices in relation to data collection methodologies and approaches, especially as they apply to field visits. Ensure that Service Providers improve data collection methodologies and approaches. These should adhere to com-

field, see David R. Thomas, “A General Inductive Approach for Analysing Qualitative Evaluation Data,” *American Journal of Evaluation*, Vol. 27. No. 2, June 2006. For a useful, if dated, primer, see Michael Quinn Patton, *How to Use Qualitative Methods in Evaluation*. SAGE Publications, 1987. For an interesting precursor of future trends, see David Byrne, “Evaluating Complex Social Interventions in a Complex World,” *Evaluation*, Vol, 19, No. 3, July 2013. It should be noted that the use of qualitative data in this review achieves only the most minimal standards in this regard. It allows for broad evidentiary trends and conclusions. However, much deeper analysis could be possible from this foundation.

mon standards for mixed-method evaluation approaches and include standard protocols and analytics for any relevant qualitative evidence. Any data collection methods should have direct relevance to established evaluation questions.

Data Analysis & Analytical Content

Given the dearth of solid evidence for quality data collection, it may not be surprising that Data Analysis/Analytical Content has the lowest ranking amongst items directly related to the report, e.g. minus the ToR and evaluation questions. The Review ranked these at 4.18 overall and with an improvement for evaluations (4.30) although this too remains the lowest average among the seven categories. ASEGE ranked this at 3.96/ $\sigma 0.35$ which, given the standard deviation, shows no substantive difference. When compared across time, data analysis shows the widest spread or most divergence from the mean. Thus, not only is the quality in this area low, there has been little change in reference to the ASEGE analysis.

While hardly conclusive, this Review suspects that this is often due to the somewhat out-dated notion that a suitable evaluation may be achieved solely through the opinion of a subject matter expert. While the line between opinion and fact is difficult in any evaluation, we found many occasions when the evaluator offered their opinion without any substantive evidence, be it from the evaluation itself or from best practices and leading research. (These instances are duly noted in the “Data & Analysis Report” and provide the basis for 2/3 of those rankings that fall below “3” in this category.) This does not imply that these subject matter experts are wrong; simply that the complex nature of international development and cooperation implies that even the most robust expertise needs to be grounded by solid evidence and analysis.³⁵

Conclusions and Recommendations

In the case of conclusions and recommendations, this Review ranked this relatively high (4.48 overall/4.62 for evaluations only); second only to reporting. This represents a slight improvement from the rankings in ASEGE (4.15/ $\sigma 0.27$).³⁶ In general

³⁵ Most accept ‘evidence based’ as a common standard for evaluations. For some literature and best practices, see: Alistair Hallam, “Harnessing the Power of Evaluation in Humanitarian Action: An Initiative to Improve Understanding and Use of Evaluation. ALNAP, June 2011; “Quality Standards for Development Evaluation.” DAC Guidelines and Reference Series, OECD, 2010; “Evaluating Development Activities: 12 Lessons from the OECD DAC.” OECD, 2013. “Monitoring & Evaluation: Some Tools, Methods & Approaches.” The World Bank, 2004; and the “Sourcebook for Evaluating Global and Regional Partnership Programs: Indicative Principles & Standards.” The Independent Evaluation Group—The World Bank, 2007.

³⁶ $\sigma 0.27$ refers to the standard deviation from the mean associated with the ASEGE data sets. In other words, anything above or below 0.27, in this case, can be considered statistically relevant. Thus a

our assessment indicates that the conclusions and recommendations were logical in that they followed from the findings that were presented (even if these were not based on solid evidence and/or analysis) and were largely actionable and directed to specific stakeholders.

Reporting

As noted, reports were well written and they presented frank opinions and findings. This Review ranked Reporting overall at 4.75 and higher for evaluations (4.88). This represents the most significant change from the ASEGE rankings that ranked reporting at 4.12/ ± 0.14 . As noted in Section 4.2.3.1, this may be attributed to Quality Assurance (QA) processes that focus on reporting.

As this analysis demonstrates, overall quality is mixed with some areas where significant improvements could be made. The FA intends to raise the standard of evaluations overall. The FA recognises that the ranked system would mean that most work would go to one supplier but saw that this supplier would be best positioned to provide high-quality evaluation services and that their expertise would grow. On the whole and as presented, evaluations are not significantly better than those assessed in ASEGE nor do they uniformly achieve common standards and best practices, especially in relation to data collection and analysis which may be considered critical aspects of any evaluation.

Comparisons between three Service Providers

The Review sought out noticeable differences in reporting and other quality measures between the three Service Providers. This is limited by the number of evaluations conducted by SIPU and Orgut; there simply isn't enough of a sample to judge different quality criteria for their work. Overall, however, there were no substantive differences. Some of the reports under review were exceptional either overall or in relation to specific criteria while others were not. Any further analysis or findings is not possible given the way that these Service Providers assume work under the FA, the volume of work they have done, and other factors beyond this review.

3.2.2 Changes in Quality

Evidence demonstrates increased quality in ToRs and in how methodologies are described, even as those for ToR remain low as compared to the ASEGE rankings. Evi-

4.4.8 overall ranking is just above this ± 0.27 statistical threshold. For more on how these metrics are calibrated, see "Calibrating ASEGE Metrics" in "Methodology" above.

dence also demonstrates positive levels of engagement and utilisation that is a feature of quality but not the sole determinant of quality. This Review attributes the change in methodologies to the stress that UTV staff put on this early in the process. The increased quality of ToRs may be attributed to the repeated use of evaluations by specific units or embassies and the increased support provided by UTV initially, as least, and by Service Providers on individual assignments. However, procurement regulations according to UPPIS prevent a discussion about the ToR between the Sida Commissioner and the first ranked Service Provider thus delaying any such discussion to the Inception Phase.

Utilisation/engagement is used with suitable regularity, given this analysis, and that there is a positive trend in evidence related to how Partner Organisations use the evaluations' results. This indicates that while the evaluations may not always achieve the highest rankings in different criteria, they may still be having a positive impact. As noted in "process" in Section 4.2.1, this does not imply that the best conclusions and recommendations are presented and may actually imply that conclusion and recommendations are simply those that are more amenable to relevant stakeholders.

3.2.2.1 *Has evaluation quality changed since the signing of the Framework Agreement? (A2.2)*

Overall quality is not much better than before the FA and it does not reach common standards and best practices in methodologies, data collection, and analysis, issues critical for any evaluation. (See Section 4.2.1.) In essence, some evaluations are good and others have significant problems and there has been little change in this over time.

As with quality overall, changes in quality are difficult to gauge without a clear baseline or comparatives. Qualitative evidence is bound to the present, as are quantitative surveys. The only time specific data available are the evaluation reports. These represent a sufficient and nearly equal spread across time from the FA's inception to July 2013. (See "Data & Analysis Report" for the dates for specific reports under assessment.) This allows us to assess any significant changes across all metrics and per category.

While most categories showed negligible change over time, Methodologies and ToR did show a significant positive change. (See Graphs 14 & 15 below.)

ToRs have the more pronounced change over time ($R^2 = +0.06399$).³⁷ Unfortunately, there is not sufficient evidence to establish what may contribute to this. Limited evidence suggests that this has occurred because of the increased stress put on the ToRs by the Service Providers and the direct guidance by UTV. As established in relation to utilisation and engagement, Sida Commissioners and Partner Organisations often remark on the comprehensive nature of the Inception Phase wherein evaluators work with stakeholders to move from the ToR to a plan for the rest of the evaluation. The Inception Phase became a primary point when the Service Providers could comment upon the overall quality and/or feasibility of the ToR. This is related to UPPIS ensuring that contractual terms are honoured, mainly that there cannot be a discussion about the ToR during procurement but only after a Service Provider is awarded a contract. Related qualitative evidence indicates that the ToRs' quality was part of these discussions. In addition, UTV has stressed the use of the ToR template included in the Sida Evaluation Manual and has launched an improved template recently.³⁸

The increase in Methodologies may be more clearly attributed to the stress that UTV placed on this during the first year of the FA. This is confirmed through qualitative remarks from respondents within UTV and Service Providers who noted this as an important aspect of the FA's implementation.

Of course, these remarks are limited to a small sample of respondents, mostly those directly involved in the FA at the time. However, the fact that nearly all Service Partner management respondents confirmed this places a sufficient level of credibility on this finding.

3.2.2.2 Have the way in which Evaluations are commissioned under the Framework Agreement changed engagement/utilization levels? (A2.2.2)

Utilisation and engagement were stressed as critical features of the FA by UTV staff directly responsible for its development and implementation. Several people from this cohort explained that, given limited resources and an exceptional need to improve quality, they focused on utilisation, especially during the inception phase, as a key to greater quality. This abides by engagement theories and best practices related to evaluation.³⁹ Instead of evaluators working at a distance from commissioning officers, as

³⁷ This means that there has been a 6% increase overall in rankings overall.

³⁸ Stefan Molund and Goran Schill, "Looking Back, Moving Forward: Sida Evaluation Manual," 2nd revised edition. Sida, 2007.

³⁹ For a summary of the theories associated with engagement, see: Mark Hansen, Marvin C. Alkin, and Tanner LeBaron Wallace, "Depicting the Logic of Three Evaluation Theories." Evaluation and Program Planning, Elsevier, March 2012. For best practices, see: Michael Quinn Patton, Utilization-Focused Evaluations, SAGE Publications, 2008. (4th edition)

a way to ensure independence and distance, engagement holds that a closer dialogue will increase appreciation for the evaluation process and results. It will generate greater understanding and acceptance for the findings and recommendations. At the same time, we note specific limitation related to engagement/utilisation and the possible pitfalls associated with deeming this as a primary feature of quality. (See Section 4.2.1)

Evidence indicates that over half of respondents have seen positive aspects of utilisation and engagement. (See Graph 16.)

GRAPH 16: Have the way in which Evaluations are commissioned under the Framework Agreement changed engagement/utilization levels? (A2.2.2; combined cohorts: Partner Organisations (evaluations), Sida commissioners, SP management, core team, consultants)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139
140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179
180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199
200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219
220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259
260	261	262	263	264	265	266													

Even when respondents' remarks were ranked 'negative' this often had to do with issues that prevented engagement/utilisation, e.g. poor/wrong teams in place and the budget and time allowed. These 'negative' responses recognise the importance of utilisation even when not possible.

While this qualitative evidence depicts an overall trend, some of the particular trends among cohorts and individuals are worth noting:

- The lack of a common understanding about engagement/utilisation amongst many respondents implies that most inferred 'engagement/utilisation' to imply solely a participatory approach. At the same time, we considered specific comments about participation as related to engagement/utilisation and so assessed these as 'positive.'
- Many Sida Commissioner and Partner Organisation respondents commented on the 'open' discussions and a focus on 'learning' that exists with the consultants.
- Some Sida Commissioner and Partner Organisation respondents commented on the benefit of having multiple perspectives being brought to bear on the evaluation's design and results.

- Some Sida Commissioner and Partner Organisation respondents commented that the open dialogue ensured that when there were issues or concerns these were treated quickly.
- Most ‘positive’ comments (from all cohorts) focus on the Inception Phase and related meetings/workshops and discussions about the report.
- Some SP core team and consultant respondents state that the ability to engage depends on the Sida Commissioners and their willingness to invest the time to do so.
- Partner organisations, as elsewhere, state that an inordinate amount of time was spent discussing the nature of the subject or organisation under evaluation rather than how the evaluation would be conducted and its expected results.
- In relation to ‘negative’ comments, all cohorts state that the time and resource pressures limited opportunities for engagement/participation.
- Some Sida Commissioners and Partner Organisation respondents questioned the increased time during the inception phase and whether this took away from the time spent during the field/data collection phase.

Perhaps more significantly, a noticeable proportion of Core Team members interviewed did not put appropriate stress on engagement/utilisation. Some expressed that this was not necessary and that it could interfere with one’s independence. (See “Data & Analysis Report” pages 271 – 281.)

Recommendation 4 (UTV & Service Partners): Service Partners should ensure that all Core Team members and team leaders recognise the importance of utilisation, especially during the inception phase, and that they are sufficiently versed in best practices for engagement. Include this as a primary element in Quality Assurance (QA), e.g. to ask and report on how Core Team members have successfully engaged stakeholders during and beyond the inception phase. UTV should also make every effort to ensure that Sida Commissioners understand the implications of engagement/utilisation and that it may require more time/resources, especially during the inception phase, to ensure its value.

3.2.2.3 *Have the way in which Evaluations are commissioned under the Framework Agreement changed the way they are used by Sida and its partner organisations (including both instrumental and process use)? (A2.2.3)*

A sufficient level of instrumental use is apparent, e.g. Sida commissioners and Partner Organisations have used the evaluations to improve performance:⁴⁰

GRAPH 17: Have the way in which Evaluations are commissioned under the Framework Agreement changed the way they are used by Sida and its partner organisations (including both instrumental and process use)? (A2.2.3: Partner Organisations)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			

The statements from respondents in this regard were largely positive, as the graph indicated. The most common statements regarded the recommendations and how these were either being implemented or led to more informed thinking about what changes could and should be made. The few ‘neutral’ and ‘negative’ comments regard recommendations that were not implemented either because they were not directly related to the organisation or because they “missed some of the complexity of the context and this has allowed certain parties to play into existing political structures; this takes a lot of time to overcome” as in the only ‘negative’ comment.

This demonstrates that nearly all Partner Organisations said that they have made changes due to the evaluation, be it the process or report. When this qualitative evidence is analysed further, it shows that Partner Organisations mention an array of changes, from specific changes to project design to improved ways of working with Sida desk officers. (See “Data & Analysis Report,” pages 283 – 285.) This was confirmed in the quantitative survey conducted as part of this Review. (See Graph 18.)

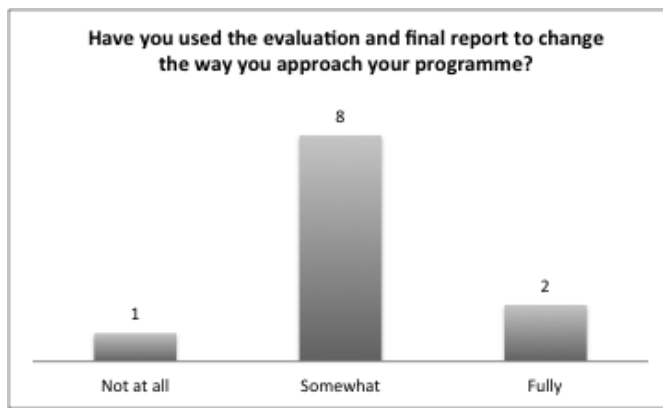
In relation to the FA, the way in which evaluations are commissioned does not have a direct bearing on their utility. While this Review demonstrates that the FA does facilitate the efficiency by which evaluation services are procured it has less of a bearing on how they are actually used. This depends instead on utilisation and engagement levels and the specificity and practicality of the resulting conclusions and recommendations. In this sense, the way evaluations are used is more related to the quality of the evaluations themselves and while this, as described throughout this Review, does

⁴⁰ For reading this and similar graphs, please refer to “Using Qualitative Graphs” in the “Methodology” section.

have some bearing on how the FA it is sufficiently removed to have less of direct correlation to their actual commissioning.

This is further supported by qualitative evidence from Sida Commissioners. When asked about the quality of the evaluation outputs, mainly the reports, a significant proportion of these respondents state that the report gave them a useful lever for working with Partner Organisations. It was described as just such a tool and one that enabled them to do more than if nothing had been done. (See “Data & Analysis Report”, pages 265 – 271.)

GRAPH 18: Quantitative Survey (Partner Organisations)



While this Review finds that the overall quality of reports does not adhere to some best practices and common standards (See Section 4.2.1), this does not necessarily impede the utility of the evaluation for specific Sida Commissioners and/or Partner Organisations. They may find the process, report and recommendations useful and may, as indicated in this Review, make changes due to these.

However, if these recommendations are not based on solid evidence, evidence drawn from appropriate data collection and analysis methodologies, they may not prove effective. If an evaluation is not evidence based, how does one know if the recommendations are the best way to improve a project or programme? If engagement/utilisation levels are sufficient, as indicated, but are so in the absence of established data collection and analysis methodologies, e.g. evidence based conclusions, how does one know if Partner Organisations implemented recommendations or used the report because of their, perhaps biased, appreciation of the consultant? Engagement/utilisation without evidenced based conclusions could lead to ‘utility’ that is simply based on how Partner Organisations ‘got along’ with the consultant.⁴¹ (See Section 4.2.1 and 4.2.2.2.)

⁴¹ This is, of course, tempered by concerted efforts to establish independence and yet the risk remains without solid evidence, a strict methodology, and standardized protocols for any interaction. As noted elsewhere, this does not imply that evaluations need to be so methodologically refined as to be mechanical. However, there should be a methodological discussion with clear rationales for different choices and tools.

3.2.2.4 *Can changes in the quality of evaluations be attributed to the Framework Agreement? (A2.2.1)*

There have been some positive changes in relation to the Terms of Reference and Methodologies. These may be related to the on-going partnership with one primary Service Provider, although there is no direct evidence for this. The FA provides a channel for UTV to express the need for greater attention to methodologies, for instance. The same could be said for the increase in quality ToRs. If this is related to more substantive inception phases and higher levels of engagement and utilisation, then this may be related to having the same Service Provider work with Sida Units and Swedish embassies repeatedly—they may improve how they work together and if the ToRs are stressed then one would expect demonstrable levels of improvement in this area.

This is even more compelling in the case of increased engagement and utilisation. This was a key element in the FA’s architecture and how UTV and SP managers expressed it in interviews. While comparative evidence does not readily exist, levels of engagement are high. If the theory holds true, this increases learning across stakeholders and can increase the understanding and acceptance of findings and recommendations. Indeed, while the causal link would be very difficult to determine, evidence above also demonstrates a fairly high degree of change in how Partner Organisations approach their programmes as a result of the evaluation process. Nonetheless, this increased utility, especially if solely related to engagement/utilisation, presents pitfall and issues that may undermine this seeming success. (See Sections 4.2.1 and 4.2.2.2.)

3.2.3 Service Partner Development

3.2.3.1 *Is the Framework Agreement supported by a standardized/common approach to the quality assurance of evaluations? (A2.2.6)*

Documentary review and qualitative evidence from Service Partner Management and from Core Team respondents indicate that quality assurance (QA) is an essential part of reports and in the analytical thinking and argumentation developed in these reports. While a similar focus on reporting existed at all three Service Partners this was, perhaps because of the volume of work and related standardisation/practice, especially pronounced at Indevelop. Qualitative evidence from Service Provider Management and Core Teams confirms this. (See Graph 19.)

This is further confirmed through the quantitative survey wherein 56% of Core Team Members replied “fully” to the question above (A2.2.6, Graph 19) with another 31% stating “somewhat.” There was a greater spread in responses amongst SP Management. (See “Data & Analysis Report,” page 24.)

GRAPH 19: Is the Framework Agreement supported by a standardized/common approach to the quality assurance of evaluations? (A2.2.6; combined cohorts: SP management, coalition partners, core team and consultants)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94					

Some of the trends in this qualitative evidence include:

- Service Partner management respondents state that the QA focuses on the selection of the team and on the different reports (inception, final).
- Core Team members state that QA includes a comprehensive review of the reports and that this is coupled with constructive dialogue about argumentation, logic, and flow from findings to conclusions and recommendations.
- Core Team responses rated ‘neutral’ state that the QA is dependent on the team itself, their qualifications and how they work together, and that this can have mixed results.
- Negative comments across cohorts were mixed, with no common issue or theme although they site issues from Sida support, to the time and resources available, to a preference to have methodologies and approaches based on the subject being evaluated rather than on set standards.

While qualitative evidence suggests a range of practices that Core Team members and others constitute as QA, like the OECD DAC standards or how best to adhere to the ToR, over half of respondents mention the QA done on the reports as the most positive and prominent aspect of QA. (See “Data & Analysis Report”, pages 293 – 297.) In other words, QA focuses on the outputs, the reports, more than the process.

We explored this report-focused QA with key members of Indevelop’s management team. They indicated that this focus was largely correct. They and Indevelop’s Core Team members’ repeatedly described QA as focused on argumentation, logic, critical thinking, and the links to broader Sida and Swedish policy issues and/or cross-cutting issues. This is confirmed in the relatively high rankings for Reporting done as part of this Review. (See Sections 4.2.1.1 & 4.2.2.3.)

Yet, it is also confirmed in the significant weaknesses noted in the same reports, especially in regards to methodology, data collection, and analysis. The fact that over half of the assessed reports in this review indicate significant weaknesses implies that there are problems with the QA system. By focusing on the actual drafting of the reports, their internal logic and argumentation, this misses the opportunity to ensure that the process itself abides to appropriate utilisation and engagement levels and established methodologies that ensure that conclusions are based on solid evidence. QA

should not be simply concerned with the outputs, the reports, but also the process and how evidence, and which sources of evidence, relate to subsequent conclusions and recommendations. This level of QA can ensure that conclusions and recommendations are comprehensive, that they get to the primary factors that contribute to success, and that they can sufficiently address any gaps or contradictions that may exist within and between data sets.

Recommendation 5 (Indevelop (short term) UTV (long term)): Develop a common QA procedure for data collection and analysis during the field/data collection phase and for engagement/utilisation, amongst all other areas of the evaluation, and demonstrate how this contributes to increased results. Include these results in Annual Reports. In relation to data collection and analysis, this may include common analytics, tools, and business processes that will ensure efficiency and create economies of scale over the long-term. In relation to utilisation, this may include a standard pre- and post-inception phase interview between the Team Leader (TL) and the Service Provider Manager to facilitate the TL's approach and understanding of the importance of utilisation/engagement. This may be undertaken by Indevelop in the short term. They may develop their tool-kit and other mechanisms in collaboration with UTV. In the long term, UTV can develop guidance on QA and what it should include, especially with regard to utilisation/engagement, methodologies, data collection and analysis, and other matters it so deems relevant.

Documentary evidence reveals that the Indevelop tool-kit and other documents are provided to Core Team members on a range of subjects. This includes quality control checklists for the inception and final reports. These include questions about process, methodology and findings, conclusions, recommendations and lessons, amongst other subjects. Overall, these constitute a substantive quality assurance tool.⁴²

When taken more broadly, this Evaluation Toolkit reflects some of the issues noted in this Review. For instance, the section on 'evidence' does not treat issues of data sources, analytics, reliability, or triangulation, as best practices might dictate, but instead on "credibility".⁴³ It is not clear just how much this Toolkit is used but drawing on best practices in key areas may strengthen it.

⁴² "Indevelop's Evaluation Toolkit, Version 4.0."

⁴³ "Indevelop's Evaluation Toolkit, Version 4.0.", p. 6. The section on analysing results is similarly skewed away from what best practices state about analysis and toward making the link to broader themes at Sida. While this may be valuable, it seems an odd emphasis when these assignments are constrained by budget and time.

Recommendation 6 (Indevelop): Update Toolkit to reflect best practices and common standards in data collection, analysis, and results. Provide an overview/training on this to all Core Team members.

3.2.3.2 Have the consultants improved their evaluation competence and/or strengthened and accumulated knowledge about Sida to offer better services? (A2.2.5)

Consultants, especially Core Team members, have increased their evaluation competence:

GRAPH 20: Have the consultants improved their evaluation competence and/or strengthened and accumulated knowledge about Sida to offer better services? (A2.2.5; combined cohort: SP Core Team and Consultants)

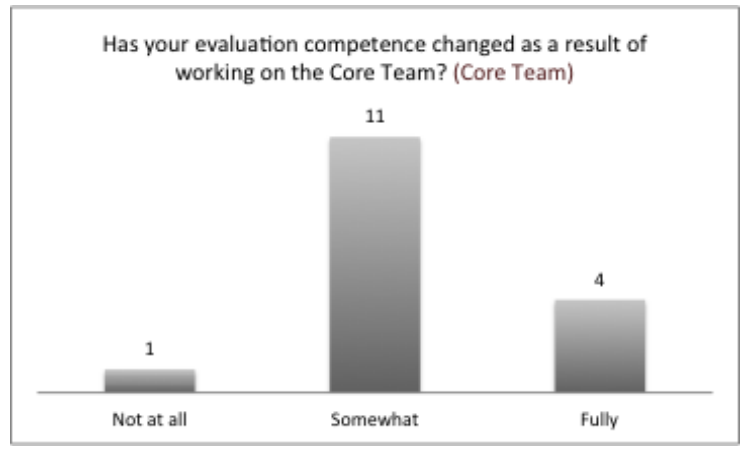
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
120	121	122																	

Common statements and trends within this qualitative evidence include:

- Respondents value the opportunity to be part of a broader ‘community’ where they discuss their work and common issues.
- Respondents said that Indevelop’s retreats were useful forums to get to know one another, share ideas and learn about new techniques.
- The Service Provider support was seen as enabling them to focus on the evaluation rather than the contracting, logistics, or other issues.
- With regard to direct improvements, some state that they have become quicker and better at writing reports and working with team members.
- The minority of ‘negative’ comments relate to the lack of resources and time for assignments and how this creates pressure to simply get the job done.

This overall positive trend is confirmed by the Review’s quantitative survey.⁴⁴ (Graph 21.)

GRAPHS 21: Quantitative Survey (Core Team)



While this constitutes self-reported data, there is little reason to dispute respondents. They were largely able to cite specific ways that their competence improved and expressed a significant level of personal satisfaction in this competence development. (See “Data & Analysis Report,” pages 287 – 293.)

This suggests that one of the most common ways that Core Team members’ have seen their competence improve is through the active community they have with other Core Team members and with the Service Provider. They uniformly recognise the value of the periodic retreats and other forums where they discuss best practices, common challenges, and other insights into evaluation.

