

Integration of Gender Equality in Program-Led Evaluations
FINAL REPORT
February 2014

Prepared by:

Beth Woroniuk, Consultant
b.woroniuk@bell.net

Submitted to the:

Evaluation Services Unit
Development Evaluation Division
Strategic Policy and Summits Branch
Department of Foreign Affairs, Trade and Development
CANADA

Executive Summary

This meta-review was commissioned by the Evaluation Services Unit (ESU) of the Development Evaluation Division of the Department of Foreign Affairs, Trade and Development (DFATD). Its objectives were to:

- *Review and assess the relevance, quality of information and level of integration of gender equality in program-led evaluations (a specific subset of DFATD/Development evaluations); and*
- *Provide recommendations for and input into a strategy to address the gaps or problems identified.*

There is a general consensus that the international development community has not lived up to its commitments to pursue gender equality objectives across all programming initiatives. Specifically, ensuring robust attention to gender equality in evaluations also remains a challenge. Thus the ESU sought to better understand the extent to which their evaluations integrated gender equality and opportunities to move forward.

The methodology involved a literature review; analysis of a sample of evaluation documents (terms of reference, work plans and final reports) provided by the ESU; interviews with DFATD/Development Gender Equality Specialists, ESU staff and DFATD program managers who had managed evaluations in the sample; and a survey of local Gender Equality Advisors (field-based, local staff who support in-country DFATD programming). The analysis of evaluation documents was rigorous and went far beyond counting the number of times the word 'gender' appeared. An assessment of the quality of the analysis and coherence with DFATD/Development's Gender Equality Policy was made.

This meta-review concluded that integration of gender equality in DFATD/Development program-led evaluation is weak. While there are some exceptions, overall the evaluation documents in the sample fail to achieve satisfactory ratings. DFATD/Development staff concurred with this assessment, noting that there are few good examples of evaluations that have successfully integrated gender equality. Furthermore, the review found that program-led evaluations rarely included new lessons relating to gender equality.

The review outlines factors that influence the integration of gender equality in DFATD program-led evaluations:

- *the experience and capacity of evaluators (including knowledge of DFATD's approach and expectations regarding gender equality results, indicators and analysis; experience with gender-sensitive evaluation methodologies; and general background in gender equality issues and challenges);*
- *the extent to which gender equality issues, questions and methodologies are included in evaluation terms of reference;*
- *the extent to which and when in the process DFATD gender equality specialists/advisors are consulted and involved;*
- *the extent to which the project/program design includes a gender-sensitive evaluation plan as well as an explicit gender analysis, gender equality results and indicators, and a sex-disaggregated baseline;*

- *the level of awareness and capacity of DFATD staff to work with gender perspectives and results across all programming areas, especially when managing evaluations;*
- *difficulties evaluating changes in gender inequalities and relations given limited time frames and limits on evaluation methodologies that often do not reach to the household/beneficiary level;*
- *the cost of including a gender equality specialist on the evaluation team; and*
- *the lack of current DFATD/Development guidance on gender equality issues in evaluation.*

Interviews with DFATD staff highlighted the importance of improving the overall context for DFATD's work on gender equality issues. Progress on gender equality in evaluations is facilitated by a strong overall climate for work on gender equality in the Department, including staff capacity on this issue. They also stressed the linkages between good program design and attention to gender equality in evaluations. If the project or program has a good gender analysis, explicit gender equality results, clear gender equality indicators and a sex-disaggregated baseline, they noted that there was a much greater probability that the evaluation would examine gender equality issues in a meaningful fashion.

As well, staff concurred that there was a need to improve capacity on this issue – both internally and among evaluators.

The meta-review concludes with a series of recommendations for DFATD/Development, the ESU and for Gender Equality Advisors and Specialists.