Of course, this missing element from this qualitative evidence concerns methodologies, data collection, analysis, and evidence based conclusions and recommendations. The only methodological competence mentioned had to do with outcome mapping and this was only mentioned by a few respondents. More commonly, Core Team members discussed their improved “thinking” and “argumentation” in the reports and how they got better at drafting reports overall. This is supported by the findings related to QA that focuses on the reports. (See Section 4.2.3.1.)

⁴⁴ While not confirmed, SP Management responses may reflect Service Providers who have not completed enough evaluations to provide for a significant level of competence.

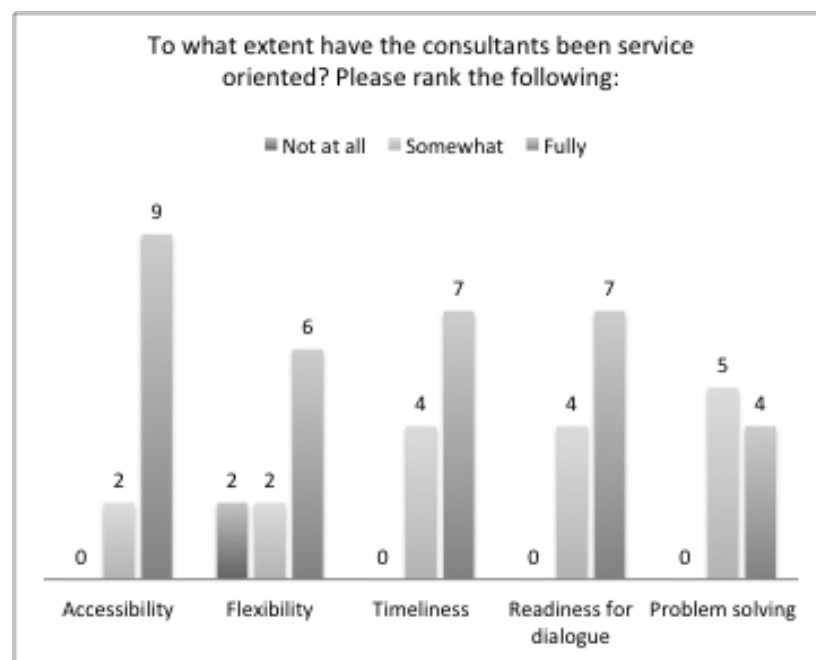
3.2.3.3 To what extent have the consultants/Service Partners been service oriented in terms of e.g., accessibility, flexibility, timeliness, readiness for dialogue and solving problems in a constructive way? (A2.2.4)

Qualitative evidence and this Review's quantitative survey indicate a high degree of service orientation. (Graph 23) This is confirmed by the quantitative survey results for partner organisations. (Graph 24)

GRAPH 23: To what extent have the consultants/Service Partners been service oriented in terms of e.g., accessibility, flexibility, timeliness, readiness for dialogue and solving problems in a constructive way? (A2.2.4; combined cohorts: Sida Commissioners and Partner Organisations)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
120	121	122	123	124															

GRAPH 24: Quantitative Survey (Partner Organisations)



Common statements and issues include:

- Many Partner Organisation comments simply state that they were 'effective' in regards to service levels; while lacking specificity this is generally judged as positive.

- Many Sida Commissioners and Partner Organisation respondents state that the communication is ‘frank’, ‘open’, ‘direct’ and friendly/constructive.
- Some Sida Commissioners commented on the ‘convenience’ of the assignment and how the consultants worked ‘fluidly’ and without many issues.
- As with ‘positive’ comments, no trend exists with ‘neutral’ and ‘negative’ comments although, as elsewhere, some issues of resources, budget, the nature of the team, and overall time allowed are mentioned.

The importance of high service levels should not be under-estimated. As an analysis of the qualitative evidence demonstrates, these service levels include a flexibility and responsiveness that is exceptionally important in complex, high-pressure assignments, as many are. While not under direct analysis, these same service levels can be seen between the Service Partners’ management teams and Sida, be it UTV, UPPIS, or individual Programme Officers.

In addition, the type of service indicated by respondents (discursive, facilitative, adaptive, open and frank, etc.) is critical for engagement/utilisation. These contribute to a context in which issues may be discussed constructively and where differences can be settled. While the casual link between these service levels and engagement/utilisation are not clear, there surely is a one.

Interestingly, this does not seem due to a standardized approach to service but instead to the effective recruitment of teams who strive for these high standards and the genuine interest that Indevlop, in particular, has shown for the issues under evaluation. This intellectual commitment and curiosity to policy trends and issues in Sweden, to the importance of cross-cutting themes and issues writ small and writ large, all speak to exceptional levels of commitment.

3.3 THE MODEL (EVALUATIONS)

3.3.1 Functionality (procurement, partnership & quality)

3.3.1.1 *Is the current model of the Framework Agreement functional? (A3.1)*

The current procurement model is largely functional in that it has served to procure over 119 evaluation services and over 80 advisory services across Sida Departments and Swedish Embassies (See Sections 4.1.1 & 4.1.2). This is supported by analysis that demonstrates that this procurement model is sufficiently more efficient, in time or money, than previous procurement methods. At the same time, there is not enough evidence to demonstrate just how much more efficient procurement may be as compared to previous methods or best practices.⁴⁵

Functionality includes how Sida, its programme officers, UTV, and UPPIS, and others work with Service Providers. As intended, this goes beyond the commissioning of services and includes the development of a body of knowledge and data about the services that may serve many strategic purposes. As described throughout this and other sections, in this regard the FA is functional. Indvelop has provided detailed Annual Reports and an analytical synthesis of the first 71 evaluations commissioned under the FA. Sida has participated in Core Team Retreats and other events. Sida and the primary Service Provider are actively engaged in a partnership - whether deeply or broadly enough is difficult to ascertain. (See Section 4.2.3.)

FA functionality relates also to its capacity to support and promote quality services. As a model, its success hinges upon a sufficiently higher and consistent level of quality than could be assured from other procurement models. As described, the FA actually fails to ensure higher and more consistent levels of quality in and of itself—the improvement in the ToR, reporting and other improvements noted in this Review have more to do with the aptitude of the people involved and in the repeated engagement between Sida Commissioners and one Service Provider, the latter of which could occur within or without a framework agreement. Quality of the evaluation ser-

⁴⁵ This level of analysis would rest upon standard benchmarks for procurement transaction costs, time (in days) and indirect and direct costs throughout the process. Ideally, this would exist, as recommended in this Review, at Sida's level and then matched with growing standards within international cooperation and development. For instance, such a harmonised set of standards exists for International Financial Institutions. See "Procurement Harmonisation" at <http://web.worldbank.org/WBSITE/EXTERNAL/PROJECTS/PROCUREMENT/0,,contentMDK:22989220~menuPK:8118597~pagePK:8271521~piPK:8271523~theSitePK:84266,00.html>

vices under review, be it based on utilisation and engagement, on finely crafted, evidence rich reports, or on an accumulated appreciation of the value and importance of measuring and thinking about performance, especially in complex and somewhat oblique subjects, is inconsistent and not does achieve levels of acceptable quality in almost half of the evaluations under review. Thus, quality is mixed and cannot be attributed to the FA in any case.

While this is explored throughout this Review, this does limit the functionality of a procurement arrangement that ensures a steady flow of work to one Service Provider over five years. If this Service Provider is not readily equipped to perform evaluations at the highest quality standards or capable of improving its performance so that that a fair level of excellence can be assured in a relatively short time, the intended trade-offs in comparison to other procurement models become questionable. While we recognise this as a fairly high standard, we would expect Sida to aim for just such standards, especially given the commercial value of the contract and Sida's current and growing needs.

3.3.2 Comparative Analysis

3.3.2.1 *What are the advantages and disadvantages of this Framework Agreement when compared to other relevant procurement models? (A3.1.9)*

The Review includes a comparison of the FA with other relevant procurement models in the Annexes. This table provides a review of the distinguishing features of a number of procurement models. The table below presents some of the advantages and disadvantages of the FA as based on this Review's analysis and best practices as summarised by a recent report by Price Waterhouse Coopers (PWC) on public procurement in the European Union.⁴⁶

⁴⁶ For additional insights into the advantages and disadvantages of different public sector procurement models, please see: Price Waterhouse Coopers (PWC), "Public Procurement in Europe: Cost and Effectiveness." Prepared for the European Commission, March 2011. We have drawn from this excellent study throughout this review. For a more direct view on practices in Sweden, see: Karin Bryntse and Carsten Greve, "Competitive Contracting for Public Services: A Comparison of Policies and Implementation in Denmark and Sweden." *International Public Management Review*, Vol. 3, 2002.

Table 6: Procurement Model Comparisons		
	Advantages	Disadvantages
Sida FA (Ranked)	- Majority of the assignments done by one SP (1st ranked one) - Greater potential for partnership with 1st ranked SP - Greater level of accountability with one firm doing most work - Core team of the 1st SP expected to improve competence on client's needs and expectations - SP to have a strong QA role; may result in higher quality - Limits propensity to 'market' organisation in proposals - Procurement is quick/efficient	- No comparison of expertise/approach is possible - Uncertainty about whether the first ranked service provider has access to the best consultants for a particular evaluation - Market implications/quasi-monopoly with one SP and greater incumbency over time and in future framework agreements/open procurement - Without strict QA, may allow SP to do the minimum necessary for any engagement given that they are ensured future work under the contract - Difficult for the second or third SP to maintain a management team or a Core Team given volume and pipeline of work. - Difficult for individual consultants other organisations to gain work - The option to go to the second SP is hardly used. This is partly because it extends the procurement period for 2 more weeks with no guarantee the second SP will become the selected supplier—diminishes advantage of procurement time.
Multiple/Mini Competitions	- Assignments spread amongst pre-qualified SPs - Some potential for partnership but limited by number of pre-qualified SPs - All SPs have opportunity to improve competence on clients' needs and expectations; may occur more slowly than in ranked model - SPs may have a strong QA role; may result in higher quality - Procurement is quick/efficient but not as much as in ranked system	- Difficult for organisations not part of framework to gain work - Less direct accountability as compared to ranked model - May lead to proposals that are more ambitious and unrealistic than is feasible, thus creating a potentially longer inception phase - Slower procurement process, as compared to ranked system - Lower levels of partnership compared to a ranked system
Open Procurement	- Maximises competition - Maximises use/access to diverse competencies and resources in multiple local/international markets - Increases overall market dynamism	- Expensive/timely procurement process - No guaranty of quality - Less opportunity to nurture partnerships with select SP - More propensity for ambitious proposals

Other advantages and disadvantages are included in the aforementioned table in the Annexes.

As these indicate, most of the standard and related advantages and disadvantages of different procurement models have more to do with the access to the purchaser, the potential impact of competition amongst potential and pre-selected suppliers, and the overall costs and efficiencies associated with different approaches. There is a presumed advantage with the FA related to accountability and partnership that could support greater quality but the links between these are not guaranteed, nor fully demonstrated by this Review.

Accountability and partnership have more to do with the relationships, time, resources, and priorities that different parties commit to quality rather than to the terms and conditions embedded in a legal agreement. The same apply to quality assurance. A framework agreement can dictate that a QA process exists but it cannot dictate how and if that QA system is applied and how and if it is applied in ways that improve quality.

In consideration of the conclusions of the aforementioned PWC report and our own findings in this Review, most of the tangible advantages and disadvantages of any framework agreement concern the ease by which services are procured, the general satisfaction with pre-selected suppliers, and the clear cost efficiencies that are derived from different models. To expect the model itself to be a guarantor of quality is not supported by this comparative evidence nor this Review's analysis of quality overall.

3.3.2.2 *What would be the pros and cons of not having a separate framework agreement for evaluation and instead let evaluations be one of the services included in Sida's different thematic framework agreements? (A3.1.10)*

As this Review demonstrates, there is a significant opportunity to make evaluation expertise more prominent in Sida's decentralised evaluations as a result of the FA. The value of this evaluation expertise is associated with having better conclusions and better recommendations for Sida Commissioners and Partner Organisations. It should result in higher quality overall. This may not be the case in thematic or other frameworks.

Other "pros" and "cons" are addressed in Section 4.3.2.1 and in the table of "Comparative Procurement Models" in the Annexes.

3.3.3 Market Implications

3.3.3.1 *What possible incentives (for the consultant and for Sida) does the Framework Agreement give rise to? (A3.1.4)*

Competition and open markets involve the science of incentives. How markets are adapted and constrained relates to both broader, often societal, interests and to individual suppliers who have the 'market scope' to 'bend' the market toward their needs. In the case of this and any framework agreement, purchasers have an interest to re-

duce costs and to increase the overall efficiency through which services are procured. They have enough ‘market scope’, e.g. enough prospective purchasing power in any area, to convince suppliers that being part of a framework agreement will be in their long-term benefit even if it may imply more on-going costs than a strict one-off transaction.⁴⁷

For Sida as the purchaser, the most substantive and proven incentives relate to cost savings. Framework agreements are largely recognised as being a cost effective way to procure on-going services. A recent Price Waterhouse Coopers (PWC) study estimates that public procurement through frameworks in the European Union (EU) reduces person-day costs by 75%. At the same time, one needs to weigh this against the overall volume of contracts procured under a framework agreement. If only a few contracts are procured this potential savings shrinks in total value. If, as in the case of the FA, there are a relatively large number of contracts, then this level of costs savings is expected.⁴⁸

As confirmed through interviews with relevant Sida officers involved in the development and implementation of the FA, additional incentives exist. By including one primary and two secondary suppliers, they expect to have increased quality and service. The logic is that one supplier could provide opportunities for the development of expertise about Sida, decentralised evaluations, and overall Sida expectations and needs. It would learn more about Sida and thus be able to provide better quality services. Sida would also be able to hold a primary service provider accountable to certain methodological issues associated with evaluation that are geared toward increases in quality. This includes a focus on engagement/utilisation, a committed Core Team, an increased focus on methodology, and other factors that went into the FA’s design and implementation.

⁴⁷ This is related to a generation of research in economics regarding purchasing power parity and agency theory. This research is appropriate to the broadest treatment of market analysis but is somewhat dated when compared to actual behavioural patterns, e.g. how people act in different market settings. For an interesting treatise on how purchasing power and agency position interact, see Charles W. Hill and Thomas M. Jones, “Stakeholder-Agency Theory.” *Journal of Management Studies*, Vol. 29, May 2007. For the landmark study on how organizations, or ‘agents’ leverage these advantages, see Michael E. Porter, *Competitive Advantage*. Free Press, 1985. So, while this section treats the issues associated with ‘incentives’ subsequent sections shift toward how different stakeholders may be behaving and/or reacting to these incentives. This latter analysis is based on the research from behavioural economics, amongst other areas.

⁴⁸ Price Waterhouse Coopers (PWC), “Public Procurement in Europe: Cost and Effectiveness.” Prepared for the European Commission, March 2011. Page 90. In fact, this study estimates that if framework agreements are only twice the amount of the average contract value, e.g. two contracts called off, framework agreements have the highest process costs when compared to other procurement models, including open, one-off procurements.

As shown in Section 4.3 in relation to evaluations, there has been some quality improvement but this has not increased the quality in key areas as compared to before the FA and as compared to best practices in evaluation. There are, as noted in Table 6, some inherent and realised advantages associated with this and other framework agreements. As this Review's analysis and cited research indicate, there is no clear correlation between the FA, per se, and these quality improvements. These tend to be based more on the commitment of the Service Providers and the early insistence on a few key areas by UTV. These may be based on the on-going relationship with one Service Provider although the same could happen with or without a framework agreement. What is more telling is that there haven't been greater improvements in quality, especially in critical areas associated with evaluation (data collection & analysis; evidence based findings and conclusions). Being the first ranked on the FA does not seem to have been enough of an incentive to make significant increases in quality in such critical areas.

The other incentive that Sida attaches to the FA concerns the need for a strategic view of what evaluations are occurring across Sida and in Swedish embassies. The FA, it is thought, could provide opportunities for institutional learning and a greater base of knowledge to facilitate Sida's development activities overall. In this, Sida has benefited from Indevalop's work. They provide detailed Annual Reports and have produced a synthesis of 71 evaluations.⁴⁹ This does not include the work being done by SIPU and Orgut, or roughly 14% of the evaluations assigned to date. At the same time, it provides substantive and strategic intelligence. What remains unclear is if this meets Sida's needs and/or if Sida, and UTV in particular, have the resources to act upon this intelligence. (See Section 4.1.3.)

For the Service Partners, market dynamics suggest that the primary incentive is commercial. The FA provides a steady flow of work over time. This provides both resources and a level of stability that supports a business's development. Even for those second and third ranked, the FA provides work and, as repeated by some SP Managers, it raises their profile with Sida and in the market. (See "Data & Analysis Report," page 313.)

There are other potential incentives for the Service Providers. They have an opportunity to gain experience in development evaluations and advisory services across Sida, increasing their skill in these areas, gaining particular experience with Sida standards and expectations. They also may have an interest in the intellectual chal-

⁴⁹ Ian Christoplos, Anna Liljelund Hedqvist, and Jessica Rothman, "Swedish Development Cooperation in Transition: Lessons and Reflections from 71 Sida Decentralized Evaluations (April 2011 – April 2013)."

allenges associated with such work and how this provides individual satisfaction while better positioning themselves for related work in the future.

There are indications that Indevelop has not had an overtly averse reaction to some of these incentives and inherent risks. It has been proactive in developing reports and other outputs that raise issues in how evaluations are being conducted and issue that relate to Sida's decentralised evaluations. The "Swedish Development Cooperation in Transition" is an example of this.⁵⁰ This indicates that Indevelop is committed to the intellectual engagement that their work provides and to supporting learning as possible. Of course, this does not imply that the other Service Providers do not have a similar commitment to intellectual engagement and learning. They surely do. Nor, does this indicate that by having one primary Service Provider there is more incentive or opportunity to engage intellectually and to support learning. In fact, common sense would indicate that there would be more opportunities and greater diversity of engagement in a more competitive environment.

All of these incentives are possible in relation to the FA. However, this Review suggests that stakeholders have responded differently to these and other more obscure interests. For instance, 'Sida's interests' are not enshrined as such; they depend on the people involved, be they in UTV or UPPIS or at a Programme desk or embassy and with the internal dynamics, related to priorities, resources, and leadership, that may highlight these interests with greater or lesser intensity. Given the changes in UTV and UPPIS, both in priorities and staff, one would expect shifts and changes in common incentives.

This is not intended to be vague. The primary incentives are clear: cost savings, increased quality, and institutional and strategic learning and knowledge for Sida; commercial, intellectual, and market interests for Service Providers. What is at issue is that these only prompt actions when given the prioritisation and value placed on them by engaged stakeholders. The FA in and of itself is no guaranty that these 'interests' will manifest in results. It is, as named, simply a framework.

⁵⁰ Ian Christoplos, Anna Liljelund Hedqvist, and Jessica Rothman, "Swedish Development Cooperation in Transition: Lessons and Reflections from 71 Sida Decentralized Evaluations (April 2011 – April 2013)."

3.3.3.2 What does the Framework Agreement imply for the long-term availability and quality of evaluation services to Sida, taking the resource base as a whole into consideration? (A3.1.5)

This question relates directly to the FA's market implications. This was a contentious issue among respondents in this Review. It was the most often cited negative aspect of the FA:⁵¹

GRAPH 25: What does the Framework Agreement imply for the long-term availability and quality of evaluation services to Sida, taking the resource base as a whole into consideration? (A3.1.5; combined cohorts: Sida commissioners, SP management, coalition partners, core team, and consultants.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68											

While a range of issues emerge from this evidence, we have combined those that are most relevant with those that are dictated by leading theories about procurement, competition, and market analysis.

Unfortunately, we were not able to find reliable studies and research on the Swedish consultancy market to use as supporting evidence. Studies exist, of course, but they treat either procurement at a government level that is not usefully compared to Sida or are geared toward consultancy competencies and subjects that are not easily relatable to those under this Review, or are simply out of date, e.g. prior to 2000.⁵² Nonetheless, several factors are worth consideration:

Competition: This is a primary element of the FA's design. By having a ranked system where the first ranked assumes over 80% of the evaluation assignments and nearly all of the advisory service assignments imply that the award of individual assignments is skewed toward Indevelop. Indevelop, of course, was selected as part of a

⁵¹ Interestingly most of these are the same issues found in a broader assessment of procurement practices across the public sector in the EU. See, Price Waterhouse Coopers (PWC), "Public Procurement in Europe: Cost and Effectiveness." Prepared for the European Commission, March 2011. Page 115.

⁵² Our research is hardly meant to be conclusive but it was what was feasible for this Review's purpose. For some of the common citations used in this Review in relation to market dynamics, see the footnotes in Section 4.3.3.1. An interesting article but that still suffers from most of these noted constraints is: Lilach Nachum, "Measurement of Productivity in Professional Services: An Illustration on Swedish Management Consulting Firms." *International Journal of Operations & Production Management*, Vol. 19, No. 9, 1999.; and Flemming Poulfelt & Adrian Paynee, "Management Consultants: Client and Consultant Perspectives." *Scandinavian Journal of Management*, Vol.10. No. 4, December 1994.

competitive process. Yet, the overall procurement scores and other evidence suggest that it was not so far removed from the second and third ranked service providers.

Furthermore, the subsequent lack of competition is not in and of itself a harbinger of quality. The evidence concerning quality suggests that the first ranked Service Provider has not provided demonstrably higher quality than before or in comparison to best practices and leading research. Thus, the lack of competition in this regard does not seem to have a positive impact on quality either way.

Incentives: As noted in section 4.3.3.1 above, leading research suggests that there are distinct incentives under a FA of this type for Service Providers and Sida. These are primarily based on rational economics theory—expecting that people and organisations act in their self-interests and in rational and logical ways. Behavioural economics demonstrate that people are often illogical in how they respond to different incentives and that organisation may be even more so.⁵³ Given this, this Review would require a much more considered view to understand how Sida and Service Partners have reacted to possible incentives.

Certainly, Indvelop has maximised its position as exhibited by the breadth of assignments it assumes. It has acted to maximise its commercial position. Evidence from Indvelop’s management team indicate that they worked to submit proposals for every call-off under the FA, regardless of whether they have demonstrated comparative advantages as compared to SIPU and Orgut. While a review of the written rationales for passing on the few assignments is not detailed enough to form a fully informed finding, those that were passed are relatively difficult, in difficult countries, and with complicated ToRs. While there has been an increase of proposals passed on in the last six months, Indvelop’s management team states that this was due to their perceptions of this Mid Term Review and its potential results.

Given this assumed protection of the commercial interests embedded in this FA, it may be expected that the first ranked service provider takes fewer chances or is less innovative than those with ‘less to lose.’ Again, there is no clear causal link but the assessment of the evaluation reports done as part of this Review reveals very few innovative practices or approaches. It is not clear, however, if this is because of market risk, a lack of proficiency with new and emerging evaluation tools and approaches, or

⁵³ For a landmark work in this regard, see Daniel Kahneman & Amos Tversky, *Choices, Values and Frames*. Cambridge University Press, 2000. This and related work won its authors the Nobel Memorial Prize in Economics in 2002. For compelling research on the absolute irrationality of individual choices, see: Dan Ariely, *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. HarperCollins, 2008.

other more obscure reasons. While Indevelop is certainly interested in protecting its commercial position, there is not enough evidence to suggest that they have become more conservative and/or risk averse. Instead, it seems that they have simply done the best they could with the team and resources available.

Competition Over Time/ Rising Benefit to Incumbents: This FA runs the risk of reducing competition over time. It gives rise to a growing level of experience in the first ranked Service Provider that may put them in a dominant position in future procurements. This is not guaranteed but it is an important consideration.

Proposal/Team Evaluation: The most commonly cited feature of the FA by Sida Commissioners concerns the lack of an opportunity to compare teams and/or approaches between different Service Providers. (See “Data & Analysis Report,” pages 312 – 314.) This seems particularly relevant given the diversity of programmes under evaluation and the different needs and available resources for these evaluations.

Individual Consultant Access: Individual consultants are generally required to go through the first ranked service provider if they want to work on development evaluations for Sida. This gives Indevelop exceptional market dominance in recruitment, selection and remuneration. Indevelop could decrease the rates paid to consultants. Indevelop could have requirements or ways of working that are not conducive to some consultants’ backgrounds or skills. Or, they could have selected Core Team members based on who was available at the time, inadvertently blocking consultants who have been on long-term assignments out of Sweden or on other development activities, a common situation given the nature of the field.

Young Consultant Access: As noted in Section 4.3.3.4, the FA implies that nearly all young evaluation consultants must position themselves with one Service Provider to gain work/experience.

Evaluation Community in Sweden: Respondents commented on how the FA has impacted the Swedish intellectual community. In essence, the development, discussions, and forums associated with these Sida evaluations are all within the purview of Indevelop. This may not only block certain local actors from participating in these forums and discussions, from accessing what is going on, the trends in Sida development evaluations, but also implies that one commercial actor determines what is and isn’t discussed.

Development of Evaluation Talent: Finally, this FA, with its prominence in development evaluations being conducted by Sida and by Swedish embassies, means that the primary talent development of evaluators is being done by one organisation. “Are Sida Evaluations Good Enough?” notes that one of the factors that lent to poor evaluation quality was “Insufficient professional development in the field of evaluation in Sweden—few courses and other training opportunities.”⁵⁴ The FA implies that most of the learning about Sida development evaluations is being done by one organisation.

All of these and other aspects of this Review indicate that the FA may have a negative impact on the long-term availability and quality of evaluation services to Sida. The dominance of one Service Provider, given the diversity of evaluation needs, the quickly developing standards associated with methodologies, tools, and approaches, the plethora of innovations and new tools that emerge internationally, and the access of different evaluators to development evaluations at Sida, implies an unfortunate constriction on what is a dynamic and quickly changing field and market.

This does not imply that the other Service Providers in this FA and beyond are not developing useful methodologies and approaches and developing the existing and emerging talent nationally and internationally. They remain active in evaluations and other work—just not in this essential part of evaluations at Sida. This FA does not suitably leverage their experiences nor does it provide an avenue for the talent that these suppliers are developing. Sida is not engaged with them and others while being wholly engaged with one Service Provider.

Recommendation 7 (UTV): The ranked system wherein the majority of the work goes to the first ranked supplier is neither a guarantor of increased quality or consistency. A system of “mini-competitions” between pre-selected suppliers is more appropriate. This will enable Sida Commissioners to have more informed decision-making about teams and approaches amongst a broader range of consultants. This will ward against the increased encumbrance of one supplier. It will increase the knowledge about Sida’s decentralised evaluations within and beyond Sweden and will re-engage the intellectual community in Sweden that may not have had access to the development in these Sida Evaluations.

⁵⁴ “Are Sida Evaluations Good Enough? An Assessment of 34 Evaluation Reports,” Kim Forss, Evert Vedung, Stein Erik Kruse, Agnes Mwaiselage, and Anna Nilsson. Sida, 2008. Page 79.

3.3.3.3 *What are the implications of the Framework Agreement for local evaluation consultants in partner countries? Could the Framework Agreement be improved to better contribute to enhanced development of capacities of evaluators in partner countries? (A3.1.6)*

Unfortunately, this Review did not discover enough evidence to answer this question fully. Obviously, there is an interest in recruiting local evaluation consultants not only for their expertise but to promote effective evaluation amongst partner countries. Service Providers have been diligent in seeking out and using local consultants.⁵⁵

Indevelop's most recent Annual Report states that 33% of evaluation teams are composed of national, e.g. 'local', evaluators.⁵⁶ It is not clear if this is by necessity, e.g. requirements or benefits from having local consultants, or as a concerted effort to increase the evaluation capacity of these consultants. Given Indevelop's commitment, as they state in the same Annual Report, "to support the development of methodology, cross-learning, and experience between evaluations" one may assume that that knowledge and learning is transferred from Core Team Members to other team members. This should support the development of this local capacity.

It is worth noting also that a range of Swedish embassies have used the FA even though this is not mandated for them. (See section 4.1.1.) They have chosen to use the FA rather than to procure locally or to use other consultants. As demonstrated, this is related to the ease associated with procuring services under the FA. It is not clear if this has implications for local consultants. Does the increased use of the FA imply that fewer local consultants are being used by these same embassies? This remains unclear.

3.3.3.4 *What has the Framework Agreement implied for young evaluators to be engaged in the provision of evaluation services to Sida? (A3.1.7)*

The first ranked supplier has made a concerted effort to include 'young evaluators' on evaluation teams. As stated in their most recent Annual Report, 29 "junior professionals have been engaged in evaluations up to date."⁵⁷ Of those included in this Report, this implies that approximately 12% of evaluators used by Indevelop are 'junior professionals'. Indevelop's proposal template includes a heading "opportunities to

⁵⁵ Indevelop has documented and reported on team composition well. Please refer to their evaluation and advisory service registries for a complete list of evaluation teams.

⁵⁶ "Final Annual Report 2012 – 2013," Indevelop. June 2013. Page 25.

⁵⁷ IBID.

engage junior consultants” that suggests that they actively consider opportunities for young professionals in every assignment.

It is beyond this Review’s remit to judge if this is sufficient. It is clear that Indevelop recruits and selects these young professionals, using their criteria and hiring practices. They have become the primary arbitrator of which young professionals are qualified for these Sida evaluations.

This implies also that there may be many other young evaluators who may not fit Indevelop’s criteria and hiring practices but that may be highly qualified otherwise. These consultants have no other access to these assignments and so must seek experience in other areas, by subject and locality.

3.3.4 Knowledge Management & Institutional Learning

3.3.4.1 *Has the Framework Agreement changed the commissioning of evaluation services, including better-formulated ToR, and if so, to what extent and in what ways? (A3.1.3)*

The commissioning of evaluation services is quicker under the FA. This is confirmed by a majority of Sida Commissioners when speaking to the quality of the FA. (See “Data & Analysis Report”, pages 249 -252.) They regularly cite the ease and speed by which services are procured.

At the same time, some complain that while the actual procurement is quicker, the Inception Phase is more detailed and longer. This is related to increased engagement/utilisation. While some Sida Commissioners may complain about this, it does not represent an area of concern. (See Section 4.2.1.)

As noted in Section 4.2.2.1, there has been a noticeable increase in the quality of ToRs. It is difficult to develop a causal chain between this improvement and the FA. Instead, this seems related to a concerted effort to facilitate Sida Commissioners’ use of the previous and new ToR template and in the repeated work with the Service Providers who emphasise the need for well-crafted ToRs.

Beyond these factors, there are no other discernable patterns that would suggest increased use of standard tools, processes or other mechanisms that may be more widespread as part of the FA.

3.3.4.2 *To what extent does the Framework Agreement contribute to Sida’s institutional learning on evaluation? (A3.1.8)*

There has not been a significant amount of institutional learning about evaluation within Sida as a result of the FA. The only area where there are indications of in-

creased learning regard the changing quality of ToRs and yet the causal links for this are unclear.

The FA, by using one Service Provider for most work, has created a bastion of knowledge and information from which Sida may draw. Yet, even this does not include the nearly 20% of evaluations that have been done by the other Service Providers. In essence, Sida has ‘outsourced’ institutional learning to the first ranked Service Partner, even if such institutional learning existed before the FA. Indevelop has delivered detailed and useful work in this regard. They have comprehensive registers of the work done, who did it, the costs, and other relevant statistics. This is an improvement given that UTV kept few statistics before the FA and these were unreliable given reporting standards. They present this in Annual Reports and other materials. They have gone beyond this in developing a synthesis of 71 reports that provides interesting insights into the trends they detect across their work.⁵⁸ This and other facets of Indevelop’s work represent a significant level of institutional learning on their part.

This ‘outsourcing’ of knowledge implies that the learning is taking place with the Service Provider rather than with Sida directly. This raises issues of credibility and biases although there has been no direct evidence of such in this Review. Instead, this presents a missed opportunity for Sida. Given that little institutional learning took place before the FA and given that this is now being done by one third party, if that third party stops delivering service, any institutional knowledge may be lost. If the FA is not continued, or continued in another form, and for whatever reason Indevelop does not play the same role, then this institutional knowledge may become quickly out-dated.

Currently, Sida has access to this information. It may draw upon it as needed.

Recommendation 8 (UTV): UTV has an opportunity to have a much more informed and broader view of the evaluations being undertaken as a result of the FA. At the moment, most of the information and institutional knowledge that serves this purpose resides with Indevelop. UTV could review what information, statistics, or materials are important for on-going institutional learning about development evaluations in collaboration with Indevelop. UTV could then standardise this for any future framework agreements and/or for other Service Providers.

⁵⁸ Ian Christoplos, Anna Liljelund Hedqvist, and Jessica Rothman, “Swedish Development Cooperation in Transition: Lessons and Reflections from 71 Sida Decentralized Evaluations (April 2011 – April 2013).”

3.3.4.3 How has Sida organised itself to support the knowledge management and institutional learning that should arise? (A1.1.9)

This is answered in Section 4.4.4.

In brief, there are few formal ways that Sida and/or UTV have organised to support the knowledge management and institutional learning that could be achieved through the FA. They have made an effort to disseminate information and reports from the Service Providers, as evidenced through qualitative evidence from this Review and in the literature desks within Sida. However, most of the tangible knowledge resides with the first ranked Service Provider.

3.3.5 Costs & Efficiencies

3.3.5.1 What have been the costs of the evaluations in the Framework Agreement? (A1.1.5)

Evaluation costs are represented in the following:

Table 7: Evaluation Costs ⁵⁹	Indevelop (82 Assignments)		SIPU (19 Assignments)		Orgut (3 Assignments)		TOTALS
	Total	Average Per	Total	Average Per	Total	Average Per	
Evaluations	19,322,030	483,051	4,744,352	431,305	600,000	600,000	24,666,382
Reviews	12,679,969	487,691	1,379,286	344,822	581,532	290,766	14,640,787
MT Re-views	5,380,174	489,107	1,852,738	617,579			7,232,912
Total Averages:	455,880		419,809		393,844		423,178
Other	5,490,290	549,029	1,198,720	1,198,720			6,689,010
TOTALS:	42,872,463		9,175,096		1,181,532		53,229,091

This shows a total cost of 53,229,091 SEK for 104 assignments across the three Service Partners, or an average cost of 516,787. Evaluations actually constitute the lowest average cost at 483,051. This is far below the maximum of 1,200,000 for the FA, as based on relevant procurement rules and other factors. Only five assignments exceed 1,000,000. While not conclusive, these amounts are based on what is budgeted for and negotiated with Sida Commissioners.

⁵⁹ This includes figures for invoiced amounts as provided by the Service Providers. As such, it does not include figures for on-going assignments. Some assignments did not have figures available. Please See "Evaluation Assignment Registry" in the Annexes.

While evidence is mixed, one of the more common quality issues raised across cohorts is the time, and as related, budget for evaluation assignments. There are few relative benchmarks or standards for how much development evaluations should cost, in total or as a percentage of programme budget, or by any other reasonable measure, including within Sida. The World Bank estimates the costs for impact evaluations between US\$250,000 and US1 million, with an average cost of sampled evaluations at US\$526,198.⁶⁰ Yet, these have requirements that largely go beyond those of the evaluations under review.

The other aspect that arises from this Review's analysis and qualitative evidence is that inception phases are long and detailed as part of engagement/utilisation. Participatory approaches are time intensive, they take more meetings, more interviews, and more analysis of the outcomes from this level of engagement. This may imply greater cost but also greater benefit when accounted for appropriately.

Finally, Sida recognises that a cap for these evaluations at 1,200,000 SEK, be it because of procurement standards or for recognition of how much such evaluations may cost. Given that the average cost is 516,787, most evaluations are far below this level. While the quality issues noted in this Review cannot all be attributed to lack of resources and budget, evidence does indicate this as an issue. Some diligence and guidance can be provided to Sida Commissioners when considering the balance between needs and budgets in evaluations.

Recommendation 9 (UTV): Develop guidance for Sida Commissioners about how much evaluations should cost as based on scope, requirements, and proposed methodologies. This can be based on the experience of the services procured under the FA and a relevant benchmark for activities going forward. A useful approach to this is minimum and maximums (min-max) for any requirement/approach. Another potential reference is to think in terms of a percentage of total programme investment. While any such guidelines may seem somewhat arbitrary at first, they will provide a foundation for more proficient standards in the long term. It should be noted that the current ranked framework agreement limits the room for negotiation of terms and costs but this may be rectified if Sida moves to mini-competitions for any future framework agreements.

⁶⁰ "World Bank Group Impact Evaluations: Relevance and Effectiveness," Independent Evaluation Group, June 2012. Page 151.

3.3.5.2 *How have the costs for evaluation services changed under the Framework Agreement? (A3.1.1)*

There is no evidence to suggest a significant change in the costs of the evaluations under the Framework Agreement comparatively, before the FA, or among those conducted under the FA. There are potential cost savings associated with a shorter procurement cycle but these may be off-set by longer induction phases. (See Section 4.2.1) The indications of greater consultant competence (Section 4.2.3.2) may lend to greater efficiency although there is no direct indication of this.

A review of the invoiced amounts per evaluation assignment do not signal any change, positively or negatively, in the costs for evaluation services. They represent what was agreed upon rather than any distinct change.

3.3.5.3 *How has the Framework Agreement impacted transaction costs/efficiency overall? (A3.1.2)*

An accurate determination of transaction costs and increased efficiency would require a benchmark. As noted in the Inception Phase, no such benchmark exists. Nonetheless, qualitative evidence and other sources indicate that this and any FA decreases transaction costs. This is confirmed in a study on procurement practices in public sector organisations across the EU:

Framework agreements aggregate in one initial stage a large part of the administrative burden of a procurement process. This would be expected to lead to significant cost savings, particularly if the number of subsequent contracts within the framework is high. The study confirms this intuition. Framework contracts have lower costs than any other form of procurement and the savings are found both for authorities and for firms. They also perform better than other forms of procurement in terms of the number of bids that they attract. This, however, should be balanced against the much lower level of competition that will exist at the time of each contract under the framework.⁶¹

In addition, this Review notes that there are efficiency gains in the procurement process and potential savings at other parts of the process. For instance, greater knowledge of Sida systems and practices may imply less time in re-drafting reports. However, this Review has not discovered nay evidence for this.

⁶¹ Price Waterhouse Coopers (PWC), "Public Procurement in Europe: Cost and Effectiveness." Prepared for the European Commission, March 2011.

3.3.5.4 *Are Sida units positioned (legal, finance and administration; business systems/technology) to ensure that expected efficiency gains/cost savings are realised? (A1.1.12)*

As noted in Section 4.1.3.2, Sida's organisation matches best practices and no issues arise about this in relation to how the FA is managed.

As noted in section 4.3.5.1, effective management of costs and efficiencies requires a concerted investment in analysis and business systems that can establish reasonable benchmarks and measure on-going work against these. There are indications that UTV is now making such investments and adopting a formal process for determining costs and efficiencies. This should prove beneficial. Not only does this provide useful insights into how much evaluations of different types do and should cost but also how cost can be an indicator of quality.

3.4 ADVISORY SERVICES ON RESULTS FRAMEWORKS (OPERATIONALISATION)

3.4.1 Overview

3.4.1.1 *How has the Framework Agreement been put to practice and used in relation to the Results Framework component? (B1.1)*

The FA has enabled Sida Commissioners to recommend these services with the comfort that there is a qualified service provider available and that the procurement process will be relatively simple. This combination of ease and confidence results in a significant number of Advisory Service assignments. (See Section 4.5.1.1.)

Given this, no significant issues emerged in this Review to suggest that the FA has not been used adequately. If anything, there may be an even greater need going forward. Given this, a framework agreement for this type of service meets a definitive need.

3.4.2 Number, Type & Costs of Advisory Services

While the following questions are duly answered as part of this Review, the information is taken from Indevlop's annual progress and financial reports. This Review discovered no reason to doubt this.

3.4.2.1 *How many advisory assignments have been conducted within the Framework Agreement and how many have they turned down? (B1.1.1)*

98 advisory services have been assigned as of June 2013. 97 of these have been conducted by Indevelop and 1 by Orgut. There is not sufficient evidence about whether this meets all potential demand. This depends on how well Programme Officers and Embassies know about these services and believe that they will be of benefit to partner organisations. (See Section 4.1.4.3.)

3.4.2.2 *What have the advisory services consisted in? (B1.1.2)*

As per documentary evidence and qualitative evidence from Partner Organisations and Sida Commissioners, these services have included capacity building, facilitative/participatory workshops, and the development of common standards, models and templates that can be used for results based reporting to Sida and other donors. (See “Data & Analysis Report”, pages 314 -320.)

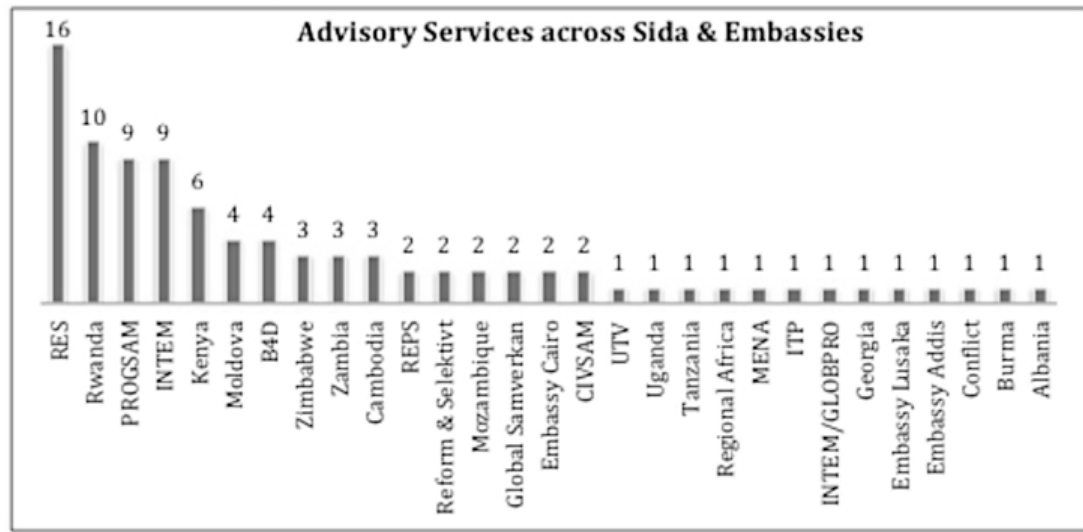
Partner Organisations confirm that this combination of services, from direct capacity building to providing templates for Sida reporting, not only met their immediate needs but also often supported their appreciation and utilisation of results based management techniques. As noted below, evidence indicates that Indevelop has an appropriate balance of standardised tools and facilitative/ participatory approaches that cater to diverse partner organisation needs and contexts.

3.4.2.3 *To what extent has the Framework Agreement been used by different Sida units/Embassies? (B1.1.3)*

As with evaluation services, Advisory Services have been conducted across Sida Departments and embassies.⁶² As Graph 26 illustrates, the Embassies in Rwanda and Kenya and RES, PROGSAM, and INTEM are the most frequent commissioning offices of these services. However, there is a fairly broad spread across different Sida units and Embassies.

⁶² Indevelop provides a list of all advisory services in “Annex 4: Assignment Registry—Advisory Services,” Indevelop’s Annual Report 2012 – 2013, June 2013. Page 77.

Graph 26: Use of Advisory Services by different Sida Unites and Swedish Embassies



3.4.2.4 Duration and costs of services (per assignment and total)? (B1.1.5)

This is provided for in Indevelop's annual progress and financial reports. This does not include the last 17 assignments. This can be summarised as follows:

Table 8: Summary of Advisory Services' Costs	
Total Invoiced	2,239,987 SEK ⁶³
Average Cost (All)	28,000 SEK
Average Cost (2 days)	18,400 SEK
Average Cost (+2 days; max 10)	67,400 SEK

45 of these assignments have been less than 2 days, leaving 35 that have been between 3 – 10 days. The value and results that have been achieved justifies an extension of the contracts beyond 2-days and perhaps closer to the 10-day maximum. As many Partner Organisations commented, the service was valuable and appreciated but often too short. (See "Data & Analysis Report," pages 25 & 26.)

The initial 2-day assignments were financed by UTV originally. This meant that it became easier for Sida Commissioners to procure these services as they did not have to prepare a formal decision for allocation of funds. An assessment of this was made

⁶³ This is slightly higher than that reported in Indevelop's annual financial report. It reports SEK 2,207,987. Our figure is based on an individual invoice amounts.

end of 2012. It was found that this arrangement was not good, as some officers tended to over-use the service (as it was “free”) and there was a tendency to keep assignments to 2 days even though more would have been appropriate, especially when face-to-face meetings would have the most impact. Of course, this also put more of an administrative burden on UTV. Therefore, UTV stopped “offering” 2-day services while still communicating that these services existed.

Overall, this does not seem to have had a negative impact on the number of days commissioned for Advisory Services as there is a relatively consistent mix of terms per assignment across all related services.

3.4.3 Consultant

3.4.3.1 *How have the consultants organized themselves to respond to specific advisory service engagements, including providing qualified evaluation teams, Core Team and quality assurance? (B1.1.6)*

Given that most of these assignments do not include substantive material outputs or other data, there is no clear way to weigh material evidence against qualitative and other evidence. This limits the evidentiary base to adequately answer this question. However, the overwhelming evidence regarding the effectiveness and impact of these services, as reported by Sida Commissioners and Partner Organisations, implies that consultants have organised themselves well.

Some respondents state concern about the fact that some consultants providing advisory services are also conducting evaluations, especially when Indevelop has recommended these services. While conflict of interests should be of distinct concern, this Review did not discover evidence of malfeasance or irregularities.

3.4.3.2 *Are consultant competencies (experience, expertise, aptitude, attitude) appropriate and adequate for specific advisory services? (B1.1.7)*

As above, a thorough analysis of each consultant’s specific skills and how they match specific organisational needs is not part of this Review’s remit. Instead, we review general qualifications, Partner Organisations’ satisfaction with their qualifications, and the overall impact they have had on results frameworks and RBM practices. This is consistent with this review’s contribution analysis approach. Section 4.7 on “Quality” does indicate that consultant competencies are largely adequate for Advisory Services.

3.4.4 Sida Organisation

3.4.4.1 *How does the dialogue and consultation between the service providers and UTV/UPPIS/Sida affect a common understanding of how to use/apply the Framework Agreement for advisory services? (B1.1.4)*

While the evidence is inconclusive, there are indications that there is certain lack of harmony in how Sida is developing its approach to performance and results, including RBM and similar frameworks, and how Indevelop is working with Partner Organisations. This is based on a few comments from key respondents and on what Partner Organisations state about the quality of the services overall. Sometimes consultants work with Partner Organisations to develop specific models and/or templates while at other times they provide coaching on how to think about results frameworks. These can be largely different from existing and emerging Sida requirements. As noted, Sida's policy is not to require any specific format. Any format should comply with the "basics" of RBM, such as clear hierarchy of objectives, good indicators, assumptions made explicit, etc.

This seems appropriate. As intended, Advisory Services are a 'service' to Partner Organisations. They are not mandated nor is there an overt insistence on their use. Instead, they represent a need and Sida's capacity to meet this need. The fact that there is not perfect harmonisation between what the Advisory Services may or may not contain and Sida standards is not of utmost importance.

3.4.4.2 *Have Sida's structures for managing the Framework Agreement for Advisory Services been appropriate and adequate, including quality assurance, learning, legal matters, support and communication? (B1.1.8)*

No issues of substantive difference were found for Advisory Services as compared to evaluation services. As such, please refer to Sections 4.3.1 and 4.3.3.

3.5 ADVISORY SERVICES ON RESULTS FRAMEWORKS (QUALITY)

As stated in the previous sections, while there is no direct evidence between consultant competencies, how they organise themselves, and how Sida standards and a QA processes may or may not be applied, actual results are positive. More than any other area of this Review, qualitative evidence demonstrates a conclusive positive trend:⁶⁴

GRAPH 27: Advisory Services Quality (ALL COHORTS) (B2)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139
140	141	142	143	144	145	146	147												

GRAPH 28: Advisory Services Quality (Partner Organisations) (B2)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
100																			

Common trends and statements in this qualitative evidence include:

- Partner Organisation respondents state that the consultants were facilitative and flexible in their approach to the organisation, their general knowledge of RBM and how they adapted to the organisational context.
- Partner Organisations state that the knowledge of RBM and related frameworks is valued more than before the services and by more relevant people across their organisations.

⁶⁴ For reading this and similar graphs, please refer to “Using Qualitative Graphs” in the *Annexes*.

- Partner Organisations state that the models and templates that the consultants shared were useful. Many have adopted these templates for standard business practices.
- Some Partner Organisations state that this service prompted them to invest in additional support and work with results frameworks, including hiring the related consultants individually.
- ‘Neutral comments’ from Partner Organisations often concern the lack of an adequate knowledge of the organisation that meant that some of the limited time was spent discussing the organisation and its work rather than results frameworks.
- Common ‘negative’ comments concern the time for the assignment and how, when combined with a lack of knowledge about the organisation, meant that sometimes there wasn’t adequate time for actual work on results frameworks.

The quantitative survey and a review of the documentation and reports related to these services support this positive trend. This evidence demonstrates that the Advisory Services are not only meeting a growing need but meeting it well.

3.5.1 Relevance

The capacity to demonstrate results and to report on these with a fair level of coherence is growing in importance. A prominent milestone in the shift toward effectiveness and demonstrable results is the Fourth High Level Forum on Aid Effectiveness in Busan, South Korea, 2011.⁶⁵ As discussed and agreed upon there, Donors require demonstrable increases in relevance, effectiveness, impact, efficiency, and sustainability. This implies a greater role for both evaluation and the improvement of how organisations demonstrate results. This is not simply to better explain these results to different constituencies but to use solid evidence to establish what works, what doesn’t, and to use this to improve performance. This is leading to approaches and tools that will radically alter the way we understand development cooperation.⁶⁶

⁶⁵ For an interesting perspective on this, see: Emma Mawdsley, Laura Savage, and Sung-Mi Kim, “A ‘post-aid world’? Paradigm shift in foreign aid and development cooperation at the 2011 Busan High Level Forum.” *The Geographical Journal*, January 2013.