Contents

Executive Summary	1
Acronyms	4
1. Introduction and Context	5
1.1 General Challenges in Ensuring Attention to Gender Equality in Evaluations.....	5
1.2 Gender Equality Issues in Evaluation: The Literature	6
1.3 DFATD Policy Context.....	7
1.4 A Final Note on Terminology	8
2. Rationale and Purpose	8
3. Overall Approach & Methodology	9
3.1 Gender Equality Issues in Evaluations – What Would It Look Like?	9
3.2 Methodology.....	10
4. Findings	12
4.1 Analysis of Evaluation Documents	12
4.2 Findings from Interviews and Surveys	17
5. Conclusions	27
6. Recommendations	29
7. Good Practice	30
Annexes	32
Annex 1 – Documentation	33
Annex 2 – List of Evaluation Reports	35
Annex 3 – Blank Document Review Grids	37
Annex 4 – Interview Guides & Survey	39
Annex 5 – People Interviewed	47
Annex 6 – Terms of Reference	48

Acronyms

CIDA	Canadian International Development Agency
DFATD	Department of Foreign Affairs, Trade and Development
DWAP	District Wide Assistance Project (Ghana)
ECG	Evaluation Cooperation Group (IFAD)
ESU	Evaluation Services Unit, Development Evaluation Division, DFATD
IFAD	International Fund for Agricultural Development
FAO	Food and Agriculture Organization
GE	gender equality
GEAP	Gender Equality Action Plan (previously CIDA now DFATD/Development)
HR	human rights
LM	logic model
OECD/DAC	Organization for Economic Co-operation and Development/Development Assistance Committee
OPEV	Operations Evaluation Department (African Development Bank)
PMF	performance measurement framework
RBM	results-based management
ToR	terms of reference
UNEG	United Nations Evaluation Group

1. Introduction and Context

In general, there is a consensus that the international development community has not lived up to its policy commitments to include gender analysis and gender equality results across all programming initiatives. Meta reviews of evaluations carried out by multilateral and bilateral development assistance agencies note that progress has been both less than expected and uneven across programming sectors and initiatives (OPEV, 2011 and ECG, 2013).¹ These corporate reviews rarely focus on evaluation but the analysis that does exist, notes that – to date – evaluations have not consistently incorporated attention to gender equality issues into their methodology and scope.

It has been pointed out that the evaluation community has not always addressed gender equality issues in evaluation practice. “The mainstream definitions of the OECD/DAC [Organization for Economic Cooperation and Development/Development Assistance Committee] criteria are neutral in terms of the HR [human rights] & GE [gender equality] dimensions. As a result, their application in evaluations often does not take into account HR & GE with the end result of producing evaluations that do not substantively assess these important and cross-cutting dimensions.” (UNEG, 2011, p. 25)

Thus the issue of how evaluations address gender equality issues is a problem throughout the development community, not just for Canadian International Development Agency (CIDA) in the past and the Department of Foreign Affairs, Trade and Development (DFATD) in the present and future.²

1.1 General Challenges in Ensuring Attention to Gender Equality in Evaluations

There are numerous challenges in ensuring that gender equality issues are incorporated into evaluation methodologies and practice – generally and for DFATD/Development in particular:

- Gender equality issues may not be prominent in the initiative design. As noted by institutional evaluations, most organizations still do not consistently incorporate gender analysis and gender equality outcomes into the design of initiatives. When the gender analysis is weak and when gender equality results and indicators are not included in the logic model and performance measurement framework, then these issues often fall off the agenda of evaluators.
- Terms of reference guidance on the gender equality issues to address during the evaluation may be overly general.
- Evaluation teams may lack expertise on gender equality and evaluation. There are no agreed standards for what constitutes an expert on gender equality issues in evaluation.
- There may be conceptual confusion, with ‘gender’ used as a synonym for ‘women.’
- There may be too many competing evaluation issues (including the other two cross-cutting issues: environment and governance) with the evaluation team unable to see the synergies.
- DFATD/Development guidance on gender equality in evaluations is dated (from 2001), rarely used, and staff have raised questions about its usefulness.

¹ For complete citations of documents see Annex 1.

² In 2013 the Canadian International Development Agency was amalgamated with the former Department of Foreign Affairs and International Trade and the new department was called the Department of Foreign Affairs, Trade and Development.

- One evaluation expert has argued that there is “little political will to make gender equality a central issue in evaluation.” (Espinosa 2013, p. 79)

1.2 Gender Equality Issues in Evaluation: The Literature

Four different strands on gender equality issues in evaluation can be found in the literature. First, as mentioned above, there is considerable literature on evaluating institutional policies and strategies relating to gender equality and there is a growing debate on the effectiveness and feasibility of ‘gender mainstreaming’ (the primary strategy or approach adopted by many institutions and organizations) (for example: OPEV, 2011; ECG, 2013; Brouwers, 2013).