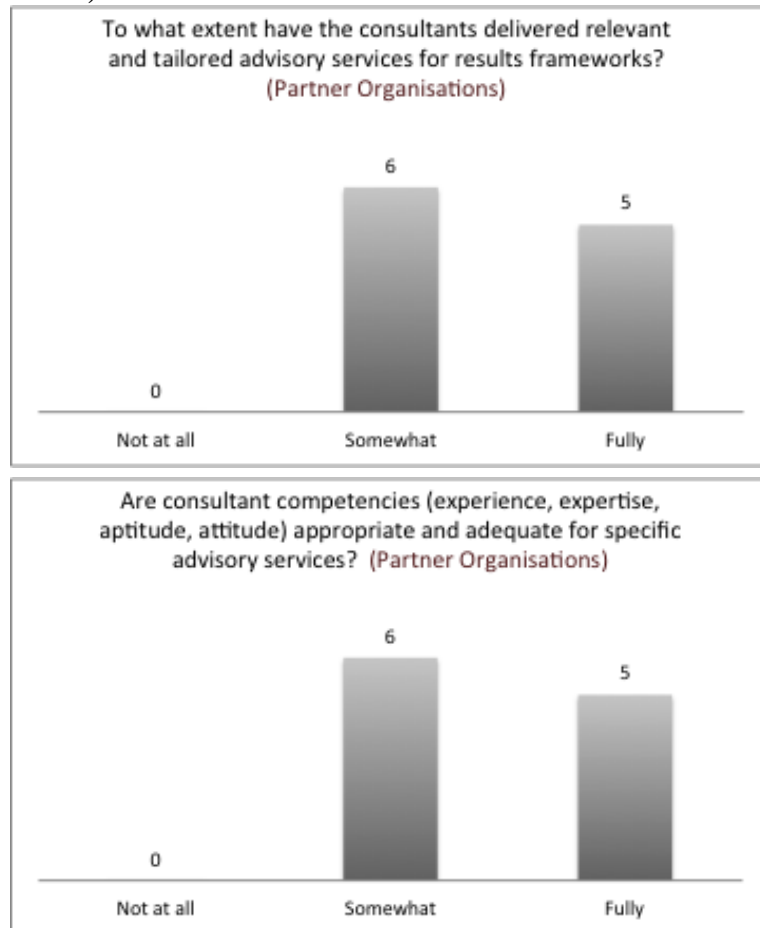
⁶⁶ For a fascinating and compelling approach that includes evidence-based randomized control trials, see: Abhijit V. Banerjee and Esther Duflo, *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. Public Affairs, 2011.

3.5.1.1 To what extent have the consultants delivered relevant and tailored advisory services for results frameworks? (B2.1.1)

While evidence is inconclusive, overall quality reported by respondents and the limited evidence directly related to this question suggest that these services have been adequately relevant and tailored to specific partner organisations. Graphs 29 & 30 demonstrate that no respondent found these services irrelevant or the consultants' competencies inappropriate or inadequate. 45% of respondents state that the services were "fully" relevant and consultant competencies "fully" appropriate and adequate.

This is confirmed in an analysis of the qualitative evidence regarding overall quality. As the graph in Section 4.6 demonstrates (Graph 30), over 70% of

GRAPHS 29 & 30: Quantitative Survey (Partner Organisations: Advisory Services)



partner organisations had positive responses regarding quality. Common themes emerge upon further analysis. Consultants took the time to understand organisations' needs and worked as 'facilitators' rather than being overly directive or pedantic. Consultants expressed a genuine interest in organisations, how they worked, and their different levels of competency and comfort with results frameworks.

Given this, consultants work from where organisations are, helping them develop rather than setting unrealistic expectations. Statements from respondents typify this sense of development and accomplishment and in that they continue to develop results framework skills, sometimes hiring consultants directly to facilitate their progress. (See "Data & Analysis Report," pages 320-324.)

3.5.2 Effectiveness & Impact

3.5.2.1 *Has the quality of results frameworks and partner competence on results frameworks and RBM improved with the Framework Agreement? (B2.1)*

Partner Organisations' competence in results frameworks has grown:

GRAPH 31: To what extent have the services contributed to competence building and better overall understanding and outlook to RBM for partners? (B2.1.3; combined cohorts: Sida Commissioners & Partner Organisations)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
120	121	122	123	124	125	126	127												

Common trends and statements in this qualitative evidence include:

- Sida Commissioners report an increase in the proficiency in which Partner Organisations use and report from results frameworks.
- Partner Organisations state that the service did improve their competence in many ways. While there is no definitive trend, common statements include the use of better models and templates for reporting, improved understanding of results frameworks amongst relative staff, and better ability to meet Sida's reporting requirements.
- Some Partner Organisations state that they appreciate this as a service from Sida, one that aims at their own work with results frameworks, and that this enables them to develop a broader understanding of RBM rather than simply getting better at Sida reporting requirements.
- 'Neutral' and 'negative' comments often dealt with the lack of specificity to their organisation and the time allowed, as emphasised also in relation to overall quality. (See Section 4.5)

When Sida Commissioners and Partner Organisations' qualitative evidence is analysed separately, the same pattern exists, mainly 77% 'positive' ranked data. Further analysis reveals that Sida Commissioners, while overwhelmingly positive as well, are more vague and tepid, referring to the basest improvements in competence, e.g. having a suitable template in place. Whereas Partner Organisations regularly cite demonstrable and practical learning, things they have learned and applied. (See "Data & Analysis Report", pages 328 - 332.)

It is possible that Partner Organisation respondents have been unduly influenced by perceptions that this is a Sida sponsored Review and, as such, it would behove them

not to be too critical. Some respondents did mention, without prompting, that this service supported their bid for more funding. At the same time, the Review Team did not sense that respondents were being disingenuous; they seemed sincere and since this sincerity was widespread across qualitative data it may be indeed genuine.

Potential weaknesses in this data set imply that the negative comments, while few, should be heeded carefully. These, as a data set, are exceptionally diverse and random. They relate to issues of Sida's requirements, to RBM jargon, to the assignments' duration, to other seemingly random frustrations and complaints. This actually supports the broader positive pattern. If there were issues that were constraining the competence building of a significant proportion of Partner Organisation respondents, we would have expected to see this as more pronounced in an even numerically limited set of negative comments.

3.5.2.2 *To what extent has the Framework Agreement improved quality on Sida's partner organisations' results frameworks? (B2.1.2)*

Given Advisory Services potential contribution to RBM competence (Section 4.6.1.1.), it may not be surprising that Partner Organisation respondents were overwhelmingly positive on this point:

GRAPH 32: To what extent has the Framework Agreement improved quality on Sida's partner organisations results frameworks? (B2.1.2; Partner Organisations)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43																

This demonstrates that over 67% of respondents provided statements that could be judged as 'positive'. When analysed more closely, Partner Organisations cited specific ways they have changed their work with Results Frameworks. While beyond this Review's remit to qualify these, they are seemingly substantive.

Of course, the only demonstrable proof of 'improved quality' on these results framework will be an increased facility and satisfaction amongst relevant Sida commissioners and increased capacity to adjust programmes and initiatives for greater performance.

Recommendation 10 (Indevelop): The increasing importance of RBM and results reporting implies that organisations must not only instil these as part of standard business practices but to use these as a tool for increased performance. Toward that, both Indevelop and Sida would be served well by a post-facto survey among Partner Organisations that have received Advisory Services. This survey could be designed to gauge how, in precise and practical terms, programmes have changed due to their increased use of RBM practices. This implies that this survey would be conducted

after at least one full cycle of reporting. This Review recognises that this represents a potential cost and yet this analysis should provide greater insight into exactly how Advisory Services contribute to increased quality and how Sida may therefore better target resources going forward.

3.5.2.3 *To what extent have the services contributed to competence building and better overall understanding and outlook to RBM for partners? (B2.1.3)*

There has been significant competence building and this can be attributed to the consultants' expertise in RBM and the way they engage with partner organisations overall. This is confirmed through qualitative evidence. (Graph 33.)

GRAPH 33: To what extent have the services contributed to competence building and better overall understanding and outlook to RBM for partners? (B2.1.3; combined cohorts: Sida Commissioners & Partner Organisations)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
120	121	122	123	124	125	126	127												

Common trends and statements from this qualitative evidence include:

- Sida Commissioners state varied aspects of competence building, often associated with the facilitative skill of the consultant and the fact that they have seen an increased quality in reporting.
- As with qualitative trends noted in Section 4.6.1.1, Partner Organisations state that their staff both directly related to the results frameworks and beyond, have improved their competence and understanding of results frameworks. They sometimes state that this includes a “de-mystification” of the subject and a better understanding of how RBM can be used as a tool for overall performance.

While this data set depends on self-reporting from Partner Organisations, as with other data related to Advisory Services, there is a definitive trend toward overall quality and positive impact. As in Section 4.6.1.1, most partner organisations note the way consultants engage with them the genuine interest they take in the organisation, as well as their level of expertise regarding RBM. This is confirmed through the quantitative survey that asks Partner Organisations about consultants' overall competence and the relevance of the services they provide.

3.5.2.4 To what extent have the consultants been service oriented in terms of e.g., accessibility, flexibility, timeliness, readiness for dialogue and solving problems in a constructive way? (B2.1.4)

As with evaluation service, a significantly high level of service is evident. (Graph 32.) This shows that over 70% of respondents had positive comments regarding client service. This is confirmed by the quantitative survey that shows high rankings on specific aspects of service. (Graph 34.)

GRAPH 34: To what extent have the consultants been service oriented in terms of e.g., accessibility, flexibility, timeliness, readiness for dialogue and solving problems in a constructive way? (B2.1.4; combined cohorts: Sida Commissioners & Partner Organisations)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70									

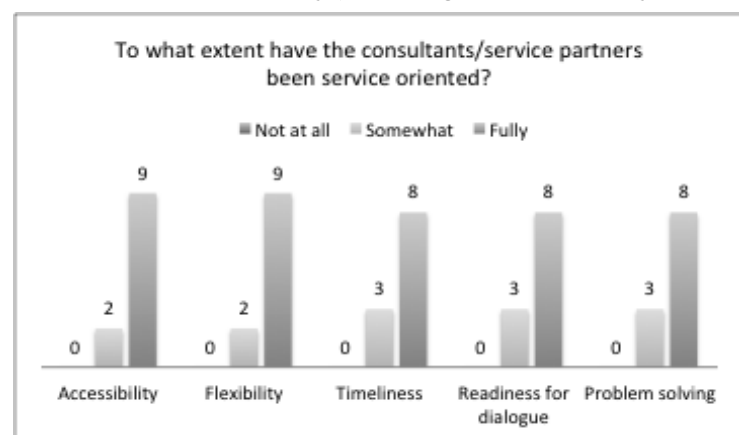
Common trends and statements from this qualitative evidence include:

- While many respondents simply noted that service levels are ‘good’ or ‘effective’ a more common issue relates to the communication skills, facilitative approach, and adaptability that the consultants brought to bear in their work.
- Partner Organisations often comment upon the consultants’ effective communication, that they were respectful of different levels of knowledge, of different types of people and their comfort with results frameworks, and with other facilitative/adaptive communication skills.
- Some Partner Organisations comment that consultants demonstrated a genuine interest in the organisation and its work.
- While no substantive trend exists in the ‘negative’ comments, there is some indication that the lack of ‘face-to-face’ meeting may have supported greater service.

While these service levels are to be expected, they are particularly important in relation to the level of facilitation, coaching, and knowledge sharing that these services imply.

This is confirmed time and again in how partner organisations describe these services. When positively impressed, they attribute

GRAPH 35: Quantitative Survey (Partner Organisations: Advisory Services)



this to the way the consultants worked with them, their staffs, and the facilitative and participatory process they used to introduce best practices. This is not only effective but important given the relatively short duration of these assignments.

3.6 ADVISORY SERVICES ON RESULTS FRAMEWORKS (MODEL)

3.6.1.1 *Is the current model of the Framework Agreement appropriate and effective? (B3.1)*

As with the description of the model in relation to evaluation services, the Model provides an efficient way for Sida programme Officers to suggest and procure needed services on results frameworks. The number of assignments procured represents a need and this need will surely grow going forward.⁶⁷ The FA serves the development of effective results based frameworks that are aligned with Sida's emerging standards.

However, it is not clear that there is a specific value to having these advisory services under the same framework as evaluation services. Key respondents indicated that this was done largely out of convenience rather than based on a specific strategy. Of course, this is aligned with UTV's role that includes both evaluations and results.

The nature of these services, being based on capacity building, facilitative/participatory workshops, and the development of common standards, models and templates, represent a particular level of service. It requires an in-depth understanding of Sida and its requirements, a particular expertise in RBM, and the ability to teach and coach diverse partner organisations. These characteristics, combined with the relatively brief duration of the contracts, are served well by a FA that awards the majority of work to one Service Partner.

⁶⁷ The most prominent milestone in the shift toward effectiveness and demonstrable results is the Fourth High Level Forum on Aid Effectiveness in Busan, South Korea, 2011. For an interesting perspective on this, see: Emma Mawdsley, Laura Savage, and Sung-Mi Kim, "A 'post-aid world'? Paradigm shift in foreign aid and development cooperation at the 2011 Busan High Level Forum." *The Geographical Journal*, January 2013. There are also interesting debates emerging about the potential conflict between the alignment with government and the need to demonstrate results. See, Martin Sjöstedt, "Aid Effectiveness and the Paris Declaration: A Mismatch Between Ownership and Results Based Management?" *Public Administration and Development*, Vol. 33; May 2013.

Recommendation 11 (UTV): Future framework agreements may include a separate contract for advisory services on results frameworks and apply a ‘winner take all’ approach, e.g. expecting to award the services to one service provider.

4 Conclusions

Evaluation Quality

Overall quality is sufficiently mixed to indicate the limitation of the FA to promote quality evaluations. This is based on an analysis of the evaluations process (utilisation), how evaluations are used (outcome), the reports themselves (outputs) as well as reference to evaluations conducted by Sida before the FA and to common best practices. Given this analysis, quality is largely inconsistent and may therefore have little to do with the FA itself.

Quality includes the entire gamut of evaluation activities, from the formulation of a ToR to the engagement/utilisation levels during Inception, to the use and analysis of relevant data, to the report and how the report is used. This review has analysed these and found that while there are some areas of noticeable quality improvement (engagement/utilisation, methodologies, and terms of reference), the overall quality lies between being “adequate” and “good.” This is based on the cumulative evidence presented in this Review and best practices and commons standards for evaluations.

The most prominent deficiency in this regard is the lack of applied methodologies for data collection and analysis. This is the basis for ‘evidence based’ evaluations and is critical when evaluating any level of evaluation or for distinguishing causality within a theory of change or from input to outputs and outcomes. It is an essential element in what makes ‘evaluation’ different from academic treaties or simply the opinions of well-meaning and learned experts. As we note in this Review and as is highlighted throughout the literature, this focus on evidence is critical for our understanding of development activities as well as for the acceptance and use of evaluation findings and recommendations. The majority would surely benefit from a much stronger evidence base for their conclusions. Many did and did so well. But, the majority did not.⁶⁸

⁶⁸ This was noted as an issue in “Are Sida Evaluations Good Enough” as well. “Are Sida Evaluations Good Enough? An Assessment of 34 Evaluation Reports,” Kim Forss, Evert Vedung, Stein Erik Kruse, Agnes Mwaiselge, and Anna Nilsdotter. Sida, 2008. Page 79.

If we refer to the previous evaluation of Sida evaluations, the “Are Sida Evaluations Good Enough?” report, we can compare the issues the authors there determined as primary quality problems with evaluations:

Are Sida Evaluations Good Enough? (Pages 77, 78)	Review	Comments of this MTR
Problems in the way evaluations are initiated and the formulation of TOR	No Change	While there has been a noticeable improvement in ToRs, many still fail to abide by the Sida template and/or provide sufficient detail for an evaluation team to efficiently develop methodologies, tools and approaches for the Evaluation.
Weak capacity among Sida desk officers to provide technical support to the evaluation process	Mixed	While evidence is mixed regarding Sida Commissioners in this regard, this is addressed somewhat through improved utilisation/engagement levels during the inception phase. UTV resources for support have decreased in the last 12 months and so this may be a greater issue going forward.
A limited number of qualified consultants, and missing skills and capacities among the evaluators with which Sida works, which might reflect a lack of competition and little variety among consultants	No Change	There are fewer consultants and less competition in their selection. The deficiencies in the reports assessed in this Review suggest a negligence of best practices and common standards.
Bias, as many of those who evaluate also plan and implement interventions	Improved	While this arose a few times and more noticeably in relation to the advisory services, no issues of bias were detected directly
Insufficient professional development in the field of evaluation in Sweden—few courses and other training opportunities	Not part of this Review	While beyond the scope of this review there is a sometimes explicitly mentioning of the expectation that the FA will facilitate the development of evaluation expertise in the Core Team and, less explicitly, within one service provider. As noted, there has been some individual development and some quality improvements in some areas but this has been minimal at best and may take far longer than the 5-year term of the FA to manifest in significant improvements. This also limits the potential of other service providers to, perhaps, start from a higher base of quality and/or to ascend the learning curve more quickly and effectively.
Poor incentives to carry out good monitoring and evaluation – both within Sida and the Swedish embassies, and for evaluators	Limited Improvement	Most evaluations are part of normal programme cycle requirements and there is an increased priority on evaluations and RBM overall. However, the overall impact of this is uncertain at this time.
A weak quality assurance system at Sida;	No Change	The QA aspect of this has largely been outsourced to the service provider thus eliminating a critical, informed third-party view to the evaluations conducted by Sida and Swedish embassies.
A low level of genuine demand for and utilisation of evaluations	Limited Improvement	Most evaluations are part of normal programme cycle requirements and there is an increased priority on evaluations and RBM overall. The overall impact of this is uncertain at this time.

The Model

All of these and other aspects of this Review indicate that the FA may have a negative impact on the long-term availability and quality of Evaluation services to Sida. The dominance of one service partner, given the diversity of evaluation needs, the quickly developing standards associated with methodologies, tools, and approaches, the plethora of innovations and new tools that emerge internationally, and the access of different evaluators to development evaluations at Sida, simply implies too much of a constriction on what is a dynamic and quickly changing field and market.

If the quality of evaluations was demonstrably higher than what had occurred before or according to common standards and best practices, then these market issues may not be of a primary concern. This is not the case and so these market issues are directly relevant to any future framework agreement for development evaluations.

Operationalisation

Sida has effectively developed and implemented a complex framework agreement.

Until now, Sida Embassies have made slightly higher use of the FA agreement (60%) than Sida headquarters' (40%), and they used it for a variety of evaluations and reviews that entailed varying approaches and tools.

The consultant base that carried out these evaluations and reviews is generally diverse (international-Swedish-local consultants) and gender-balanced, but there are fewer female team leaders (36%). Sida Commissioners perceive consultants' competencies acceptable overall although this in and of itself is not directly related to quality, especially regarding the technical aspects of evaluations that may be beyond Sida Commissioners' fields of expertise.

On Sida's side, issues remain about the role and legal interpretation of the FA and high staff turnover impedes the creation of institutional memory both for the operational use of the FA and lessons learned for the organisation. Commissioning staff generally manage on their own and rely on guidance material from UTV, although UTV supports them when asked. There still remains an opportunity for Sida (UTV) to communicate the use and benefits of the FA in a strategic way across all departments and embassies.

Advisory Services

Advisory services for results based management have adequately been used and met immediate needs. Indevalop has shown an appropriate balance of standardized tools and facilitative/participatory approaches in providing relevant and tailored services to Partner Organizations; consultant competencies are perceived to be largely adequate.

Partner Organizations express strong appreciation of the dedicated way the Service Provider interacted with them and reported high degrees of utilization of the services received. They criticize the short provision of initially 2 days and recognise how this has changed during the duration of the FA. In brief, UTV originally funded these 2-day assignments thus prompting Sida commissioners to use these services but not to exceed this 2-day UTV funded limit. UTV stopped “offering” 2-day services while still communicating that these services existed.

There is a lack of harmony in how Sida is developing its approach to performance and results. And there seems not to have been much communication about the availability of these services at headquarters and to all Embassies, which presents a missed opportunity.

The nature of these services, different from evaluations, requires a tailor-made approach that focuses on capacity building and participatory workshops as well as in-depth knowledge of Sida specific requirements on results based management. Combined with the brief duration of the contract, the current ‘the winner takes it all’ model in place seems well suited for the purpose of advisory services.

5 Lessons Learned

While many lessons are implied in the Findings presented above, there are several worth highlighting.

Quality

- Evaluation is a dynamic and complex activity. This means that there are few standards or processes that can dictate fully how evaluations are conducted. This is why some Sida Commissioners and Partner Organisations may find an evaluation that shirks these basic standards as effective. Simply because one is ‘satisfied’ doesn’t imply that maximum value was obtained.

There are enough standards and principles to guide evaluation quality, especially from Sida’s own Manual, OECD DAC, UNEG, and the World Bank’s IEG. These go beyond issues of independence and credibility and forge standardised evidence based tools and approaches. Big data analytics and architectures are also providing models for how to organise and analyse voluminous, complicated, and diverse qualitative and quantitative data. These too should be incorporated into evaluations as the subjects they treat demand. These emerging standards and best practices are there and should result in even higher quality and better results.

- The assumption that the FA could provide a significant increase in quality in and of itself, or simply because of the partnership potential with a single supplier, is somewhat negligent of market dynamics, commercial interests, and best practices regarding procurement. Framework agreements stand as models for procuring services, reducing the costs associated with that procurement, and ensuring a steady supply of services over a set period. The quality of those services depends on the competence and expertise of the suppliers, their on-going commitment and continuous improvement, and on the procurer’s ability to consistently hold them accountable. This can happen in a one-off contract as much as in a continuous stream of work.
- UTV can enhance the way it supports Sida Commissioners within Sida at Swedish embassies. While having standard templates within Sida for commissioning evaluations and reviews are of undisputable value, the outreach about their availability and terms of use, e.g. that not all aspects have to be followed if they are not essential for assessing an intervention, could be enhanced. The publication of a template is not enough. It needs to be supported with informal and formal support and a consistency of messaging about the value and im-

portance of these templates and the quality that they assure. UTV has proven its ability to shift practices within Sida, Swedish Embassies, and among the Service providers. It can and should continue to do this, to be the harbinger of what evaluation quality means and how it can be achieved.

Model

- When developing a procurement model, the broadest market implications must be carefully considered. This is not only the case in Sweden but internationally. Framework agreements are compromises. They provide greater efficiencies and potential costs savings while constraining competition. Competition in and of itself is not a guarantor of quality but it does enable those procuring service to make a more informed choice, to compare team and approaches and overall levels of service. This is especially important for complex services like evaluations. This field is developing quickly and it is within the market where advances are being made.
- When developing a model that has implications for strategy, knowledge management, and institutional learning, current and projected resource requirements should be calculated diligently. There is considerable value in having a centralised view of what evaluations and advisory services are being conducted, by whom, for what purpose, and toward what results. It is, of course, Sida's primary responsibility to develop its own institutional learning and knowledge management but this does not imply that it cannot draw on Service Providers to support this. The FA hopes to facilitate this. Yet, the out-sourcing of this does raise issues of both credibility (What does a Service Provider choose to highlight? What don't they highlight?) and completeness (In this FA, almost 20% of evaluations have been done by the 2nd and 3rd ranked Service Providers).
- A highly committed, professional, and intellectually driven Service Provider's value cannot be underestimated. Indvelop has consistently worked to both meet and exceed expectations in many areas. However, one may assume that SIOU and Orgut would do the same if given the chance.

Operationalisation

- When UTV focuses on particular issues, like the ToR and the Methodology, they have a noticeable and positive impact. This is a key success and should be recognised as a useful lever going forward.
- Efficiency and cost savings are based on numbers. The number of hours it takes to do a process, the cost of those person hours, the total cost of services, etc. Benchmarks and on-going measurements must be part of standard business practices to accurately gauge changes in efficiency. These can then be

compared internally and to emerging standards for procurement processes across the EU and beyond. This should be a distinct priority.

- Greater attention can be made to the wider intellectual community. Many consultants, individual and institutional, express dismay about the FA and a concern about the quality of Sida evaluations under the FA. These may be without merit but Sida should not ignore the fact that it is a significant player in this community and actively engage in it whenever possible. This may imply having public forums for all interested consultants and firms as well as Sida staff where these evaluations may be discussed and knowledge shared. They can conduct their own synthesis of evaluations and disseminate this within and beyond Sida. Even if one Service Provider provides the majority of services, this does not mean that Sida cannot provide other forums and opportunities for the broader consultant community.

Advisory Services

- Demand for Advisory Services for improved results frameworks by Sida, by its partner organizations, and globally has increased. This is a growing area of need and Sida stands in a leading position to facilitate progress with results frameworks. The Advisory Services are working and should be extended to as many Partner Organisations as is possible.

6 Recommendations

These are drawn directly from the “Findings” section above and have been organised by intended stakeholder and whether intended for the current Framework Agreement or for after its conclusion in April 2015. This includes a reference to the section from which they are drawn.

Current Framework

Recommendation 2 (UTV): UTV should stress the need for evaluation questions that follow OECD DAC evaluations levels to Sida Commissioners. While there may be reasons to divert from or change the OECD DAC levels, a set of comprehensive and valid questions should be included in every ToR. UTV should also ensure that Service Providers follow these questions, developing an evaluation matrix that includes all questions, their data sources, limitations and constraints, and risks and opportunities. This Evaluation Matrix should be mandatory for all Inception Phase Reports. (Section 4.2.1.1)

Recommendation 3 (UTV & Service Providers): Provide Service Providers with expected standards and practices in relation to data collection methodologies and approaches, especially as they apply to field visits. Ensure that Service Providers improve data collection methodologies and approaches. These should adhere to common standards for mixed-method evaluation approaches and include standard protocols and analytics for any relevant qualitative evidence. Any data collection methods should have direct relevance to established evaluation questions. (Section 4.2.1.1)

Recommendation 4 (UTV & Service Partners): Service Partners should ensure that all Core Team members and team leaders recognise the importance of utilisation, especially during the inception phase, and that they are sufficiently versed in best practices for engagement. Include this as a primary element in Quality Assurance (QA), e.g. to ask and report on how Core Team members have successfully engaged stakeholders during and beyond the inception phase. UTV should also make every effort to ensure that Sida Commissioners understand the implications of engagement/utilisation and that it may require more time/resources, especially during the inception phase, to ensure its value. (Section 4.2.2.1)

Recommendation 5 (Indevelop (short term) UTV (long term)): Develop a common QA procedure for data collection and analysis during the field/data collection phase and for engagement/utilisation, amongst all other areas of the evaluation, and demonstrate how this contributes to increased results. Include these results in Annual

Reports. In relation to data collection and analysis, this may include common analytics, tools, and business processes that will ensure efficiency and create economies of scale over the long-term. In relation to utilisation, this may include a standard pre- and post-inception phase interview between the Team Leader (TL) and the Service Provider Manager to facilitate the TL's approach and understanding of the importance of utilisation/engagement. This may be undertaken by Indevelop in the short term. They may develop their tool-kit and other mechanisms in collaboration with UTV. In the long term, UTV can develop guidance on QA and what it should include, especially with regard to utilisation/engagement, methodologies, data collection and analysis, and other matters it so deems relevant. (Section 4.2.3.1)

Recommendation 6 (Indevelop): Update Toolkit to reflect best practices and common standards in data collection, analysis, and results. Provide an overview/training on this to all Core Team members. (Section 4.2.3.1)

Recommendation 10 (Indevelop): The increasing importance of RBM and results reporting implies that organisations must not only instil these as part of standard business practices but to use these as a tool for increased performance. Toward that, both Indevelop and Sida would be served well by a post-facto survey among Partner Organisations that have received Advisory Services. This survey could be designed to gauge how, in precise and practical terms, programmes have changed due to their increased use of RBM practices. This implies that this survey would be conducted after at least one full cycle of reporting. This Review recognises that this represents a potential cost and yet this analysis should provide greater insight into exactly how Advisory Services contribute to increased quality and how Sida may therefore better target resources going forward. (Section 4.6.1.2.)

Recommendation 8 (UTV): UTV has an opportunity to have a much more informed and broader view of the evaluations being undertaken as a result of the FA. At the moment, most of the information and institutional knowledge that serves this purpose resides with Indevelop. UTV could review what information, statistics, or materials are important for on-going institutional learning about development evaluations in collaboration with Indevelop. UTV could then standardise this for any future framework agreements and/or for other Service Providers. (Section 4.3.4.2)

Recommendation 9 (UTV): Develop guidance for Sida Commissioners about how much an evaluation should cost as based on scope, requirements, and proposed methodologies. This can be based on the experience of the services procured under the FA and a relevant benchmark for activities going forward. A useful approach to this is minimum and maximums (min-max) for any requirement/approach. Another potential reference is to think in terms of a percentage of total programme investment. While any such guidelines may seem somewhat arbitrary at first, they will provide a foundation for more proficient standards in the long term. It should be noted that the current ranked framework agreement limits the room for negotiation of terms and costs but

this may be rectified if Sida moves to mini-competitions for any future framework agreements. (Section 4.3.5.1)

Post-Framework Agreement

Recommendation 1 (UTV): If a similar Framework Agreement is used in the future, Sida should insist on a small core team (max 10) whose primary expertise and role is to serve as the evaluation lead on any assignment. If this is maintained across multiple Service Providers, as would be the case in ‘mini-competitions’, then diversity of talent and perspective could be maintained. This will lend to higher levels of quality overall by having a smaller group accountable to the highest possible evaluation standards. At the same time, if consultants adhere to common standards and best practices in evaluation and rigorous QA is done for data collection and analysis, as recommended, the result should be higher quality of evaluation reports no matter the consultant. While this does present a limitation, in view of the diversity of consultants who may bring different perspectives and innovative tools, this is superseded by the need to have a more consolidated and standardized approach to how evaluations are conducted. (Section 4.1.3.1)

Recommendation 7 (UTV): The ranked system wherein the majority of the work goes to the first ranked supplier is neither a guarantor of increased quality or consistency. A system of ‘mini-competitions’ between pre-selected suppliers is more appropriate. This will enable Sida Commissioners to have more informed decision-making about teams and approaches amongst a broader range of consultants. This will ward against the increased encumbrance of one supplier. It will increase the knowledge about Sida’s decentralised evaluations within and beyond Sweden and will re-engage the intellectual community in Sweden that may not have had access to the development in these Sida Evaluations. (Section 4.3.3.2)

Recommendation 11 (UTV): Future framework agreements may include a separate contract for advisory services on results frameworks and apply a ‘winner take all’ approach, e.g. expecting to award the services to one service provider. (Section 4.7.1.1)

7 Annex

7.1 TERMS OF REFERENCE



Appendix 5

Reference number: 13/000428

Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks

TERMS OF REFERENCE

1. Background/ Introduction

1.1 Function and use of Sida Evaluations

Evaluation is an important prerequisite for results based management and learning. It helps us understand *how* and *why* certain results were achieved whereas others were not and thereby complements regular monitoring. Evaluation also includes a normative assessment of results against set criteria. Evaluation should “*provide useful and credible evidence to strengthen accountability for development results or contribute to learning processes, or both.*” (OECD DAC Quality Standards for Development 2010:p.7)⁶⁹

Evaluations are defined by Sida as careful, systematic and retrospective assessments of the design, implementation and results of development activities. Sida conducts two categories of evaluations:

Evaluations managed by Sida’s Monitoring & Evaluation Unit (UTV) are referred to as “strategic evaluations”. These typically deal with issues of agency-wide interest and are principally used for strategic steering, reporting and learning. A strategic evaluation plan is elaborated by UTV and decided upon by the Director General.

Evaluations managed by other Sida units are referred to as “decentralised evaluations” (formerly “reviews”). Serving primarily the needs of the commissioning unit or Embassy, these are often of operational character, and are mostly used for decision making regarding Sida funded projects, programmes and strategies. The decentralised evaluations shall preferably be included in each unit’s annual operational plan. Sida is commissioning about 80 decentralised evaluations per year.

1.2 The Framework Agreement

To make Sida’s evaluation system more efficient and improve the quality and use of the evaluations, Sida procured a framework agreement that came into effect on 11 April 2011 and is valid until 10 April 2015. Before this Framework Agreement most of Sida’s evaluations were commissioned through different thematic framework agreements, defined by sector, at different departments and units. The aim of bringing all Sida’s evaluations and results advisory services under one Framework Agreement was to ensure the provision of high quality evaluation and advisory services. The framework agreement covers services in two main areas:

⁶⁹ Useful evaluation guidelines from the OECD DAC, including on quality standards, managing joint evaluations, key terms in results based management and evaluation, evaluating humanitarian assistance, conflict- and peace building activities etc. are found at http://www.oecd.org/document/11/0,3746,en_35038640_35039563_35126667_1_1_1_1,00.html.

1. The design and implementation of *strategic and decentralised evaluations*⁷⁰ of contributions funded by Sida, including pre-studies/ concept papers as input to the design of evaluations, with a contract value not exceeding 1,200,000 SEK.⁷¹
2. *Advice on results framework* for contributions funded by Sida, with a contract value not exceeding 100,000 SEK.⁷²

The Framework Agreement should always be used when Sida HQ procures the above-mentioned services. Embassies and field offices *may choose but are not required* to use this framework agreement (alternatively they can procure using the usual procedures in accordance with the Swedish Public Procurement Act).

The decision to establish a framework agreement was partly a consequence of the findings and conclusions from a study on the quality of Sida evaluations: *“Are Sida Evaluation good Enough? An Assessment of 34 Evaluation Reports” (2008)*. The overall conclusion of the study was that many Sida evaluations were weak, assessed from the DAC/OECD quality standards. Weaknesses were found both on the commissioning side (such as weak ToR) and on the implementing side (for example inadequate methods, analysis and formulation of recommendations).

The Framework Agreement is a so-called “ranked” framework agreement. Contracts were signed with three consultancy firms: 1) Indvelop AB, 2) The Consortium SIPU International, 3) The Consortium ORGUT Consulting AB and Sida shall always first ask for services from the consultant ranked highest. If the first company is unable to respond to the request (for reasons of capacity, conflict of interest or other) the request passes to the company ranked second (and so on.)

The Framework Agreement was designed with the intention to improve the quality of ToR, reduce the transactions costs, get access to more competent consultants and to promote quality improvements through dialogue with the consultants.

⁷⁰ Decentralised evaluations were called ‘reviews’ until 2011. Thus, the term ‘review’ is used consistently in the framework agreements (signed 2011), and is to be seen as synonymous to ‘decentralised evaluation’.

⁷¹ In case of extension/add-on of an ongoing mission, the extended services may have maximum additional cost of 300,000 SEK in addition to the original budget.

⁷² During 2011 and 2012, call-off and financing of the advisory services were made in two different ways depending on the scope of the assignment; for so-called ‘limited advisory services’ not exceeding 20 000 SEK, a simplified call-off process was applied and UTV financed the assignment. As of March 2013, the possibility to use this simplified process was phased out and hence a regular call-off is to be applied and regardless of the scope of the assignment, and each calling-off unit has to finance the assignment through its respective operational appropriation(s).

The results framework component was a new service intended to raise the quality of partner's results frameworks and thereby also the quality of Sida evaluations and Sida-supported programmes. Advisory services are to be used in a way that strengthens the cooperation partner's ability to follow up and achieve results. The services should not be used for Sida's own capacity development on results management. The requirement of a clear link to the programme/project and the involved partners Sida supports, also means that the framework agreement does *not* include advisory services for the preparation or development of Sida's own "results proposals" (for the preparation of government strategy or follow-up of the implementation of Swedish strategies. Finished

To get access to competent consultants and sustain a constructive dialogue between Sida and consultants, the consultancy firms within the Framework Agreement are required to have a fixed team that focuses solely on evaluations (Core Team) and a management and quality assurance function in place (Management Team). The Core Team shall include a number of evaluators who possess advanced knowledge and expertise in conducting evaluations. The team also needs advanced knowledge in the kind of participatory processes which are required for so-called *utilisation-focused* evaluations (See Michael Patton). The Management Team should at a minimum include:

- Project director with the responsibility to ensure the appropriate quality control of services and communication with Sida.
- Project manager/coordinator with the responsibility to manage the contract and dealing with the selection and recruitment of experts for each specific assignment.
- Administrator/financial officer responsible for contractual and financial matters.

There shall also be a quality assurance system at a minimum including:

- A quality assurance plan.
- Policy and goals relating to quality assurance.
- A document management system.
- Names and responsibilities of the persons proposed for the quality assurance management and monitoring.

In order to ensure cost-effective communication, the Consultant shall have a representation in Sweden.

The Framework Agreement stresses the need for evaluations to be *utilisation-focused* meaning that all evaluations should be designed and implemented in a way that puts the intended users, and their use of the evaluation, on centre stage. This requires that the commissioner state who the intended users are, for whom the evaluation is undertaken and involve them to contribute to the evaluation design. It also requires that the evaluator carefully considers how activities in all stages of the evaluation (design, data collection, analysis, distribution, etc.) affect the intended users and uses.

Services listed under the Framework Agreement are:

- Manage and undertake utilization-focused evaluations/reviews of various kinds, including: summative, formative, theory-driven, impact assessments, lessons learning, rapid assessments; real time, meta- and synthesis.
- Manage and undertake feasibility studies/pre-studies/concept papers as input to the design of evaluations and reviews.
- Provide advice and input to Terms of References and the design of evaluations/reviews, including methodology.
- Conduct and facilitate participatory processes, workshops and seminars as part of the evaluation/review process.
- Searching, sampling and analysing information in Sida databases (currently E-doc and Plus), where some of the information is in Swedish.
- Analyse documents such as government bills, policies, strategies etc., of which some material will be in Swedish.
- Proof read and prepare reports for publication.
- Provide advisory services concerning results frameworks of Sida financed projects and programmes.

During the first year 45 evaluations were contracted within the Framework Agreement; 40 were conducted by Indevelop, four by SIPU International and one by ORGUT Consulting.

The first year 29 advisory services assignments were called-off; 28 were conducted by Indevelop and one by ORGUT Consulting.

2. General Purpose

The MTR shall serve as an input to improvements to be made in the second half of the agreement period as well as for the decision on how Sida will design a possible framework agreement when the current agreement expires. To this end, the idea is to study if the intentions of the Framework Agreement have been fulfilled. Has the Framework Agreement improved the quality and use of Sida evaluations and reduced the transactions costs? Has the Framework Agreement improved the quality of partner organisations' results frameworks? Do the Framework Agreement services meet the OECD/DAC quality standard and the needs of Sida? The MTR shall consider quality both in relation to the evaluation process and the evaluation reports. It shall also assess the quality of advisory services for results framework.

It is stated in the Framework Agreement's Terms of Reference that Sida will review the Framework Agreement after two years to ensure that it is meeting the evolving needs of Sida. Sida's Monitoring and Evaluation Unit (UTV) will manage the MTR, to be carried out by an external, independent consultant in collaboration with the stakeholders.

3. Users and Intended Use

There are two different groups of primary stakeholders – Sida/embassies and the evaluation consultancy firms - involved in the Framework Agreement. Sida can in turn be divided into two different “users” of the Framework Agreement: UTV and the procurement unit (UPPIS) responsible for *managing* the Framework Agreement and Sida units/embassies *commissioning and using* services under the Framework Agreement.

Representatives from Sida (UTV and UPPIS) will form an evaluation management group, which together with a reference group of representatives from Sida’s operational departments, will provide an oversight and quality assurance of this part of the evaluation. The reference group has contributed with input to these Terms of Reference and will be involved in every step of the evaluation process, and read and comment on inception report and draft reports. Also the consultancy firms have had the opportunity to comment on these ToR and will be engaged in dialogue and comment on all drafts during the MTR process.

The MTR is intended as a learning process for both Sida and consultancy firms, in their respective roles as commissioners/users of evaluations and providers of evaluation/advisory services. There is also an element of accountability, in respect to whether the consultancy firms, UTV/UPPIS and Sida commissioning units have fulfilled their obligations under the agreement.

In short, the intended use of the MTR

is For UTV and UPPIS:

- To ensure that the framework agreement has been properly complied to.
- As an input to the dialogue with the consultant regarding necessary adjustments/improvement during the second half of the contract period.
- Identification of areas for adjustments/improvement in relation to the commissioning Sida units and the relations between the commissioning units and the consultants.
- As an input to the decision on how Sida will design a possible future Framework Agreement.

Commissioning Sida units

- As an input to more effective use of the framework agreement and the evaluation tool at large.
- To identify where improvements are needed in relation to ToRs, matching scale of assignments with resources, etc.

Consultants:

- To get a better notion of Sida's needs and as an input to possible adjustment/improvements of services.
- To identify possible areas where the consultant needs to strengthen its capacity.

4. Evaluation Questions

The principal object of evaluation is the Framework Agreement and how it has been operationalized. The framework agreement can be said to represent one "model" of procuring evaluation services. Analysing the Framework Agreement in turn requires studying the quality of services and the interplay between Sida's commissioning units, UTV/UPPIS and the consultant.

The MTR shall analyse if and if so, how, the Framework Agreement has contributed to better quality of Sida evaluations and partner results frameworks and thereby better results based management. The MTR shall also analyse if the framework agreement model has reduced the transaction cost in Sida's evaluation processes. In order to do this plausibly, the evaluation shall reason in terms of a counterfactual. Two alternatives to be considered would be i) no separate framework agreement for evaluations and ii) a framework agreement constructed differently (for example, not ranked).

More specific evaluation questions are listed below.

4.1 Evaluation Component

A. Operationalisation of the Framework Agreement

How has the Framework Agreement been put to practice and used, and what change can be seen in compare to before the Framework Agreement?

- How many evaluations have been conducted within the Framework Agreement (total and per consultant) and how many have they turned down?
- What kinds of evaluations/services have been provided?
- To what extent has the Framework Agreement been used also by the embassies?
- How has the dialogue and collaboration between the consultants UTV/UPPIS/Sida been?
- What have been the costs of the evaluations in the Framework Agreement?
- What have the teams looked like in terms of gender, Swedish/national/international team members and age?

- How have the consultants organized themselves to respond to the Framework agreement, including for administration, providing qualified evaluation teams, establishment of a Core Team and quality assurance? Has it been adequate and appropriate?
- What competence has been offered by the consultants? Has it been adequate and appropriate?
- How has Sida organized itself to manage and utilize the framework agreement, including institutional learning, quality assurance, legal and administrative matters, support and communication? Has it been adequate and appropriate?

B. Quality of Services

Has the evaluation quality changed since the signing of the Framework Agreement?

- Has evaluation quality improved?
- To what extent are Sida evaluations under the Framework Agreement utilisation focused?
- To what extent and in what different ways are Sida's evaluations used by Sida and its partner organisations (including both instrumental and process use)?
- How have the evaluated partner organisations experienced the evaluation process?
- To what extent have the consultants been service oriented in terms of e.g., accessibility, flexibility, timeliness, readiness for dialogue and solving problems in a constructive way?

C. Model

Is the current model⁷³ of the Framework Agreement functional?

- Overall, is there evidence that the Framework Agreement has improved evaluation quality?
- How have the costs for evaluation services changed under the Framework Agreement?

⁷³ “Current model” refers to a ranked Framework Agreement as expressed in Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Framework (Case number: 2010-001697), including a fixed Core Team and services with a maximum value of SEK 1 200 000 respectively SEK 100 000.

- Have the transaction costs been reduced since the Framework Agreement was implemented?
- Has the Framework Agreement led to a better commissioning of evaluation services, including better formulated ToR and better quality assurance, and if so, to what extent and in what ways?
- Have the consultants improved their evaluation competence and/or strengthened and accumulated knowledge about Sida to offer better services?
- What possible incentives (for the consultant and for Sida) does the Framework Agreement give rise to?
- What does the Framework Agreement imply for the long term availability and quality of evaluation services to Sida, taking the resource base as a whole into consideration?
- What are the implications of the Framework Agreement for local evaluation consultants in partner countries? Could the Framework Agreement be improved to better contribute to enhanced development of capacities of evaluators in partner countries?
- What has the framework agreement implied for young evaluators to be engaged in the provision of evaluation services to Sida?
- To what extent does the Framework Agreement contributed to Sida's institutional learning on evaluation?
- What would be the advantages and disadvantages of having a framework agreement with "renewed competition" (as opposed to a ranked agreement)?
- What would be the pros and cons of not having a separate framework agreement for evaluation and instead let evaluations be one of the services included in Sida's different thematic framework agreements?

4.2 Advisory Services on Results Framework Component

D. Operationalisation of the Advisory Services

How has the Framework Agreement been put to practice and used in relation to the Results Framework component?

- How many advisory assignments have been conducted within the Framework Agreement and how many have they turned down?
- What have the services consisted in?
- To what extent has the Framework Agreement been used by different Sida units/Embassies? What have been the needs and how have the services responded to these needs?
- How has the dialogue and collaboration between the

consultants UTV/UPPIS/Sida been?

- Duration and costs of services (per assignment and total)?
- How have the consultants organized themselves to respond to the Framework Agreement, including providing qualified evaluation teams, Core Team and quality assurance?
- What competence has been offered by the consultants? Has it been adequate and appropriate?
- Have Sida's structures for managing the framework contract been appropriate and effective, including quality assurance, legal matters, support and communication?

E. Quality of Services

Has the quality of results frameworks and partner competence on results frameworks and RBM improved with the Framework Agreement?

- To what extent have the consultants delivered effective and relevant, tailored advisory services for results frameworks?
- To what extent has the Framework Agreement improved quality on Sida's partner organisations results frameworks?
- To what extent have the services contributed to competence building and better overall understanding and outlook to RBM for partners? Including HRBA, perspectives of the poor, emphasis on gender equality etc. in partners' results frameworks.
- To what extent have the consultants been service oriented in terms of e.g., accessibility, flexibility, timeliness, readiness for dialogue and solving problems in a constructive way?

F. Model

Is the current model of the Framework Agreement appropriate and effective?

5. Conclusions, Lessons and Recommendation

The MTR shall include conclusions and lessons in respect to the above mentioned questions, as well as clear recommendations, in relation to the earlier mentioned evaluation use. The recommendations should be directed to UTV/UPPIS, Sida's commissioning units and to the consultants respectively.

6. Methodology

The evaluator shall in its tender suggest a suitable method for the MTR to be further elaborated in an inception report.

Evaluation Component

The MTR shall assess the quality, utilisation focus and use of a representative sample of the Framework Agreement evaluations, considering both the reports and the evaluation processes. Quality refers to OECD/DAC quality standard (independence, credibility and utility) and Sida Evaluation Guidelines 2010.

For the quality assessment, the study "Are Sida Evaluations Good Enough? An Assessment of 34 Evaluation Reports" should be used as point of reference.

Where relevant, the MTR shall apply the same method as "Are Sida Evaluations Good Enough?". This study can thereby function as a "baseline" for the quality assessments of evaluations and ToR. Additionally, for process issues and questions on perception of services and their use, the evaluator(s) shall interview Sida program officers, partner organisations and consultants from the three consultancy firms within the Framework Agreement as well as a few consultants that tendered but were not selected for an understanding of how the Framework "model" has been implemented and functioned.

This can be done through meetings in Sweden, telephone or mail. No field visits are foreseen.

Advisory Services on Results Frameworks Component

Advisory service is a new service so no baseline exists. Interviews with and perceptions of Sida program officers, partner organisations and consultants will be a central part of the evaluation methodology, here as well.

No field visits are foreseen.

7. Time schedule and Reporting

The MTR should be carried out over the period August-November 2013.

An inception report shall be presented at an early stage of work, including but not necessarily limited to:

- A further elaboration on the evaluation questions and criteria
- Methodology, including data collection and analysis
- Selection issues
- Further detailing on stakeholder participation
- Possible delimitations to be agreed upon with Sida
- Risks and risk mitigation
- A detailed work programme
- A draft communication plan.

Findings shall be presented in a draft report to be delivered for comments from Sida, the reference group and framework agreement consultants. A final draft for Sida's approval is to be prepared by the consultant no later than two weeks after receipt of comments. A seminar/workshop shall also take place at this stage, for learning and feed-back to the consultants.

Reports shall be written in English and adhere to the OECD/DAC Glossary of Key Terms in Evaluation and Results Based Management. The final report shall not exceed 50 pages, excluding annexes. Format and outline of the report shall follow the guidelines in Sida Evaluation Manual "Looking Back, Moving Forward" – Annex B, Format for Sida Evaluation Reports.

Tentative timeframe	
Start of work	15 August
Inception report from consultants	15 September
First draft report	31 October
Final report	30 November

8. Resources

The MTR is estimated to require about 15-18 person weeks.

9. Evaluation Team Qualification

The MTR should be conducted by a consultancy team (2-3 persons) with considerable theoretical and practical experience of evaluation. At least one team member shall possess experience of evaluation in a development context. All team members shall have an advanced academic degree, i.e. a minimum of a Master's degree or equivalent.

Competence that must be represented in the team is in-depth knowledge about evaluation systems, experience of evaluation quality assessment, experience of working with results frameworks and RBM, experience of utilisation-focused evaluation, language skills in form of excellent English and Swedish.

One person shall be assigned team leader with the overall responsibility for the evaluation. The team leader shall have extensive experience of leading evaluations assignments.

To ensure impartiality, the evaluation shall be conducted by independent consultants without previous involvement in the Framework Agreement or the tendering process (neither persons nor consulting companies as legal entities) and having no stake in the outcome of the evaluation.

10. References

Are Sida Evaluation Good Enough? An Assessment of 34 Evaluation Reports
<http://www.oecd.org/derec/sweden/41390724.pdf>

Sida's General Condition For Framework Agreement (2002) - Appendix A

Framework Agreement For Sida Reviews, Evaluations And Advisory Services On Results Framework- Appendix B TERMS OF REFERENCE (Case number: 2010-001697)

Sida Evaluation Guidelines 2010- Appendix E

DAC Guidelines and Reference Series Quality Standard for Development Evaluation- Appendix F
<http://www.oecd.org/development/evaluation/dcdndep/36596604.pdf>

Indevelop Final Annual Report 2011-2012- Framework Agreement for Sida Review, Evaluations and Advisory Services on Results Frameworks

Sipu Annual Report 2011-2012- Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks, Case number: 2010-001697

7.2 INCEPTION PHASE REPORT (AS A SEPARATE ANNEX)

7.3 DATA & ANALYSIS REPORT (AS A SEPARATE ANNEX)

7.4 COMPARATIVE PROCUREMENT MODELS

Please see tables at the end of this section for abbreviations.