A second strand is the emerging discussion of ‘feminist evaluation’ (Podems 2010 and Hay 2011). “Feminist evaluation theorists tend to describe feminist evaluation as flexible and do not recommend a strict approach or provide a framework; rather it is described as a way of thinking about evaluation.”³ However, six tenets have been distilled from the thinking on feminist evaluation:

- Feminist evaluation has as a central focus the gender inequities that lead to social injustice.
- Discrimination or inequality based on gender is systemic and structural.
- Evaluation is a political activity; the contexts in which evaluation operates are politicized; and the personal experiences, perspectives, and characteristics evaluators bring to evaluations (and with which we interact) lead to a particular political stance.
- Knowledge is a powerful resource that serves an explicit or implicit purpose.
- Knowledge should be a resource of and for the people who create, hold and share it. Consequently, the evaluation or research process can lead to significant negative or positive effects on the people involved in the evaluation/research. Knowledge and values are culturally, socially and temporally contingent. Knowledge is also filtered through the knower.
- There are multiple ways of knowing, with some ways privileged over others (depending on the perspectives of the evaluator) (Podems 2010).

Some evaluation specialists, such as Hay (2011), have argued that both feminist and equity-focused evaluations point out that merely documenting inequities is inadequate and that “quality evaluations using either approach should also seek to reduce those inequities.” (p. 40) Feminist evaluation methodologies thus raise issues about what is evaluated, how it is evaluated and how the evaluation findings are used.⁴

A third strand in the literature is the discussion of the challenges of evaluating initiatives specifically designed to address women’s empowerment or gender equality (for example, see Battliwala & Pittman, 2010). The argument is that changing gender relations and building women’s empowerment is difficult to measure, involves timeframes that exceed normal project planning periods of 3-4 years, is often highly contested and may often involve backlash. Thus, it is contended that traditional, linear evaluation methodologies and frameworks may not be the most effective at assessing this type of change.

The final strand and the one of most concern for this meta-evaluation, is the discussion of how evaluations can and should take up gender equality issues (which may or not explicitly adopt the

³ Podems, 2010, p.4.

⁴ Podems (2010) argues that there is a difference between feminist evaluations and evaluations that incorporate a gender approach. She demonstrates, however, that elements of these two approaches can be successfully combined.

terminology and all of the tenets of 'feminist' evaluation). CIDA/DFATD has been wrestling with this issue for over a decade, seeing this question as part of the implementation of its Gender Equality Policy. For example CIDA was one of the first organizations to publish guidance on gender equality dimensions in evaluations (CIDA, 2001)

More recently, this has been the focus of efforts within the United Nations Evaluation Group (UNEG). It has developed guidance on integrating human rights and gender equality in evaluations (Eriksen *et al.*, 2011; Sanz, 2010; UNEG, 2011). This guidance notes that evaluations that address human rights and gender equality share inter-related principles such as:

- Equality and non-discrimination;
- Participation and inclusion;
- Empowerment as a goal; and
- Accountability and the rule of law (Sanz, 2010, p. 140).

Other evaluation practitioners and academics have also explored how evaluation practice could better incorporate gender equality dimensions. Several organizations have developed manuals and guidelines (SDC, no date; World Bank et al, 2009). Very recently, the Food and Agriculture Organization published *Guidelines for Quality Assurance of Gender Mainstreaming in FAO Evaluations* (FAO, 2013). Academics have argued that evaluation can be an important tool in furthering efforts to build greater equality between women and men and improve evaluation practice (de Waal, 2006; Espinosa, 2013; Faúndez-Meléndez, 2012). Espinosa, in particular, offers an interesting overview of what it means to carry out a 'gender sensitive' evaluation and some of the challenges in operationalizing this type of approach.

1.3 DFATD Policy Context

DFATD/Development's (CIDA's) Gender Equality Policy clearly spells out the commitment to using a gender analysis and incorporating gender equality results in all programs, policies and initiatives. DFATD evaluations have an important role to play in assessing the extent to which this policy commitment is being carried out. A reasonable question to ask during an evaluation is: how has the initiative being evaluated contributed to one or more of the Policy objectives, namely:

- To advance women's equal participation with men as decision makers in shaping the sustainable development of their societies;
- To support women and girls in the realization of their full human rights; and
- To reduce gender inequalities in access to and control over the resources and benefits of development.