	FA scope	FA model (Single/ Ranked/ Multiple)	Contracting authorities (CA) reference	Number of SPs	Mini-competition	Core team	Pool of experts	Several CVs pro- posed/position	Mixed SP teams (upon CA selection)
Evaluation FA	Evaluation FA for all regions and all thematic sectors and different type of evaluations	Ranked	SIDA	3	NO	YES	YES	NO	NO
			Type of evaluation (*)	Utilisation focused evaluations/reviews including: summative, formative, theory driven, impact assessments, lessons learning, rapid assessments, real-time, meta and synthesis					
			Observations (**)	Majority of the assignments done by one SP (1st ranked one). Core team of the 1st SP expected to be trained on client's expectations. SP to have a strong QA role. This is limited to the 1st ranked SP; impossible for the second or third SP to maintain a management team or a Core Team given volume and pipeline of work. It is quick to procure but no comparison of expertise/approach is possible. The option to go to the second SP is hardly used. This is partly because it extends the procurement period for 2 more weeks with no guarantee the second SP will become the selected supplier.					
		Multiple	GEFA-DFID	N° SPs	Mini competition	Core team	Pool of experts	Several CVs/ position	Mixed SP teams
				27	10	NO	NO	NO	NO
			Type of evaluation (*)	All types of evaluations (e.g. Mid-term, Final, Impact) and design of multiyear M&E frameworks					
			Observations (**)	Highly competitive FA (sometimes up to 10 bids are received). Trend is to undertake long-term/multiyear evaluations (incl. baseline and end line survey, plus Mid-term evaluation) or the design and implementation of multiyear M&E frameworks in parallel to the implementation of the actual DFID programme to increase the accountability and measure impacts. Evaluation costs can vary from a few hundred thousand to several million pounds (Standard is 5% of the implementation budget). Proposals are in average 50 pages long (methodology). Commercial offer has long narrative part describing Value for Money (VFM), is benchmarking costs, outlines economy of scales and asked for explanations of the business process behind delivering on time and within budget.					
	for Policy evaluations, 1 for Strategy evaluations, 1 for Impact evaluations, Synthesis	Multiple	WFP	N° SPs	Mini competition	Core team	Pool of experts	Several CVs/ position	Mixed SP teams
				5	3	YES	YES	NO	NO
			Type of evaluation (*)	Long term Agreement for policy, strategy, synthesis, impact, cross-cutting evaluations					

			Observations (**)	WFP decides on number of applicants for each tender/proposals. Not all LTA members are always invited, sometimes even outside competitors can be asked due to specific evaluation requirements. But focus is on pre-selected members of the LTA. Evaluations do have clear budget ceilings. Expert costs are predetermined and are not part of proposal selection. Methodology and team description is max. 15 pages long. Feedback on proposals can be requested anytime. Elaborated uptake system of evaluation results to decision maker level. WFP has multiannual evaluation plans in place. Companies can prepare responses in advance. Each evaluation has a volume of app. 200.000 to 500.000 USD (with exceptions).					
	Evaluation FA per thematic sector (ex.: 1 for Governance, 1 for Agriculture, 1 for Health, 1 for Education)	Single	UNDEF	N° SPs	Mini competition	Core team	Pool of experts	Several CVs/ position	Mixed SP teams
				1	NO	YES	YES	NO	NO
			Type of evaluation (*)	Country-based project final evaluations, regional project final evaluations and thematic (cluster) final evaluations covering multiple projects. In total 90 country-regional evaluations and 6 cluster evaluations are to be conducted during the contractual period.					
			Observations (**)	Low volume evaluations (10 days maximum for an evaluation core team expert and a local expert). CA plans with the SP the evaluations yearly with trimestral updates. Core team of evaluation experts (8) with strong knowledge of the methodology and client's expectations systematically mobilised as Team leaders and supported by a local expert selected for each mission from a pool of experts, CA reject a local expert CV and ask for additional proposal until approval. Important quality assurance in place (service provider + the Core team leader) and the evaluations are standardized.					
			CTB	N° SPs	Mini competition	Core team	Pool of experts	Several CVs/ position	Mixed SP teams
				1	NO	YES	NO	NO	NO
			Type of evaluation (*)	Only mid-term and final evaluations per thematic area (Education, Governance, Infrastructure etc.) with limited evaluation scope on answering 5 DAC criteria					
			Observations (**)	CTB selects one provider per thematic area. Individual experts are also predetermined through a thorough expert profile selection for 3 years. (e.g. expert for Higher Education or with experiences in VET or curricula development). Evaluations are undertaken by one international and one local expert. Time frame is max. 25 days for each expert and mission. Annual summary reports are requested per thematic area by each SP. Evaluations are used mostly by operating teams to design next phase and or continue implementation efforts.					
		Multiple	EU ECHO	N° SPs	Mini competition	Core team	Pool of experts	Several CVs/ position	Mixed SP teams
				4	4	YES	YES	NO	NO
			Type of evaluation (*)	All types of evaluations in the field of Humanitarian activities and Civil Protection action					
			Observations (**)	Always 4 SPs who are submitting proposals and competing against each other. Full proposals and methodologies are requested. Core team of team leaders, core team of QA for each SP. Costs/ evaluations are around 200.000 to 500.000 Euros. Limited uptake approach and limited dissemination approach.					

Non-specific Evaluation FA	Evaluation FA per region (ex.: 1 for Asia, 1 for Africa, 1 for Latin America)	Single	EU ROM	N° SPs	Mini competition	Core team	Pool of experts	Several CVs/ position	Mixed SP teams
				1	1	YES	YES	NO	NO
			Type of evaluation (*)	Short external assessment/monitoring of projects and programmes and budget support					
			Observations (**)	Monitoring missions are 10 days long. Service provider is responsible for a whole region/continent (ACP, ALA, Europe, etc.). Core team and pool of experts are undertaking the missions across all subject matter sectors. Very strict reporting format. Data are captured in EC wide project/programme database. System has led to low pricing between competitors and quality of mission has been suffering over the years. Currently new ROM system under discussion and change of focus seems necessary. For the next years ROM is planned to be undertaken mostly for projects/programs in critical phases.					
	Global (TA and evaluation) FA per thematic sector (ex.: 1 for Governance, 1 for Agriculture, 1 for Health, 1 for Education)	Single	Previous ETF	N° SPs	Mini competition	Core team	Pool of experts	Several CVs/ position	Mixed SP teams
				1	NO	NO	YES	?	NO
			Type of evaluation (*)	TA to support the implementation work programmes but no evaluation					
			Observations (**)	CA looks for innovative research/academic expertise, constantly renewed. Prices are contractually fixed and systematically applied. QA requirements are set contractually however the CA has strong internal expertise and the staff in charge tend to bypass the SP for QA and practical coordination of the mission. At least 2 CVs requested for each position and given the FA set (no competition) the second one is often proposed to comply with the contractual requirement. No discussion about the ToR approaches as such though a lot of clarifications are provided by the CA during the tender period. No no-bid cases. Contractual delays to provide a proposal tend to be extended.					
		Ranked	Previous ETF	N° SPs	Mini competition	Core team	Pool of experts	Several CVs/ position	Mixed SP teams
				3	NO	NO	YES	?	NO
			Type of evaluation (*)	TA to support the implementation work programmes but no evaluation					
			Observations (**)	The first ranked SP declined almost all requests to the second who did the majority of the missions. The CA looks for innovative research/academic expertise, constantly renewed. Prices contractually fixed and systematically applied. QA requirements are set contractually however the CA has strong internal expertise and the staff in charge tend to bypass the SP for QA and practical coordination of the missions. No discussion about the ToR approaches as such though a lot of clarifications are provided by the CA during the tender period. The option to decline a SP proposal (here the second one) extends the tendering period (10 days to bid + 5 days to evaluate for each SP) and was never used. At least 2 CVs requested for each position and the opportunity to present 2 strong options is sometimes used though the contract set and implementation on the long run (no rejected proposals) does not encourage this. Very few call-offs declined by the 2nd SP (too heavy work load at a time on the FA).					

		Multiple	EuropeAid EU BENEF Lots 1 to 12	N° SPs	Mini competition	Core team	Pool of experts	Several CVs/ position	Mixed SP teams
				5 to 9	3 to 5	NO	YES	NO	NO
			Type of evaluation (*)	Ex-ante, interim, ex-post evaluations, monitoring of projects and programmes					
			Observations (**)	CA evaluates up to 5 teams/proposals. High competition between SP and all experts can potentially work with all SP. Maximum fee rates are set contractually on the basis of each SP proposal at the global tender phase. SP maximum fees rates are published and financial offers for specific missions tend to align on the cheapest SP rates. No possibility to mix SP teams (would not be manageable in terms of responsibility over the deliverables). QA standardized at the SP level, following CA guidelines, but diversity of approaches at the SP level is probably still important.					
			ETF	N° SPs	Mini competition	Core team	Pool of experts	Several CVs/ position	Mixed SP teams
				3	3	NO	YES	YES	YES
			Type of evaluation (*)	TA to support the implementation of work programmes but no evaluation					
			Observations (**)	SP can propose up to 2 CVs per position and this opportunity is almost systematically used as it maximises the chances to win the bid. CA can mix different SP teams. This option is used for missions where experts work independently from each other (ex: various training sessions): the CA can evaluate up to 6 CVS for the same position. In case of mixed SP teams a lot a micro-management is needed on the CA side to coordinate experts (whereas this is dealt by SP directly otherwise). No competition on prices as fee rates are set contractually and systematically applied. CA bypasses SP to coordinate missions and activities and the QA of SP is de facto very limited.					

Acronyms	
CA	Contracting Authority
CTB	Coopération Technique Belge
ECHO	Commission's European Community Humanitarian Office
ETF	European Training Foundation
FA	Framework Agreement
ROM	Result Oriented Monitoring
SIDA	Swedish International Development Cooperation Agency
SP	Service provider

Definition	
Ranked	The first ranked SP is requested to submit a proposal within the determined deadline. If he is not able to submit an offer or if the CA is not satisfied with the offer, the request must be passed on to the second ranked SP, etc.
Multiple	At least 3 pre-selected SPs are requested to submit a proposal within the determined deadline. The CA evaluates and selects the best one.
Single	The pre-selected SP is requested to submit a proposal within the determined deadline.

(*) source=contracts/TDR/dedicated official web-pages

(**) source=interviews at SIDA and Transtec

TA	Technical Assistance
TL	Team leader
UNDEF	United Nations Democracy Fund
EU	European Union
WFP	World Food Programme

7.5 REPORTS ASSESSED AS PART OF THIS REVIEW

The following reports were assessed as part of this Review. This included most reports conducted under the Framework Agreement. Some reports were not included as their purpose and content was not appropriate to the metrics being used for the assessment.

	Indevelop Evaluations	Reviewed?
1	Evaluation of the of Co-operation with the Statistical office Kosovo (RE1)	x
2	Revu externe du Projet d'Appui au Système Statistique National du Mali 2009-2012 (RE2)	
3	Evaluation of the MENA-OECD Investment Programme 2008-2010 (RE4)	x
4	Evaluation of Swedish Health Sector Programme Support in Uganda 2000-2010 (RE6)	x
5	Desk study of Diakonia (RE7)	x
6	Evaluation of Swedish East African Music Network (RE8)	x
7	Evaluation of Cultural Heritage without Borders (RE9)	x
8	Overview of International Training Programme "Management of Hydropower Development and Use" (RE10)	x
9	Utveckling av Metod för Resultatinsamling inom International Utbildnings Program (ITP) (RE11)	
10	External review of the Moz-SAKSS programme in Mozambique (RE12)	x
11	Evaluation of Life and Peace Institute (RE13)	x
12	Mid-term review of Sida support to the Centre for Science and Environment (RE14)	x
13	Evaluation of Diakonia Zimbabwe Programme 2009-2011 (RE15)	x
14	Evaluation of the Upland Research and Capacity Development Programme in Laos (RE16)	x
15	Review of the Technical Barriers to Trade Mentorship Programme 2008-2011 (RE17)	x
16	Evaluation of Olof Palme International Centre (RE18)	x
17	Review of the Greater Mekong Sub-Region Regional Power Trade (RE19)	x
18	Review of civil society organisations in Moldova (RE20)	x
19	Mapping of Sida funded projects in area C, Jerusalem and Seam Zones through the Palestinian Authority, UN, Palestinian, Israeli and International NGO's (RE21)	x
20	Evaluation of the Water Network Management Programme in Erbil, Iraqi Kurdistan (RE22)	x
21	Mid-Term Review and Evaluation of the Rural Electrification Programme in Zambia (RE23)	x
22	Review of UNDP Georgia Good Governance project (RE24)	x

23	Review of Cultural Heritage for Sustainable Development, a Programme of Museum Cooperation in Southeast Asia (MuSEA) (RE25)	x
24	Mid-Term Review of the Regional Human Rights Programme of the Ombudsman Institution (Defensoría del Pueblo) in Colombia (RE26)	x
25	Mid-term Review of the Capacity Development Project of the Rural Energy Agency, Tanzania (RE27)	x
26	Outcome Assessment and Lessons Learnt from Swedish Development Cooperation with Macedonia (1999–2012) (RE28)	x
27	Review of the EAC and MENA Carbon Footprint of Products Pilot Projects 2010–2011 (RE29)	x
28	Review of FOS, the Fund for the Colombian Civil Society for Peace, Human Rights and Democracy (RE30)	x
29	Evaluation of Sida's Support to the National University of Rwanda (NUR) 2007–2011 (RE31)	x
30	Interim Evaluation of Support to the Church of Sweden's Community Based Psychosocial Support in Humanitarian Assistance 2006–2011 (RE32)	x
31	Mid-term Review of the Bio-fuel project support to the Ministry of Energy and Minerals, Tanzania (RE33)	x
32	External Review of Core Support under Joint Financial Agreement to Zambia National Farmers Union (ZNFU) (RE34)	x
33	Evaluation of World Children's Prize (RE35)	x
34	Review of the CREDO Herzegovina (RE37)	x
35	Evaluation of the Maghreb Regional Training Programme Development of Public Employment Services in Algeria, Morocco and Tunisia (RE38)	x
36	Uncovering the Power of Performing Arts: An Evaluation of the Tamasi Programme (RE39)	x
37	Review of the Sida-funded Project Education for Sustainable Development in Action (ESDA) (RE40)	x
38	Study of Sida's Support to the Swedish Civil Contingencies Agency (MSB) 2006–2011 (RE41)	x
39	Evaluation of the BetterAid and Open Forum Programmes (RE42)	x
40	Evaluation of Forum Syd and Diakonia's Democracy and Human Rights programmes in Cambodia (RE43)	x
41	Evaluation of the Barbro Johansson Model Girls' Secondary School in Tanzania (RE44)	x
42	Evaluation of Legal and Human Rights Centre (LRHC) in Tanzania (RE46)	x
43	Resilience, Risk and Vulnerability at Sida (RE47)	x
44	Review of the project Capacity Building & Improved Client Services at the National Agency of Public Registry (NAPR) in Georgia (RE48)	x
45	Evaluation of Promoting the Integrity of Civil Data in Georgia (RE49)	x
46	Review of the project Capacity Building of the Georgian Leadership Community for Improved Decision-making and Negotiation Skills (CBGL) (RE50)	x
47	Evaluation of the Raoul Wallenberg Institute's regional programme "Building Human Rights Knowledge and Resources in the Middle East and North Africa" (RE51)	x

48	Review of the partnership cooperation between the National Bureau of Enforcement of Georgia and the Swedish Enforcement Agency (RE52)	x
49	Civil Society Support Modalities at Sida Headquarters and Swedish Embassies (RE53)	
50	Evaluation of the Sida and DFID funded Public Policy Information, Monitoring and Advocacy (PPIMA) project in Rwanda (RE54)	x
51	Evaluation of the Project “Protection of Children at Risk and Children in Contact with the Justice System in Bosnia and Herzegovina” (RE55)	x
52	Évaluation Externe du Plan Stratégique de Promotion de la Gouvernance Démocratique au Burkina Faso (RE56)	
53	Civil Society on Climate Change: Evaluation of the work on climate change adaptation and mitigation of Swedish Society for Nature Conservation in cooperation with partner organisations in the South (RE57)	x
54	External Evaluation of the Partnership Agreement for Sustainable Development of Lake Victoria Basin (RE58)	x
55	Evaluation of the International Foundation for Science (RE59)	x
56	Midterm Review of the ‘Capacity Development of Municipal Associations’ Programme (CDMA) (RE60)	x
57	Metodstöd vid genomförande av resultatsseminarium för ITP 268 LGBT and Human Rights (RE61)	
58	Summative Evaluation of phase out in East Timor (RE62)	Report NA
59	Evaluation of Save the Children Sweden’s Support to Organisations (RE64)	x
60	Evaluation of Sida’s Support to Environment Infrastructure and Reforms in Central and Eastern Europe and Western Balkans 1995-2010 – A Desk Study (RE65)	x
61	Mid-Term Review for the Quality Infrastructure and Standards Programme (QUISP) in Uganda (RE66)	x
62	Mid-Term Review of the AGIR Programme (RE67)	x
63	Review of the Sida-funded Project “Advisory Support to the Ministry of Finance of Ukraine” (RE68)	x
64	Reform cooperation in the Western Balkans – regional cooperation: experiences, constraints and opportunities (RE69)	x
65	Study on Results of Cooperation under the Swedish Strategy for Development Cooperation with Kosovo, 2009-2012 (RE71)	x
66	Strategic review and advice on regional comparative advantages in Eastern Europe and the Eastern Partnership (RE72)	x
67	Evaluation of Sida & NIR Core Support Program (2009 – 2012) (RE73)	
68	Review of Swedish Support to the World Customs Organization (WCO) Capacity Building, 2008-2012 (RE74)	x
69	Mid-term Review of the Institutional Strengthening Support to the Association of Local Authorities of Tanzania - ALAT (RE75)	x
70	Evaluation of Femina Health Information Project (HIP) Strategic Plan 2006-2012 (RE77)	x
71	Review of the Swedish cooperation strategy with the Russian Federation in the areas of human rights and democracy (RE78)	x
72	Evaluation of the Institutional Cooperation Between Ministry of Lands, Kenya	x

	and Lantmäteriet, Sweden on the Project for Improving Land Administration in Kenya 2009-2012 (PILAK) (RE79)	
73	Monitoring of Rwanda Peacebuilding Programme (RE80)	Ongoing
74	Evaluation of Results of Sweden's Development Cooperation Strategy Albania (RE81)	x
75	Evaluación del "Programa Liderazgo Joven Construyendo Democracia" Julio 2010 – Diciembre 2012 (RE82)	
76	Mid-term Review of the Sida Supported Femmes Africa Solidarité (FAS) programme "Enhancing Civil Society in Human Security, Conflict Prevention and Peacekeeping" during the period 2010-2012 (RE83)	x
77	Mid-term Review of the Diakonia Strategic Peace Building Programme in Zimbabwe (RE84)	x
78	Evaluation of the Cooperation Strategy with Kenya 2009-2013 (RE85)	Ongoing
79	Review of the Results of Sweden's Development Cooperation Strategy in Georgia (RE86)	x
80	Outcome Assessment of Swedish Bilateral Cooperation in Bolivia 2009-2012 (RE87)	x
81	Sida-ICRC study on value for money in the humanitarian sector (RE88)	Ongoing
82	Evaluación de Medio Término del Programa de Desarrollo Agropecuario Sostenible PROAGRO - Bolivia (RE89)	
83	Review of the National Democratic Institute (NDI) project "Building Public Confidence in the Electoral Process in Georgia", 2009-2013 (RE90)	x
84	Review of the Core Support for The Eurasia Partnership Foundation (EPF), 2008-2013 (RE91)	x
85	Review of the Swedish Development Cooperation within the Breakaway Region of Abkhazia, Georgia, 2011-2013 (RE92)	x
86	Review of Sweden's support to the South African HIV programme during 2004-2013 (RE93)	Ongoing
87	Evaluation of Counselling Services Unit in Zimbabwe (RE94)	Ongoing
88	Review of the DDG Humanitarian Mine Action Programme in Afghanistan (RE95)	x
89	Mid-Term Review of the Trade Policy Training Centre in Africa, trapca (RE96)	x
90	Review of the DDG Humanitarian Mine Action Support to the National Strategy through Clearance and Enhanced Quality Project in Afghanistan (October 2010 – September 2013) (RE97)	Ongoing
91	Review of the Joint Climate Change Initiative project (RE98)	Ongoing
92	Översyn av utbildningsprogram inom Sidas uppdrag för resursbasutveckling (RE99)	Ongoing
93	Review of the Swedish civil Society Support in Liberia (RE100)	Ongoing
94	Review of Minor Field Studies (RE101)	Ongoing
95	Evaluation of the Swedish Strategies for Special Initiatives for Democratisation and Freedom of Expression (RE102)	Ongoing
96	Review of Reso Climat Mali (RE103)	Ongoing
97	Evaluation of Market Transformation Initiative (MTI) (RE104)	Ongoing

98	Evaluation of the international standard (ISO 26000) on Social Responsibility (RE105)	Ongoing
99	Evaluation of “UN Women’s program Truth, Justice and Reparation for Women in Colombia” (RE106)	Ongoing
100	Impact evaluation of Afghanistan Research and Evaluation Unit (AREU) and Afghanistan Analysts Network (AAN) (RE107)	Ongoing
101	Review of the Sida-funded Institutional Cooperation in the Field of Environment in Ukraine (RE108)	Ongoing
102	Evaluation of the Sida supported program of the International Association of Universities (RE109)	Ongoing
103	Final Evaluation of ERRC 2012-2013 (RE110)	Ongoing
SIPU Evaluations		
104	Evaluation of cooperation results under the Swedish Strategy for Development Cooperation with Turkey, 2010-- 2013	x
105	Evaluation of the contribution made by the Swedish Peace and Arbitration Society to the project “Human Rights and Democracy Promotion in Chechnya”	x
106	2011 Evaluation of ACCORD, SAIIA and CCR – South Africa	x
107	Evaluation of “Sustainable Livelihood Programme through Community Mobilization and Establishing Knowledge Resource Centre in Mazar-e-Sharif”	x
108	Evaluation of Sida’s support to Kvinna till Kvinna (KtK) and its programme: “Palestinian women seek greater power and influence to organise for democratic state building” 2011-2013	x
109	Informe Final de Evaluación Externa; Proyecto Fortalecimiento Democrático de las Organizaciones Políticas de Bolivia, PNUD – Embajada de Suecia	x
110	Evaluation of Support to Private and Decentralised Forestry in Kosovo, 2009-- 2013	x
111	Mid Term Review of the Malonda Program July 2010-June 2013	x
112	A Review of Sida’s Meeting Point Programmes with the Swedish Trade Council and the Swedish Geological Survey	x
113	Review/evaluation of RFSU’s regional programme: “Improving Lesbian, Gay, Bisexual Transgender (LGBT) rights and health (including HIV) in South Asia through strengthening civil society organisations”	x
114	Evaluation of the Swedish Support to the Tax Administration in Albania, Phase 2, 2008-2011	x
115	Evaluation of CENTEC (Centre for Environmental Technology) at the Embassy in Beijing	x
ORGUT Evaluations		
116	External Review (MTR) of the BABA CARAPA Community Forestry Programme, in benefit of the forest and its inhabitants, in Bolivia	x
117	Review Kvinna till Kvinna	x
118	Identifying and measuring results for sustainable systemic change in market development: A pre study for strategic evaluation of Sida’s market portfolio	
119	Evaluation of the Open Fun Football School Project 2010-2013, implemented by CCPA	
121	Advisory services for Results Framework to REC	

7.6 LIST OF PERSONS INTERVIEWED DURING FIELD PHASE

	Organisation	Name (First)	Name (Second)	Title/Role/ Partner Organisation
1	Sida (OPS)	Lennart	Peck	Evaluation Specialist; Focal person for review.
2	Sida (OPS)	Tomas	Bergenholtz	Heading Sida's working group on Results.
3	Sida (OPS)	Katarina	Kotoglou	Formerly at UTV, currently program officer at Sida's unit for cooperation with conflict countries
4	Sida (OPS)	Ingela	Juthberg	Program officer at PROGSAM, Stockholm
5	Sida (OPS)	Susanne	Mattsson	Has been stationed abroad for Sida but now works at UTV.
6	Sida (OPS)	Annika	Sandell	Former Manager of Framework Contract at Procurement Unit
7	Sida (OPS)	Annika	Nordin-Jayawardena	Deputy Director of Department for Program Cooperation (PROGSAM)
8	Sida (OPS)	Carin	Valtré	Manager of Framework Contract at Procurement Unit
9	Sida (OPS)	Joakim	Molander	Head of Development Cooperation in Rwanda. Former head of UTV, including when FA was established.
10	Sida (OPS)	Mattias	Jonsjö	Focal point at UTV for the framework agreement, up to June this year.
11	Sida (ADV)	Jessica	Pellrud	RES/Västra Balkan
12	Sida (ADV)	Anna	Tjärvar	Reform & Selektivt Samarbete
13	Sida (ADV)	Pezo	Mateo Phiri	Embassy Lusaka
14	Sida (ADV)	Andrei	Darie	Embassy in Moldova
15	Sida (ADV)	Jenny	Åkerbäck	B4D
16	Sida (ADV)	Anna	George	RES
17	Sida (ADV)	Maria	Myrman	RES
18	Sida (ADV)	Johan	Sundberg	INTEM
19	Sida (ADV)	Maria	Vink	Embassy in Kenya
20	Sida (ADV)	Fredrik	Uggla	Embassy in Egypt
21	Sida (ADV)	Lennart	Jemt	Sida Rwanda, Human rights and democracy senior PM
22	Sida (ADV)	Maja	Edfast	Sida HQ, previously field (Moz).
23	Sida (ADV)	Joergen	Schoenning	Sida HQ, previously BKK
24	Sida (ADV)	Elisabeth	Folkunger	Embassy Kenya
25	Sida (ADV)	Jorgen	Eriksson	Sida HQ
26	Sida (Eval & Adv)	Peeter	Kaaman	Georgia
27	Sida (Eval & Adv)	Charlotte	Lundqvist	Embassy in Georgia
28	Sida (Eval & Adv)	Rezarta	Katuci	Embassy in Albania
29	Sida (Eval & Adv)	Magnus	Carlquist	Embassy in Zimbabwe
30	Sida (Eval & Adv)	Stephen	Mwakifwamba	Embassy in Tanzania
31	Sida (Eval & Adv)	Katarina	Perrolf	Embassy in Kenya
32	Sida (Eval & Adv)	Désiré	Ballo	Embassy in Mali
33	Sida (Eval & Adv)	John	Nakedde	Sida, Embassy Uganda
34	Sida (Eval & Adv)	Jenny	Bjoerk	Sida HQ
35	Sida (Eval & Adv)	Kerstin	Gylhammar	Sida HQ

36	Sida (Eval & Adv)	Jonas P.	Bergstroem	Embassy Bogota, Colombia
37	Sida (Eval & Adv)	Kristina	Kuhnel	Embassy Cambodia
38	Sida (Eval & Adv)	Brigitte	Junker	Swedish Embassy, Addis, Ethiopia
39	Sida (EVAL)	Maria	Bjernevi	HUM
40	Sida (EVAL)	Patrik	Stålgren	Embassy in Moldova
41	Sida (EVAL)	Nedim	Bukvic	Embassy in Bosnia
42	Sida (EVAL)	Lisa	Mossberg	Global Cooperation /CIVSAM
43	Sida (EVAL)	Mario	Vignjevic	Embassy in Bosnia
44	Sida (EVAL)	Eva	Smedberg	Embassy in Georgia
45	Sida (EVAL)	Anna	Tufvesson	Embassy in Kenya
46	Sida (EVAL)	Kakha	Khimiashvili	Embassy in Tbilisi
47	Sida (EVAL)	Maria	Ljungman	Embassy in Kenya
48	Sida (EVAL)	Maher	Doudi	Embassy in Jerusalem
49	Sida (EVAL)	Jennifer	Matafu	Embassy Tanzania
50	Sida (EVAL)	Veronica	Melander	Embassy Guatemala
51	Sida (EVAL)	Carin	Zetterlund	Sida HQ, hum. Assistance
52	Sida (EVAL)	Karin	Olofsson	Sida HQ (formerly Embassy Colombia)
53	Sida (EVAL)	Nito	Matavel	Embassy Mozambique
54	Sida (EVAL)	Maria	Melbing	Embassy Kosovo
55	Sida (EVAL)	Olga	Sandakova	Sida, Embassy Ukraine
56	Sida (EVAL)	Karin	Faellman	Sida, Civil Society Unit, Senior Policy Specialist
57	Sida (EVAL)	Jonathan	Francis	Sida Sarajevo, BiH
58	Sida (EVAL)	Ravi	Behara	Sida New Delhi, India
59	Sida (EVAL)	David	Holmertz	Sida HQ, Dpt for Program Coordination
60	Sida (EVAL)	Michael	Otto	Sida HQ
61	Partner (ADV)	Bineta	Diop	FAS Femmes du Solidarite
62	Partner (ADV)	Gisella	Reina	ICJ - Int. Commission of Jurists
63	Partner (ADV)	Christine	Oram	Global Witness
64	Partner (ADV)	Lee	Wallis	Burn Diagnostics
65	Partner (ADV)	Kok-Thay	Truth	DCCam
66	Partner (ADV)	Carin	Karlsson	Energimyndigheten
67	Partner (ADV)	Deo	Kamweya	QUISP
68	Partner (ADV)	Lina	Aurell	Swedish chamber of trade
69	Partner (ADV)	Federico	Properzi	UN WATER
70	Partner (ADV)	Laura	McVeigh	Internationella PEN
71	Partner (ADV)	Maria	Åberg	ICLD
72	Partner (ADV)	Mohamed	Samir	AGADIR ATU
73	Partner (ADV)	Nand	Kishor Agrawal	ICIMOD; Neera Shresta Pradhan
74	Partner (ADV)	Linda	Nilsson	Riksförbundet Narkotikafritt Samhälle
75	Partner (ADV)	Malin	Grape	SMI
76	Partner (ADV)	Pravina	Makan-Lakha	ACCORD, South Africa

77	Partner (ADV)	Ebenizario	Conguica	OKACOM, Botswana
78	Partner (ADV)	Susanne	Kozak	GM and coordination expert, UN Women Zimbabwe
79	Partner (ADV)	Ylva	Reinhard	Naturvardsverket
80	Partner (ADV)	Kenneth	Hotz	Swedish Peace and Arbitration Society
81	Partner (ADV)	Beyant	KABWE	Save the Children, Tanzania
82	Partner (ADV)	Ebenizario	Chonguica	OKACOM
83	Partner (ADV)	Mads	Jörgen	Save the Children Zambia
84	Partner (ADV)	Håkan	Henning	ILAC
85	Partner (EVAL & ADV)	Coumba	Fall-Ven	FAS Femmes du Solidarite
86	Partner (EVAL & ADV)	Felipe	Atkins	Country Director, Norwegians People's Aid, Rwanda
87	Partner (Eval)	Graham	Haylor	The International Foundation For Science (Sweden)
88	Partner (Eval)	Hanna	Johnsson	Raoul Wallenberg Institute of Human Rights and Humanitarian Law
89	Partner (Eval)	Lars	Klint	Kronofogden
90	Partner (Eval)	Lena	Skiold	Arbetsformedlingen (Swedish Employment Service)
91	Partner (Eval)	Oley	Dibba-Wadda	FAS Femmes du Solidarite
92	Partner (Eval)	Peter K	Sjögren	Life & Peace Institute
93	Partner (Eval)	Minou	Fuglesang	Femina HIP (Tanzania)
94	Partner (Eval)	Keti	Gomelaury	Evaluation of promoting the integrity of civil data in Georgia
95	Partner (Eval)	Olena Galina	Pometun Gupan	Review of the Sida-funded Project Education for Sustainable Development in Action (ESDA)
96	Partner (Eval)	Carina	Svensson	SIS
97	Partner (Eval)	Karl	Magnusson	National Museums of World Heritage
98	Partner (Eval)	Anna	Axelsson	Evaluation of SSNC's sub-program "Climate Change" and sub-components "Cooperation with partners in the South"
99	Partner (Eval)	Carl	Mossberg	Evaluation of the Upland Research and Capacity Development Programme in Laos
100	Partner (Eval)	Eka	Msekhidze	National Agency of Public Registry, Georgia
101	Partner (Eval)	Todd	Benson	IFPRI, HQ
102	Partner (Eval)	Anne-Claire	Dufay	Dep. Representative, UNICEF BiH
103	Partner (Eval)	Lene	Rasmussen	Danish Demining Group, Amman
104	Partner (Eval)	David	Kalaba	TRAPCA, Arusha
105	Partner (Eval)	Claudia Jimena	Arenas Ferro	ForumSyd, Colombia
106	Partner (Eval)	Herman	Musahara	National University of Rwanda
107	Partner (Eval)	Tanja	Gohlert	EURO Burma
108	SP--Man	Anna	Liljelund Hedqvist	Project Manager, SP
109	SP--Man	Jessica	Rothman	Programme Manager, SP
110	SP--Man	Ian	Christoplos	Project Director, SP
111	SP--Man	Viktoria	Hildenvall	Business Area Coordinator Evaluation & PCM; SIPU
112	SP--Man	Anders	Olin	Managing Director SIPU
113	SP--Man	Christian	Carlbaum	Project Manager SIPU
114	SP--Man	Niklas	Herrmann	Deputy Managing Director, Orgut
115	SP--Man	Sandra	Mårtensson	Project Director, Orgut

116	SP--Core Team	Joakim	Anger	Core Team, SP
117	SP--Core Team	Sanne	Chipeta	Core Team, SP
118	SP--Core Team	Annika	Nilsson	Core Team, SP
119	SP--Core Team	Adam	Pain	Core Team, SP
120	SP--Core Team	Annica	Holmberg	Core Team, SP
121	SP--Core Team	Stefan	Dahlgren	Core Team, SP
122	SP--Core Team	Vera	Devine	Core Team, SP
123	SP--Core Team	Johanna	Lindström	Core Team, SIPU
124	SP--Core Team	Kim	Fors	Core Team, SIPU
125	SP--Core Team	Leif	Danielsson	Core Team, SIPU
126	SP--Core Team	Erik	Byrld	Core Team Member (TL) & Management (Tana)
127	SP--Core Team	Riccardo	Polastro	Core Team Member (TL) & Management
128	SP--Core Team	Pontus	Modeer	Conusltant, Right House
129	SP--Core Team	Jerome	Gouzou	Core Team Member (TL)
130	SP--Core Team	Jock	Baker	Core Team Member (TL)
131	SP--Core Team	Cecilia	Ljungman	Core Team Member (TL)
132	SP--Core Team	Ali	Gastgeer	Core Team Member (TL)
133	SP--Con	Bernt	Andersson	Consultant, SP
134	SP--Con	Krister	Eduards	Consultant, SP
135	SP--Con	Madeleine	Elmqvist	Consultant, Orgut
136	SP--Con	Trond	Norheim	Consultant, Orgut
137	Coalition Partner	Katy	Johnson	OPM
138	Coalition Partner	Maria	Bak	Channel Research
139	Coalition Partner	Brade	Johanson	Coalition Partner (SIPU)
140	Coalition Partner	Johan	Homberg	Coalition Partner (Orgut)
141	Coalition Partner	Martin	Schmidt	Coalition Partner (Orgut)
142	Coalition Partner	Krister	Eduards	Stockholm Group for Development Studies;
143	Other Consultant	Ann	Kampe	
144	Other Consultant	Cecilia	Karlstedt	

7.7 EVALUATIONS CONDUCTED UNDER FRAMEWORK AGREEMENT

	Indevelop Evaluations	Type	Unit/Embassy	Cost (Invoiced)	Team Leader	Gender
1	Evaluation of the of Co-operation with the Statistical office Kosovo (RE1)	Evaluation	PROGSAM /Unit for Western Balkans and Turkey	824,971	Martin Schmidt	M
2	Revu externe du Projet d'Appui au Système Statistique National du Mali 2009-2012 (RE2)	Review	Embassy in Mali	586,943	Jan Valdelin	M
3	Evaluation of the MENA-OECD Investment Programme 2008-2010 (RE4)	Evaluation	Conflict and post-conflict cooperation/MENA	649,597	Stefan Dahlgren	M
4	Evaluation of Swedish Health Sector Programme Support in Uganda 2000-2010 (RE6)	Evaluation	Dept. for Prog Cooperation/Health and social security support	1,163,621	Janet Gruber	F
5	8.2.8 Desk study of Diakonia (RE7)	Other	Global Cooperation /CIVSAM	368,600	Annica Holmberg	F
6	8.2.9 Evaluation of Swedish East African Music Network (RE8)	Evaluation	Programme Cooperation/REPS	273,025	Annica Holmberg	F
7	8.2.10 Evaluation of Cultural Heritage without Borders (RE9)	Evaluation	PROGSAM /Unit for Western Balkans and Turkey	509,996	Cecilia Ljunggren	F
8	8.2.11 Overview of International Training Programme "Management of Hydropower Development and Use" (RE10)	Other	Global Cooperation/Capacity Development	398,137	Alicia Månsson	F
9	Utveckling av Metod för Resultatinsamling inom International Utbildnings Program (ITP) (RE11)	Review	Global Cooperation/Capacity Development	96,000	Joakin Anger	M
10	8.2.12 External review of the Moz-SAKSS programme in Mozambique (RE12)	Review	Embassy in Mozambique	410,650	Ian Christoplos	M
11	8.2.13 Evaluation of Life and Peace Institute (RE13)	Evaluation	Programme Cooperation/REPS	665,707	Finn Skadkear Pedersen	M
12	8.2.14 Mid-term review of Sida support to the Centre for Science and Environment (RE14)	Mid-Term Review	Embassy in India	334,030	Adam Pain	M
13	8.2.15 Evaluation of Diakonia Zimbabwe Programme 2009-2011 (RE15)	Evaluation	Embassy in Zimbabwe	256,980	Jerome Gozu	M
14	8.2.16 Evaluation of the Upland Research and Capacity Development Programme in Laos (RE16)	Evaluation	Reform and Selective Cooperation/Asia	601,660	Dirk Van Gansberghe	M

15	8.2.17 Review of the Technical Barriers to Trade Mentorship Programme 2008-2011 (RE17)	Review	Programme Cooperation	193,253	Jens Andersson	M
16	8.2.18 Evaluation of Olof Palme International Centre (RE18)	Evaluation	Global Cooperation /CIVSAM	463,853	Martin Schmidt	M
17	8.2.19 Review of the Greater Mekong Sub-Region Regional Power Trade (RE19)	Review	Programme Cooperation /Asia	965,915	Juhani Antikainen	M
18	8.2.20 Review of civil society organisations in Moldova (RE20)	Review	Reform and Selective Cooperation/Eastern Europe	358,505	Klas Markensten	M
19	8.2.21 Mapping of Sida funded projects in area C, Jerusalem and Seam Zones through the Palestinian Authority, UN, Palestinian, Israeli and International NGO's (RE21)	Other	Embassy in Jerusalem	458,801	Trish Silkin	F
20	8.2.22 Evaluation of the Water Network Management Programme in Erbil, Iraqi Kurdistan (RE22)	Evaluation	Conflict and post-conflict cooperation/MENA	657,244	Bernt Andersson	M
21	8.2.23 Mid-Term Review and Evaluation of the Rural Electrification Programme in Zambia (RE23)	Mid-Term Review	Embassy in Zambia	858,176	Ali Dastgeer	M
22	8.2.24 Review of UNDP Georgia Good Governance project (RE24)	Review	Embassy in Tbilisi	224,802	Vera Devine	F
23	8.2.25 Review of Cultural Heritage for Sustainable Development, a Programme of Museum Cooperation in Southeast Asia (MuSEA) (RE25)	Review	Programme Cooperation /Asia	299,648	Stefan Dahlgren	M
24	8.2.26 Mid-Term Review of the Regional Human Rights Programme of the Ombudsman Institution (Defensoría del Pueblo) in Colombia (RE26)	Mid-Term Review	Embassy in Colombia	585,794	Vegard Bye	M
25	8.2.27 Mid-term Review of the Capacity Development Project of the Rural Energy Agency, Tanzania (RE27)	Mid-Term Review	Embassy in Tanzania	606,309	Leif Danielsson	M
26	8.2.28 Outcome Assessment and Lessons Learnt from Swedish Development Cooperation with Macedonia (1999–2012) (RE28)	Assessment	Embassy in Macedonia	771,772	Joakim Anger	M
27	8.2.29 Review of the EAC and MENA Carbon Footprint of Products Pilot Projects 2010–	Review	Programme Cooperation	939,211	Kimi Pedersen	F

	2011 (RE29)					
28	8.2.30 Review of FOS, the Fund for the Colombian Civil Society for Peace, Human Rights and Democracy (RE30)	Review	Embassy in Colombia	421,530	Merete Hansen	F
29	8.2.31 Evaluation of Sida's Support to the National University of Rwanda (NUR) 2007 – 2011 (RE31)	Evaluation	Embassy in Rwanda	1 116 518	Adam Pain	M
30	8.2.32 Interim Evaluation of Support to the Church of Sweden's Community Based Psychosocial Support in Humanitarian Assistance 2006–2011 (RE32)	Evaluation	Global Cooperation /CIVSAM	687,394	Sue Enfiled	F
31	8.2.33 Mid-term Review of the Bio-fuel project support to the Ministry of Energy and Minerals, Tanzania (RE33)	Mid-Term Review	Embassy in Tanzania	442,630	Klas Markensten	M
32	8.2.34 External Review of Core Support under Joint Financial Agreement to Zambia National Farmers Union (ZNFU) (RE34)	Review	Embassy in Uganda	561,070	Sanne Chipeta	M
33	8.2.35 Evaluation of World Children's Prize (RE35)	Evaluation	Global Cooperation /CIVSAM	638,744	Annica Holmberg	F
34	8.2.36 Review of the CREDO Herzegovina (RE37)	Review	Embassy in Bosnia	265,624	Andrea Spear	F
35	8.2.37 Evaluation of the Maghreb Regional Training Programme Development of Public Employment Services in Algeria, Morocco and Tunisia (RE38)	Evaluation	Conflict and post-conflict cooperation/MENA	553,747	Jups Kluyskens	M
36	8.2.38 Uncovering the Power of Performing Arts: An Evaluation of the Tamasi Programme (RE39)	Evaluation	Embassy in Cairo	683,045	Cecilia Ljunggren	F
37	8.2.39 Review of the Sida-funded Project Education for Sustainable Development in Action (ESDA) (RE40)	Review	Embassy in Ukraine	490,334	Vera Devine	F
38	8.2.40 Study of Sida's Support to the Swedish Civil Contingencies Agency (MSB) 2006-2011 (RE41)	Other	HUM	1,436,924	Jock Baker	M
39	8.2.41 Evaluation of the BetterAid and Open Forum Programmes (RE42)	Evaluation	Global Cooperation /CIVSAM	1,041,429	Angela Christie	F

40	8.2.42 Evaluation of Forum Syd and Dia-konia's Democracy and Human Rights programmes in Cambodia (RE43)	Evaluation	Embassy in Cambodia	499,683	Erik Bryld	M
41	8.2.43 Evaluation of the Barbro Johansson Model Girls' Secondary School in Tanzania (RE44)	Evaluation	Embassy in Tanzania	564,569	Ali Dastgeer	M
42	8.2.44 Evaluation of Legal and Human Rights Centre (LRHC) in Tanzania (RE46)	Evaluation	Embassy in Tanzania	662,031	Stefan Dahlgren	M
43	8.2.45 Resilience, Risk and Vulnerability at Sida (RE47)	Evaluation	International Organizations and Policy Support	496,751	Ian Christoplos	M
44	8.2.46 Review of the project Capacity Building & Improved Client Services at the National Agency of Public Registry (NAPR) in Georgia (RE48)	Review	Embassy in Georgia	227,607	Jan Eriksson	M
45	8.2.47 Evaluation of Promoting the Integrity of Civil Data in Georgia (RE49)	Evaluation	Embassy in Georgia	223,579	Krister Eduards	M
46	8.2.48 Review of the project Capacity Building of the Georgian Leadership Community for Improved Decision-making and Negotiation Skills (CBGL) (RE50)	Review	Embassy in Georgia	220,071	Bernt Andersson	M
47	8.2.49 Evaluation of the Raoul Wallenberg Institute's regional programme "Building Human Rights Knowledge and Resources in the Middle East and North Africa" (RE51)	Evaluation	Embassy in Egypt	851,701	Annika Nilsson	F
48	8.2.50 Review of the partnership cooperation between the National Bureau of Enforcement of Georgia and the Swedish Enforcement Agency (RE52)	Review	Department of Reform and Selective Cooperation	339,987	Klas Markensten	M
49	Civil Society Support Modalities at Sida Headquarters and Swedish Embassies (RE53)	Review	CIVSAM	232,487	Annika Nilsson	F
50	8.2.51 Evaluation of the Sida and DFID funded Public Policy Information, Monitoring and Advocacy (PPIMA) project in Rwanda (RE54)	Evaluation	Embassy in Rwanda	1,099,133	Ali Dastgeer	M
51	8.2.52 Evaluation of the Project "Protection of Children at Risk and	Evaluation	Embassy in Bosnia	264,055	Vera Devine	F

	Children in Contact with the Justice System in Bosnia and Herzegovina” (RE55)					
52	Évaluation Externe du Plan Stratégique de Promotion de la Gouvernance Démocratique au Burkina Faso (RE56)	Evaluation	Embassy in Burkina Faso	462,443	Jerome Gouzou	M
53	8.2.53 Civil Society on Climate Change: Evaluation of the work on climate change adaptation and mitigation of Swedish Society for Nature Conservation in cooperation with partner organisations in the South (RE57)	Evaluation	CIVSAM	657,750	Annica Holmberg	F
54	8.2.54 External Evaluation of the Partnership Agreement for Sustainable Development of Lake Victoria Basin (RE58)	Evaluation	Embassy in Kenya	573,522	Gunilla Göransson	F
55	8.2.55 Evaluation of the International Foundation for Science (RE59)	Evaluation	FORSK	298,558	Ian Christoplos	M
56	8.2.56 Midterm Review of the ‘Capacity Development of Municipal Associations’ Programme (CDMA) (RE60)	Mid-Term Review	Embassy in Bosnia	319,009	Andrea Spear	F
57	Metodstöd vid genomförande av resultatseminarium för ITP 268 LGBT and Human Rights (RE61)	Review	KAPSAM	-	Joakim Anger	M
58	Summative Evaluation of phase out in East Timor (RE62)	Evaluation	Embassy in East Timor	-	Trish Silkin	F
59	8.2.57 Evaluation of Save the Children Sweden’s Support to Organisations (RE64)	Evaluation	CIVSAM	689,285	Annica Holmberg	F
60	8.2.58 Evaluation of Sida’s Support to Environment Infrastructure and Reforms in Central and Eastern Europe and Western Balkans 1995-2010 – A Desk Study (RE65)	Evaluation	Embassy in Moldova	921,782	Eric Buhl-Nielsen	M
61	8.2.59 Mid-Term Review for the Quality Infrastructure and Standards Programme (QUISP) in Uganda (RE66)	Mid-Term Review	Embassy in Uganda	654,928	Jens Andersson	M
62	8.2.60 Mid-Term Review of the AGIR Programme (RE67)	Mid-Term Review	Embassy in Mozambique	809,063	Kevin Kelpin	M
63	8.2.61 Review of the Sida-funded Project “Advisory Support to the Ministry of Finance of	Review	Embassy in Ukraine	792,650	Vera Devine	F

	Ukraine” (RE68)					
64	8.2.62 Reform cooperation in the Western Balkans – regional cooperation: experiences, constraints and opportunities (RE69)	Other	Western Balkins	401,606	Joakim Anger	M
65	8.2.63 Study on Results of Cooperation under the Swedish Strategy for Development Cooperation with Kosovo, 2009-2012 (RE71)	Other	Embassy in Kosovo	379,032	Jim Newkirk	M
66	8.2.64 Strategic review and advice on regional comparative advantages in Eastern Europe and the Eastern Partnership (RE72)	Other	Eastern Europe	521,651	Krister Eduards	M
67	Evaluation of Sida & NIR Core Support Program (2009 – 2012) (RE73)	Evaluation	B4D	287,216	Erik Bryld	M
68	8.2.65 Review of Swedish Support to the World Customs Organization (WCO) Capacity Building, 2008-2012 (RE74)	Review	PROGSAM	232,020	Andrea Spear	F
69	8.2.66 Mid-term Review of the Institutional Strengthening Support to the Association of Local Authorities of Tanzania - ALAT (RE75)	Mid-Term Review	Embassy in Tanzania	513,444	Joakim Anger	M
70	8.2.67 Evaluation of Femina Health Information Project (HIP) Strategic Plan 2006-2012 (RE77)	Evaluation	Embassy in Tanzania	622,574	Sanne Chipeta	F
71	8.2.68 Review of the Swedish cooperation strategy with the Russian Federation in the areas of human rights and democracy (RE78)	Review	Russia	308,785	Vera Devine	F
72	8.2.69 Evaluation of the Institutional Cooperation Between Ministry of Lands, Kenya and Lantmäteriet, Sweden on the Project for Improving Land Administration in Kenya 2009-2012 (PIL-AK) (RE79)	Evaluation	Embassy in Kenya	459,709	Bernt Andersson	M
73	Monitoring of Rwanda Peacebuilding Programme (RE80)	Other	Embassy in Rwanda	271,935	Kevil Kelpin	M
74	8.2.70 Evaluation of Results of Sweden's Development Cooperation Strategy Albania (RE81)	Evaluation	Embassy in Albania	562,841	Ian Christoplos	M

75	Evaluación del “Programa Liderazgo Joven Construyendo Democracia” Julio 2010 – Diciembre 2012 (RE82)	Evaluation	Embassy in Guatemala	-	Annica Holmberg	F
76	8.2.71 Mid-term Review of the Sida Supported Femmes Africa Solidarité (FAS) programme “Enhancing Civil Society in Human Security, Conflict Prevention and Peace-keeping” during the period 2010-2012 (RE83)	Mid-Term Review	PROGSAM	335,855	Jerome Gouzou	M
77	8.2.72 Mid-term Review of the Diakonia Strategic Peace Building Programme in Zimbabwe (RE84)	Mid-Term Review	Embassy in Zimbabwe	259,998	Jerome Gouzou	M
78	Evaluation of the Cooperation Strategy with Kenya 2009-2013 (RE85)	Evaluation	Embassy in Kenya	359,627	Angela Christie	F
79	8.2.73 Review of the Results of Sweden’s Development Cooperation Strategy in Georgia (RE86)	Review	Embassy in Georgia	150,000	Vera Devine	F
80	8.2.74 Outcome Assessment of Swedish Bilateral Cooperation in Bolivia 2009-2012 (RE87)	Assessment	Embassy in Bolivia	-	Vegard Bye	M
81	Sida-ICRC study on value for money in the humanitarian sector (RE88)	Other	HUM	359,893	Jock Baker	M
82	Evaluación de Medio Término del Programa de Desarrollo Agropecuario Sustentable PROAGRO - Bolivia (RE89)	Review	Embassy in Bolivia	-	Jokob Kronik	M
83	Review of the National Democratic Institute (NDI) project “Building Public Confidence in the Electoral Process in Georgia”, 2009-2013 (RE90)	Review	Embassy in Georgia	95,869	Krister Eduards	M
84	Review of the Core Support for The Eurasia Partnership Foundation (EPF), 2008-2013 (RE91)	Review	Embassy in Georgia	121,190	Jim Newkirk	M
85	Review of the Swedish Development Cooperation within the Breakaway Region of Abkhazia, Georgia, 2011-2013 (RE92)	Review	Embassy in Georgia	135,431	Vera Devine	F
86	Review of Sweden's support to the South African HIV programme during 2004-2013 (RE93)	Review	Embassy in South Africa		Annika Nilsson	F