The Gender Equality Policy does not explicitly take up evaluation issues, but it does set out suggestions for performance measurement:

- Gender equality results are expressed, measured and reported on using qualitative and quantitative indicators;
- Data, disaggregated by sex, as well as by age and socio-economic and ethnic groups, is collected;
- Qualified gender equality specialists (especially locally-based ones) are involved in performance measurement;
- Information on progress in reducing gender inequalities is collected and analyzed as an integral part of performance measurement;

- A long-term perspective is taken (i.e. social change takes time); and
- Participatory approaches are used, where women and men actively take part in the planning of performance measurement frameworks, in their implementation, the analysis, and in decision-making stemming from the analysis.

1.4 A Final Note on Terminology

The amalgamation of CIDA into DFATD has created challenges for language in this report. When referring to past projects and internal units and programs, the previous (then relevant) names have been used. When referring to current or future initiatives, the new structure and titles have been used.

As well, the phrase ‘integrate gender equality into evaluations’ has been used to include attention to gender analysis and gender equality results/indicators in the initiative or organization being evaluated as well as incorporating gender-sensitive methodologies (as discussed below). If an evaluation ‘integrates gender equality,’ it is assumed that it incorporates gender equality issues in a substantive manner throughout the evaluation. For example, the methodology used in this meta-review to assess evaluation documents goes beyond counting the number of times the phrase ‘gender equality’ is used or looking to see if there is a gender equality section. It is a substantive review of whether or not the evaluation took up gender equality issues in a consistent manner across all areas of the evaluation, not just inconsequential mentions sprinkled throughout a document. This is explored more thoroughly in the section on methodology.

2. Rationale and Purpose

DFATD/Development’s policy on Gender Equality requires that all development policies, programs and projects integrate a gender equality perspective. The 2008 evaluation of then-CIDA Gender Equality Policy pointed to an increasing trend to include gender equality in formative evaluation terms of reference (ToR), and to a high percentage of assessments of gender equality issues in summative evaluations. Unfortunately, the evaluation of gender equality integration in ToR and reports was based on counting the times the word ‘gender’ was mentioned in those documents. A useful assessment of gender equality integration requires a stronger methodological approach and analysis.

As well, the DFATD/Development Gender Equality Action Plan includes the commitment to ‘increase feedback from project and program evaluations on gender equality results’ (Strategy 2.3) and the Development Evaluation Division is currently looking to develop a strategy to ensure a more effective and consistent approach to assessing and reporting on gender equality performance.⁵

The specific objectives of this evaluation are to:

- Review and assess the relevance, quality of information and level of integration of gender equality in program-led evaluations; and
- Provide recommendations for, and input into, a strategy to address the gaps or problems identified.

The primary users of the evaluation report will be the Development Evaluation Division, development programs and the Gender Equality Specialists/Advisors throughout the development arm of DFATD.

⁵ The rationale and purpose outlined here are identical to the rationale and purpose included in the terms of reference for this study.

3. Overall Approach & Methodology

This is a meta-evaluation to review and assess the extent to which gender equality issues have been incorporated into the key documents of program-led evaluations (from a sample provided by the Evaluation Services Unit (ESU)).⁶ Interviews with CIDA staff allowed for the exploration of issues generated by the document analysis in more detail and the identification of ways of moving forward.

3.1 Gender Equality Issues in Evaluations – What Does It Mean?

Given that there is often confusion and vagueness around ‘gender in evaluations’, it is important to set out key characteristics of an evaluation that fully incorporates gender equality issues. What criteria could be used to assess whether or not an evaluation ‘integrates gender equality’? This list is based on CIDA’s policy commitments to gender equality and a review of the literature:

- Consistent with an RBM approach, the evaluation examines if/how the initiative contributed to the objectives of the Gender Equality Policy.
- The evaluation looks at **both** how the initiative was designed and implemented (e.g. explicit identification of gender equality results and issues, systematic monitoring and reporting on gender equality performance) **and** at gender equality elements that were not part of the original design (unintended results, excluded stakeholders, missed opportunities, etc.).
- The evaluation assesses the extent to which the initiative was informed by gender analysis. Gender is not used as a synonym for women. There is acknowledgement that changes in gender relations and inequalities involve changes at the formal level (laws, policies, formal representation, etc.) and in areas more difficult to assess, such as culture, beliefs and daily practices. There is a consideration of the negative impacts of the intervention on gender relations and the position of diverse groups of women.
- In the case of project/program initiatives, the evaluation notes the gender equality code given to the initiative and assesses if this was the correct code to apply to the completed initiative.⁷
- The evaluation methodology incorporates elements such as mixed methods (including participatory methodologies) and disaggregates inputs/contributions from women and men (ideally with an understanding of diversity). The methodology uses a gender lens throughout. Data is sex-disaggregated (and if possible also disaggregated for factors such as age, ethnicity, socio-economic position, etc.). This data is not only included but also analyzed for trends and gaps. The evaluation methodology also notes the constraints in reaching marginalized populations (especially women) and develops strategies to overcome these constraints. It also includes ethical standards involved in dealing with specific stakeholder groups.⁸

⁶ Given the availability of documentation, they include ToR, work plans and evaluation reports. Management responses are not included.

⁷ DFATD/Development assigns a gender equality code to all project/programs. Even though this specific guidance note is relatively recent (2013), the practice of assigning codes has been going on for numerous years.

⁸ The United Nations Evaluation Group (UNEG) notes that ethical considerations are a critical element in evaluations and has developed guidelines on ethics and behaviour for evaluators (UNEG, 2011)

- If institutional capacity is a key element in the initiative, there is an investigation of the institutional capacity to support gender equality results.⁹
- UNEG and other guidance point to the importance of how the evaluation is used and whether or not the findings are shared with stakeholders, and especially local stakeholders.

3.2 Methodology

This meta-review used the following methodologies:

- *Analytical review of evaluation documents using explicit analysis grids.* Evaluation Services Unit identified 29 evaluations that had at least one of the following: evaluation terms of reference, a work plan, final evaluation report (see Annex 2).¹⁰ This sample included evaluations of specific projects (funded by geographic programs) and organizations (funded by the previous Partnership with Canadians Branch). Sixteen evaluations had all three documents. These were program-led evaluations carried out between 2010 and 2012 by independent evaluators. These reports were reviewed and scored using the evaluation grids in Annex 3.

The review assessed the extent to which gender equality considerations were incorporated into the key sections of each document on a four point scale:

0: no mention of gender equality issues	Unsatisfactory attention to gender equality issues
1: brief/superficial mention of gender equality issues ¹¹	
2: there is some elaboration/discussion of gender equality issues ¹²	Satisfactory attention to gender equality issues
3: there is a robust discussion of gender equality issues, constituting a good practice	

This methodology is markedly different from earlier assessments done by CIDA of gender equality integration in evaluations. Consistent with the criteria set out in the previous section (3.1), this is an in-depth assessment of whether or not the document included key gender equality issues (such as gender-sensitive methodology, sex-disaggregated data, etc.) and incorporated gender equality dimensions across the major sections of the document. This methodology has been used to assess the cross-cutting nature of gender equality issues in evaluation and goes beyond merely ‘ticking the box’ if the words ‘gender equality’ are mentioned.

⁹ Tool 2 in CIDA’s *Framework for Gender Equality Results* outlines issues to consider when looking at multilateral organizations. This framework provides initial guidance on the types of questions that could be explored in other situations.

¹⁰ The group of evaluations provided by the ESU included several projects which had gender equality objectives as their primary focus (gender-equality specific projects). These were removed from the sample, as the focus of the review is to examine evaluations that had gender equality as a cross-cutting theme (not the primary purpose).

¹¹ If the reference to ‘gender equality’ was totally without context or elaboration and appeared to be primarily tokenistic, this was scored as a ‘1’. For example, if the evaluation design matrix said ‘gender equality will be treated as a cross-cutting theme’ but there was no further evidence of this integration, this was scored as a ‘1’.

¹² Sections of reports that only contained the words ‘gender’, ‘gender equality’, ‘women’, ‘girls’ or ‘gender equality results’ without elaboration were scored as 1. In order to receive a 2 rating, the section required, at a minimum, some indication of what the inclusion of these words meant.

If the document did not contain a specific section (i.e., no evaluability assessment or no section on lessons) this was rated as 'not included.' No score was given and this category was not included in the averaging calculation for the overall rating on the document. Thus the document was not penalized for not having that particular section.

The overall scores were analyzed for trends and good practices were noted.