87	Evaluation of Counseling Services Unit in Zimbabwe (RE94)	Evaluation	Embassy in Zimbabwe		Jerome Gouzou	M
88	Review of the DDG Humanitarian Mine Action Programme in Afghanistan (RE95)	Review	Sida/KONFLIKT		Jock Baker	M
89	Mid-Term Review of the Trade Policy Training Centre in Africa, trapca (RE96)	Mid-Term Review			Bernt Andersson	M
90	Review of the DDG Humanitarian Mine Action Support to the National Strategy through Clearance and Enhanced Quality Project in Afghanistan (October 2010 – September 2013) (RE97)	Review			Adam Pain	M
91	Review of the Joint Climate Change Initiative project (RE98)	Review			Stefan Dahlgren	M
92	Översyn av utbildningsprogram inom Sidas uppdrag för resursbasutveckling (RE99)	Review			Joakim Anger	M
93	Review of the Swedish civil Society Support in Liberia (RE100)	Review			Annica Holmberg	F
94	Review of Minor Field Studies (RE101)	Review			Joakim Anger	M
95	Evaluation of the Swedish Strategies for Special Initiatives for Democratisation and Freedom of Expression (RE102)	Evaluation			Cecilia Ljungman	F
96	Review of Reso Climat Mali (RE103)	Review			Jerome Gouzou	M
97	Evaluation of Market Transformation Initiative (MTI) (RE104)	Evaluation			Sanne Chipeta	F
98	Evaluation of the international standard (ISO 26000) on Social Responsibility (RE105)	Evaluation			Bernt Andersson	M
99	Evaluation of “UN Women’s program Truth, Justice and Reparation for Women in Colombia” (RE106)	Evaluation			Kimberly Inksater	F
100	Impact evaluation of Afghanistan Research and Evaluation Unit (AREU) and Afghanistan Analysts Network (AAN) (RE107)	Evaluation			Erik Bryl	M
101	Review of the Sida-funded Institutional Cooperation in the Field of Environment in Ukraine (RE108)	Review			Eric Buhl-Nielsen	M
102	Evaluation of the Sida supported program of the International Association	Evaluation			Ian Christoplos	M

	of Universities (RE109)					
103	Final Evaluation of ERRC 2012-2013 (RE110)	Evaluation			N/A	
	SIPU Evaluations					
104	Evaluation of cooperation results under the Swedish Strategy for Development Cooperation with Turkey, 2010--2013	Evaluation	Embassy in Istanbul	592,995	Roland Blomeyer	M
105	Evaluation of the contribution made by the Swedish Peace and Arbitration Society to the project "Human Rights and Democracy Promotion in Chechnya"	Evaluation	Eastern Europe Unit, Sida	168,800	Roger Hållhag	M
106	Evaluation of the African Centre for the Constructive Resolution of Disputes (ACCORD), the South African Institute of International Affairs (SAIIA) and the Centre for Conflict Resolution, CCR	Evaluation	Team for Empowerment, Peace and Security in Africa	552,500	Niels Eilschow Olesen	M
107	Evaluation of "Sustainable Livelihood Programme through Community Mobilization and Establishing Knowledge Resource Centre in Mazar-e-Sharif"	Evaluation	Conflict and post-conflict cooperation/MENA	450,000	Sarah Gray	F
108	Evaluation of Sida's support to Kvinna till Kvinna (KtK) and its programme: "Palestinian women seek greater power and influence to organise for democratic state building" 2011-2013	Review	Embassy in Palestine	326,764	Henrik Brade Johansen	M
109	Informe Final de Evaluación Externa; Proyecto Fortalecimiento Democrático de las Organizaciones Políticas de Bolivia, PNUD – Embajada de Suecia	Evaluation	Embassy in Bolivia	565,500	Francesca Jessup	F
110	Evaluation of Support to Private and Decentralised Forestry in Kosovo, 2009-- 2013	Evaluation	NA	363,520	Åke Nilsson	M
111	Mid Term Review of the Malonda Program July 2010-June 2013	Mid-Term Review	Embassy in Mozambique	828,400	Andrew Lambert	M
112	A Review of Sida's Meeting Point Programmes with the Swedish Trade Council and the Swedish Geological Survey	Review	NA	242,000	Karlis Goppers	M

113	Review/evaluation of RFSU's regional programme "Improving Lesbian, Gay, Bisexual, Transgender (LGBT) rights and health (including HIV) in South Asia through strengthening civil society organisations" (Nepal, India, Bangladesh, Sri Lanka, Pakistan)	Evaluation	NA	844,720	Paul Balogun	M
114	Evaluation of the Swedish Support to the Tax Administration in Albania, Phase 2, 2008-2011	Evaluation	NA	454,080	Allan Gustavsson	M
115	Evaluation of CENTEC (Centre for Environmental Technology) at the Embassy in Beijing	Evaluation	RES/Selektivt samarbete	227,637	Martin Froberg	M
116	Evaluation of the Liberia Dujár Association	Evaluation	Liberia/Sierra Leone Unit	249,600	Leif Danielsson	M
117	Review of Sida's Program for Development research (Uforsk)	Review	U-forsk	400,000	Mats Hårsmar	M
118	Utvärdering av Sidas arbetsplan "Mänskliga rättigheter för personer med funktionshinder"	Evalaution	INTEM/Tema	275,000	Ulrika Ribohn	F
119	Mid Term Review of the MENA-OECD Investment Programme	Mid-Term Review	N.A	647,458	Ali Dastgeer	M
120	Experiences and lessons learned from partner Driven Cooperation in the seven selective cooperation countries	Other	RES/Selektivt samarbete	1,198,720	Klas Markensten	M
121	Evaluación de Medio Término MTR del Proyecto: "NODO de saneamiento Sostenible Descentralizado como Plataforma de Conocimiento y Generación de Impacto en Soluciones Sostenibles NSSD 2012-2015"	Mid-Term Review	Embassy in Bolivia	376,880	Thomas Alveteg	M
122	Review of the Results of the Support to the Centre for Science and Environment (CSE) and Toxics Link (TL)	Review	Embassy in India	410,522	Thomas Alveteg	M
ORGUT Evaluations						
123	External Review (MTR) of the BABA CARAPA Community Forestry Programme, in benefit of the forest and its inhabitants, in Bolivia	Review	Embassy in Bolivia	282,200	Trond Norheim	M
124	Review Kvinna till Kvinna	Review	Embassy in Georgia	299,332	Madeleine Elmqvist	F

125	Identifying and measuring results for sustainable systemic change in market development: A pre study for strategic evaluation of Sida's market portfolio	Evaluation	Sida	600,000	Johan Holmberg	M
126	Evaluation of the Open Fun Football School Project 2010-2013, implemented by CCPA	Evaluation	Sida		Madeleine Elmqvist	F

7.8 QUALITATIVE DATA INTERVIEW PROTOCOLS

These protocols provide standard questions for each cohort. These may be adapted for specific persons. Each interview allows for additional comments or information that respondents may deem important.

All protocols include a control question as an opening and select protocols have additional controls to gauge behaviours that may be influencing respondents at the time of the interview.

Each question includes a reference to the Evaluation Question from which it is derived.

Overview

Sida is currently conducting a Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks.

This Mid Term Review focuses on the operationalisation of the Framework Agreement, how and if services have improved, and the overall functionality of the Framework Agreement Model, amongst other issues.

This interview is designed to capture data that can be analysed to determine relevant evidentiary trends. As such, no single person's input may dominate the analysis while ensuring that all inputs are taken *prima facie*.

There are standard questions for all respondents and then specialized questions per cohort. We hope to capture 2 – 3 points per question, reciting these verbatim as possible. These will then be compiled in a database for broader analysis.

Obviously, we don't want to prevent any respondent from raising issues that they deem important and that may not be directly related to the questions in this protocol. This is expected and these points will be collected and analysed as well.

*Any questions or comments before we begin?*⁷⁴

For Sida operational (UTV; UPPIS) units:

1. Please describe your role in relation to the Framework Agreement?
-

⁷⁴ This is a control question. If unsolicited comments are overtly negative or positive, make note of this. If they are not or if they are relatively neutral overall, mark these responses as "Neutral".

2. What do you feel are the most critical things to consider about the Framework Agreement?⁷⁵
3. What were the challenges and/or opportunities associated with making the Framework Agreement operational? (A1.1)
4. What advantages and disadvantages do you see in relation to the Framework Agreement? (A2.2.3; A3.1; B1.1)
5. What examples do you have of dialogue and consultation between the service providers and UTV/UPPIS/Sida that was critical to a common understanding of the Framework Agreement? That detracted from this common understanding? (A1.1.4)
6. Has the Framework Agreement been communicated across Sida to support a common understanding of its purpose, services, and functionality? If so, please describe how. (A1.1.13)
7. Can the quality and performance of services provided under the Framework Agreement be improved? If so, how? (A2.2.2)
8. Do you have specific examples of how the Framework Agreement has reduced direct and/or indirect costs? (A3.1.1)
9. Does the Framework Agreement contribute to Sida's institutional learning on evaluation? If so, please describe. (A3.1.8)

For Sida Commissioners (Evaluations):

1. Please describe the evaluations you have commissioned under the Framework Agreement.
2. What do you feel are the most critical things to consider about the Framework Agreement?⁷⁶
3. What advantages and disadvantages have you experienced in relation to the Framework Agreement? (A2.3.3; A3.1)
4. Please describe how you worked with the consultants on these assignments? What worked, what didn't, what could be different? (A2.2.2)
5. Please describe the final report or other outputs. What worked, what didn't, and what could have been different? (A2.2.2)
6. How do service levels, e.g. accessibility, flexibility, timeliness, readiness for dialogue and solving problems in a constructive way, in these evaluations as compared with other evaluations? (A2.2.4)

⁷⁵ This is a control question designed to gauge current behaviours, attitudes, or events that may be unduly influencing a respondent.

⁷⁶ Ibid.

7. What types of guidance and support have you received, e.g. from UTV? (A1.1.10)
8. Overall, how would you rate the consultant's work? Not effective? Effective? Very Effective?
9. What else should we know about how evaluations are commissioned and implemented?

For Sida Commissioners (Advisory Services):

1. Please describe the advisory services you have commissioned under the Framework Agreement.
2. What advantages and disadvantages have you experienced in relation to the Framework Agreement? (A2.3.3; A3.1)
3. Please describe how you worked with the consultants on these assignments? What worked, what didn't, what could be different? (B2.1.1)
4. What could be done to improve the advisory services overall?⁷⁷ Please provide some examples. (B2.1)
5. Have services contributed to competence building and a better overall understanding of RBM for partners? If so, please describe. (B2.1.3)
6. How do service levels in these advisory services compare with other consultancy services you have experienced? (B2.1.4)

For Service Partners (Management):

1. What are the advantages and disadvantages of working with the Framework Agreement? (A3.1)
2. What were the challenges and/or opportunities associated with making the Framework Agreement operational? (A1.1)
3. What's different, if anything, about how relevant Sida partners are working with results frameworks/RBM as related to the advisory services? For instance, do you have examples how their work with results frameworks has improved or worsened? (B2.1.3)
4. What examples do you have of dialogue and consultation between the service providers and UTV/UPPIS/Sida that was critical to a common understanding of the Framework Agreement? That detracted from this common understanding? (A1.1.4)

⁷⁷ Purposefully open-ended. Should solicit a range of possible responses.

5. What has changed about how you work with Sida during the course of the Framework Agreement? Can you provide examples of what's worked, what hasn't, and what could be different? (A3.1.8)
6. How do you ensure quality overall? What has worked best, least and what could be different going forward? (A2.2)

For Service Partners (Core Teams):

1. Please describe briefly the evaluations/advisory services in which you acted as team leader. How many assignments have you done under the Framework Agreement?
2. How have you approached evaluation assignments differently as a Core Team member, if so how and why? (A2.2)
3. What type of support do you receive from the Service Partner and how does this impact our work? (A2.2)
4. How would you describe the collaboration with Sida overall?⁷⁸ (A2.2)
5. What are the most rewarding and frustrating aspects of the evaluation process? (A2.2.2)
6. What is done to ensure the quality and performance of evaluations?⁷⁹ (A2.2.1)
7. What do you wish could be different about how your work on the Core Team?⁸⁰
8. What are 3 - 5 ways that Core Team consultants have become better at evaluations and/or advisory services? (A2.2.5)
9. From your perspective, how does a Framework Agreement like this impact the consultancy market in Sweden? (A3.1.5; A3.1.6)

For Service Partners (Consultants):

1. How many assignments have you had under the Framework Agreement?
2. What are the 3 most important things you do as a consultant working under this Framework Agreement?⁸¹ (A2.1)
3. What are the most rewarding and frustrating aspects of the evaluation process? (A2.2.2)
4. What do you do to ensure that the process and reporting of an evaluation maximizes potential impact and results? (A2.2.2)⁸²

⁷⁸ This is purposefully open-ended and will hopefully elicit a variety of responses.

⁷⁹ This is purposefully positively assumptive: they should be able to list a range of elements and these should match those of the Service partner management team.

⁸⁰ This is a control question designed to gauge current behaviours, attitudes, or events that may be unduly influencing a respondent.

⁸¹ This seeks a response related to utilization.

5. What else should we know about your role on evaluations in relation to the Framework Agreement?

For Partner Organisations (Evaluations & Advisory):

1. Please describe the services (evaluations, advisory, or both) that you have engaged with in relation to the Framework Agreement. (May need to list specific engagements.)
2. Please describe how you worked with the consultants on these assignments? What worked, what didn't, what could be different? (A2.2.2)
3. Please describe the final report or other outputs. What worked, what didn't, and what could have been different? (A2.2.2)
4. What have you done differently, if anything, as a result of the evaluation and/or advisory services? (A2.2.3; B2.1.3)
5. How do service levels, e.g. accessibility, flexibility, timeliness, readiness for dialogue and solving problems in a constructive way, in these evaluations as compared with other evaluations? (A2.2.4)

What else should we consider in relation to this Review?

Do you have any questions for me?

Thank you.

⁸² This is purposefully positively assumptive: they should be able to list a range of elements and these should match those of the Service partner management team.

7.9 EVALUATION QUESTIONS

A. Evaluation Component	
A1	Operationalisation of the Framework Agreement
A1.1	How has the Framework Agreement been put to practice and used, and what change can be seen in compare to before the Framework Agreement?
A1.1.1	How many evaluations have been conducted within the Framework Agreement (total and per consultant) and how many have they turned down?
A1.1.2	What kinds of evaluations have been provided? (Thematic area, department, embassy, approach/methodology, etc.)
A1.1.3	To what extent has the Framework Agreement been used for commissioning evaluations by Sida Units/Embassies?
A1.1.4	What are the primary attributes of the collaboration between Sida and the Service Partners and how has this impacted the way the FA was implemented and how quality services are assured?
A1.1.5	What have been the costs of the evaluations in the Framework Agreement?
A1.1.6	What have the teams looked like in terms of gender, Swedish/national/international team members?
A1.1.7	Are consultant competencies (experience, expertise, aptitude, attitude) appropriate and adequate for specific evaluations?
A1.1.8	How has Sida organized itself to manage and utilize the Framework Agreement, including legal and administrative matters, support and communication? Has it been adequate and appropriate?
A1.1.9	How has Sida organised itself to support the knowledge management and institutional learning that should arise?
A1.1.10	How has Sida organised itself to provide guidance and support for using the Framework Agreement?
A1.1.11	How has Sida organised itself to develop strategic analysis and trends as based on how the Framework Agreement is used?
A1.1.12	Are Sida units positioned (legal, finance and administration; business systems/technology) to ensure that expected efficiency gains/cost savings are realised?
A1.1.13	Has the Framework Agreement been communicated in ways across Sida to support a common understanding of its purpose, services, and functionality?
A2	Quality of Services
A2.1	What is the overall quality of evaluation services under the Framework Agreement?
A2.2	Has the evaluation quality changed since the signing of the Framework Agreement?
A2.2.1	Can changes in the quality of evaluations be attributed to the Framework Agreement?
A2.2.2	Have the way in which Evaluations are commissioned under the Framework Agreement changed engagement/utilization levels?
A2.2.3	Have the way in which Evaluations are commissioned under the Framework Agreement changed the way they are used by Sida and its partner organisations (including both instrumental and process use)?
A2.2.4	To what extent have the consultants/Service Partners been service oriented in terms of e.g., accessibility, flexibility, timeliness, readiness for dialogue and solving problems in a constructive way?
A2.2.5	Have the consultants improved their evaluation competence and/or strengthened and accumulated knowledge about Sida to offer better services?
A2.2.6	Is the Framework Agreement supported by a standardized/common approach to the quality assurance of evaluations?
A3	Model
A3.1	Is the current model of the Framework Agreement functional?
	Overall, is there evidence that the Framework Agreement has improved evaluation quality?
A3.1.1	How have the costs for evaluation services changed under the Framework Agreement?
A3.1.2	How has the Framework Agreement impacted transaction costs/efficiency overall?

A3.1.3	Has the Framework Agreement changed the commissioning of evaluation services, including better-formulated ToR, and if so, to what extent and in what ways?
A3.1.4	What possible incentives (for the consultant and for Sida) does the Framework Agreement give rise to?
A3.1.5	What does the Framework Agreement imply for the long-term availability and quality of evaluation services to Sida, taking the resource base as a whole into consideration?
A3.1.6	What are the implications of the Framework Agreement for local evaluation consultants in partner countries? Could the Framework Agreement be improved to better contribute to enhanced development of capacities of evaluators in partner countries?
A3.1.7	What has the Framework Agreement implied for young evaluators to be engaged in the provision of evaluation services to Sida?
A3.1.8	To what extent does the Framework Agreement contribute to Sida's institutional learning on evaluation?
A3.1.9	What are the advantages and disadvantages of this Framework Agreement when compared to other relevant procurement models?
A3.1.10	What would be the pros and cons of not having a separate framework agreement for evaluation and instead let evaluations be one of the services included in Sida's different thematic framework agreements?
B.	Advisory Services on Results Framework Component
B1	Operationalization of the Advisory Services
B1.1	How has the Framework Agreement been put to practice and used in relation to the Results Framework component?
B1.1.1	How many advisory assignments have been conducted within the Framework Agreement and how many have they turned down?
B1.1.2	What have the advisory services consisted in?
B1.1.3	To what extent has the Framework Agreement been used by different Sida units/Embassies?
B1.1.4	How does the dialogue and consultation between the service providers and UTV/UPPIS/Sida effect a common understanding of how to use/apply the Framework Agreement for advisory services?
B1.1.5	Duration and costs of services (per assignment and total)?
B1.1.6	How have the consultants organized themselves to respond to specific advisory service engagement, including providing qualified evaluation teams, Core Team and quality assurance?
B1.1.7	Are consultant competencies (experience, expertise, aptitude, attitude) appropriate and adequate for specific advisory services?
B1.1.8	Have Sida's structures for managing the Framework Agreement for Advisory Services been appropriate and adequate, including quality assurance, learning, legal matters, support and communication?
B2	Quality of Services
B2.1	Has the quality of results frameworks and partner competence on results frameworks and RBM improved with the Framework Agreement?
B2.1.1	To what extent have the consultants delivered relevant and tailored advisory services for results frameworks?
B2.1.2	To what extent has the Framework Agreement improved quality on Sida's partner organisations results frameworks?
B2.1.3	To what extent have the services contributed to competence building and better overall understanding and outlook to RBM for partners? Including HRBA, perspectives of the poor, emphasis on gender equality etc. in partners' results frameworks.
B2.1.4	To what extent have the consultants been service oriented in terms of e.g., accessibility, flexibility, timeliness, readiness for dialogue and solving problems in a constructive way?
B3	Model
B3.1	Is the current model of the Framework Agreement appropriate and effective?

7.10 BIBLIOGRAPHY

The following list includes all materials, articles and books referenced as part of this Report.

- “Aid Effectiveness and the Paris Declaration: A Mismatch Between Ownership and Results Based Management?” Martin Sjöstedt *Public Administration and Development*, Vol. 33; May 2013.
- “A ‘post-aid world’? Paradigm shift in foreign aid and development cooperation at the 2011 Busan High Level Forum,” Emma Mawdsley, Laura Savage, and Sung-Mi Kim. *The Geographical Journal*, January 2013.
- “Are Sida Evaluations Good Enough? An Assessment of 34 Evaluation Reports,” Kim Forss, Evert Vedung, Stein Erik Kruse, Agnes Mwaiselage, and Anna Nilsson. *Sida*, 2008.
- *Choices, Values and Frames*, Daniel Kahneman & Amos Tversky. Cambridge University Press, 2000.
- *Competitive Advantage*, Michael E. Porter. Free Press, 1985.
- “Competitive Contracting for Public Services: A Comparison of Policies and Implementation in Denmark and Sweden,” Karin Bryntse and Carsten Greve. *International Public Management Review*, Vol. 3, 2002.
- “Depicting the Logic of Three Evaluation Theories,” Mark Hansen, Marvin C. Alkin, and Tanner LeBaron Wallace. *Evaluation and Program Planning*, Elsevier, March 2012.
- “Evaluating Complex Social Interventions in a Complex World,” David Byrne. *Evaluation*, Vol, 19, No. 3, July 2013.
- “Evaluating Development Activities: 12 Lessons from the OECD DAC,” OECD, 2013.
- “Evaluation in the Nordic Countries,” Gunilla Holm. *Journal of MultiDisciplinary Evaluation*, 2007.
- “Evidence-based Evaluation of Development Cooperation: Possible? Feasible? Desirable?,” Kim Forss and Sara Bandstein. *Network of Networks on Impact Evaluation (NONIE)*, World Bank, January 2008.
- “Final Annual Report 2012 – 2013,” Indevalop. June 2013.
- “A General Inductive Approach for Analyzing Qualitative Evaluation Data,” David R. Thomas. *American Journal of Evaluation*, Vol. 27. No. 2, June 2006.

- “Harnessing the Power of Evaluation in Humanitarian Action: An Initiative to Improve Understanding and Use of Evaluation,” Alistair Hallam. ALNAP, June 2011.
- *How to Use Qualitative Methods in Evaluation*, Michael Quinn Patton. SAGE Publications, 1987.
- “Influencing Change: Building Evaluation Capacity to Strengthen Governance,” Ray C. Rist, Marie-Helene Boily, & Federic Martin. The International Bank for Reconstruction and Development (IBRD)/The World Bank, 2011.
- *Looking Back, Moving Forward: Sida Evaluation Manual*, Stefan Molund and Goran Schill. 2nd revised edition. Sida, 2007.
- “Management Consultants: Client and Consultant Perspectives,” Flemming Poulfelt & Adrian Paynee. *Scandinavian Journal of Management*, Vol.10. No. 4, December 1994.
- “Measurement of Productivity in Professional Services: An Illustration on Swedish Management Consulting Firms,” Lilach Nachum. *International Journal of Operations & Production Management*, Vol. 19, No. 9, 1999.
- “Monitoring & Evaluation: Some Tools, Methods & Approaches.” The World Bank, 2004.
- “The OECD/DAC Criteria for International Development Evaluations: An Assessment and Ideas for Improvement,” Thomas Chianca. *Journal of MultiDisciplinary Evaluation*, Vol. 5. No. 9, March 2008.
- *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, Abhijit V. Banerjee and Esther Duflo. Public Affairs, 2011.
- *Predictably Irrational: The Hidden Forces That Shape Our Decisions*, Dan Ariely. HarperCollins, 2008.
- “Public Procurement of Consulting Services: Evidence and Comparison with Private Companies,” Filip Roodhooft & Alexandra Van den Abbeele. *International Journal of Public Sector Management*. Vol. 19, 2006.
- “Public Procurement in Europe: Cost and Effectiveness.” Price Waterhouse Coopers (PWC). European Commission, March 2011.
- “Quality Standards for Development Evaluation.” DAC Guidelines and Reference Series, OECD, 2010.
- “Sida Evaluations and Audits 2007: A Synthesis of Findings and Conclusions,” Lennart Peck & Verena Knippel. UTV Working Paper 2008:1.
- “Sustainable Development, Evaluation And Policy-Making: Theory, Practise and Quality Assurance,” Anneke von Raggamby & Frieder Rubik. Institute for Ecological Economy Research (IÖW), Germany, 2012.

- “Sourcebook for Evaluating Global and Regional Partnership Programs: Indicative Principles & Standards.” The Independent Evaluation Group—The World Bank, 2007.
- “Stakeholder-Agency Theory,” Charles W. Hill and Thomas M. Jones. *Journal of Management Studies*, Vol. 29, May 2007.
- “Swedish Development Cooperation in Transition: Lessons and Reflections from 71 Sida Decentralized Evaluations (April 2011 – April 2013),” Ian Christoplos, Anna Liljelund Hedqvist, and Jessica Rothman. Indevelop, 2013.
- “Terms of Reference for this Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks.” Reference No. 13/000428.
- *Using Individual Assessments in the Workplace*, Leonard D. Goodstein and Erich P. Prine. Pfeiffer, 2006
- *Utilization-Focused Evaluations*, Michael Quinn Patton. SAGE Publications, 2008. (4th edition)
- “World Bank Group Impact Evaluations: Relevance and Effectiveness,” Independent Evaluation Group, June 2012.



Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks

This is a mid-term review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks which serves to reduce procurement costs, provide a more centralised view of reviews and evaluations being done by Sida as well as increase the quality of services overall. The review, using a contribution analysis approach, includes the development of an analytical framework and collection of data from multiple sources (documentary, interviews, surveys, and comparative analysis) to establish evidentiary trends. From the evidence and analysis of the review, conclusions and recommendations are drawn for future framework agreements.

SWEDISH INTERNATIONAL DEVELOPMENT COOPERATION AGENCY

Address: S-105 25 Stockholm, Sweden. Office: Valhallavägen 199, Stockholm

Telephone: +46 (0)8-698 50 00. Telefax: +46 (0)8-20 88 64

E-mail: info@sida.se. Homepage: <http://www.sida.se>