- *Document review* (background documentation – institutional reports, manuals, grey literature and academic studies) on gender equality issues in evaluation (see Annex 1).
- *Key informant interviews* were held with:
 - two out of three staff of the decentralized evaluation group in the Development Evaluation Division (one male/one female);
 - six (out of eight) and two past headquarters Gender Equality Specialists (all female);¹³
 - six program staff (or just under one-third of those contacted) who managed evaluations in the sample (two male/four female);¹⁴ and
 - A representative of the Human Development and Gender Equality Unit in the former Strategic Policy, Planning and Performance Branch (female).

See Annex 4 for the interview guides. See Annex 5 for the names of people interviewed.

An attempt was made to analyze different trends stemming from male/female interviewees. First, in the two categories with both male and female interviewees, the sample was very small. Second, there were no discernible differences in the answers from women and men in these samples. Therefore, the responses are not disaggregated by sex in the analysis.

- *A survey of local Gender Equality Advisors.* The survey (in French and English) was sent to all the specialists in the 'countries of concentration' (16 specialists, all female).¹⁵ Twelve surveys were returned.¹⁶
- *Frequency analysis.* Where possible, a frequency analysis was done of the points raised in all interviews and surveys (by question). Given the wide ranging nature of the interviews, the frequency analysis does not note how many times an issue was raised throughout the interview, but rather calculates responses to specific questions. Given the small numbers, percentages have not been used to explain the number of responses.

Information sources, including documents, were triangulated whenever possible.

¹³ This is all of the Gender Equality Specialists, with the exception of two new specialists who were only recently appointed to their position and had few opportunities, if any, to participate in evaluations. Given the recent departure of 3 Gender Equality Specialists from TSSD, 2 interviews were done with staff who previously held these positions. One Specialist provided written response to the questions, rather than participate in an interview.

¹⁴ Nineteen program staff were invited to participate in interviews, however only six opted to participate in interviews. This sample cannot be considered to be representative of program staff managing evaluations in general. Instead the responses given in the interviews are considered to be illustrative and anecdotal.

¹⁵ Some countries of concentration did not have local Gender Equality Advisors at the time the survey was sent. As well, the local Gender Equality Advisor for Peru, Estela Gonzalez, participated in a phone interview along with a project manager, so her responses/perspectives were captured in that interview.

¹⁶ Surveys were returned from Bolivia, Ghana (2), Ethiopia (2), Mozambique, Tanzania, Indonesia, Afghanistan, Pakistan, Mali and Vietnam.

4. Findings

4.1 Analysis of Evaluation Documents

Finding 1: In general, the evaluation documents reviewed had weak integration of gender equality issues and were poor in terms of gender analysis. The average score of the 16 evaluations that had all three reports was 1.22 out of 3.0 (thus in the unsatisfactory range). See the Methodology section for an explanation of the points rating structure. Only three out of this sample of 16 evaluations (or 19%) were considered satisfactory (receiving a rating of 1.5 or higher).

Table 1 summarizes the average scores, minimum scores and maximum scores for all three types of documents.¹⁷ Observations from this Table include:

- In general, the scores are low across all three types of evaluation documents, with final reports scoring – on average – slightly higher than both terms of reference and work plans.
- Despite the fact that few terms of reference (22%) and work plans (14%) received overall satisfactory ratings, 27% of final reports received satisfactory ratings. In other words, the final reports were actually slightly better than the initial guiding documents.
- There is a wide range in the scores for all three types of reports. In all three categories, at least one document scored very poorly (either 0 or close to 0) and at least one document scored excellent or close to excellent (3 or close to 3). The fact that at least one document in each category received either 3 or just below 3 highlights that there are some evaluation documents that do integrate gender equality issues in a satisfactory fashion.

Type of Document	Average Score (rated 0-3) ¹⁸	Number of documents reviewed	Minimum score	Maximum Score	Number (percentage) receiving a 'satisfactory' rating ¹⁹
Terms of Reference	1.04	23	0	2.8	5 (22%)
Work Plans	0.93	21	0.2	2.7	3 (14%)
Reports	1.22	26	0.2	3.0	7 (27%)

Finding 2: There is no statistical relationship between the ratings on terms of reference and the ratings on the final report for each evaluation.

Even though one would assume that there would be a correlation between the quality of the terms of reference and quality of the final evaluation report (this was a common assumption of many of the

¹⁷ Table 1 only includes the subsample of evaluations that included all three documents.

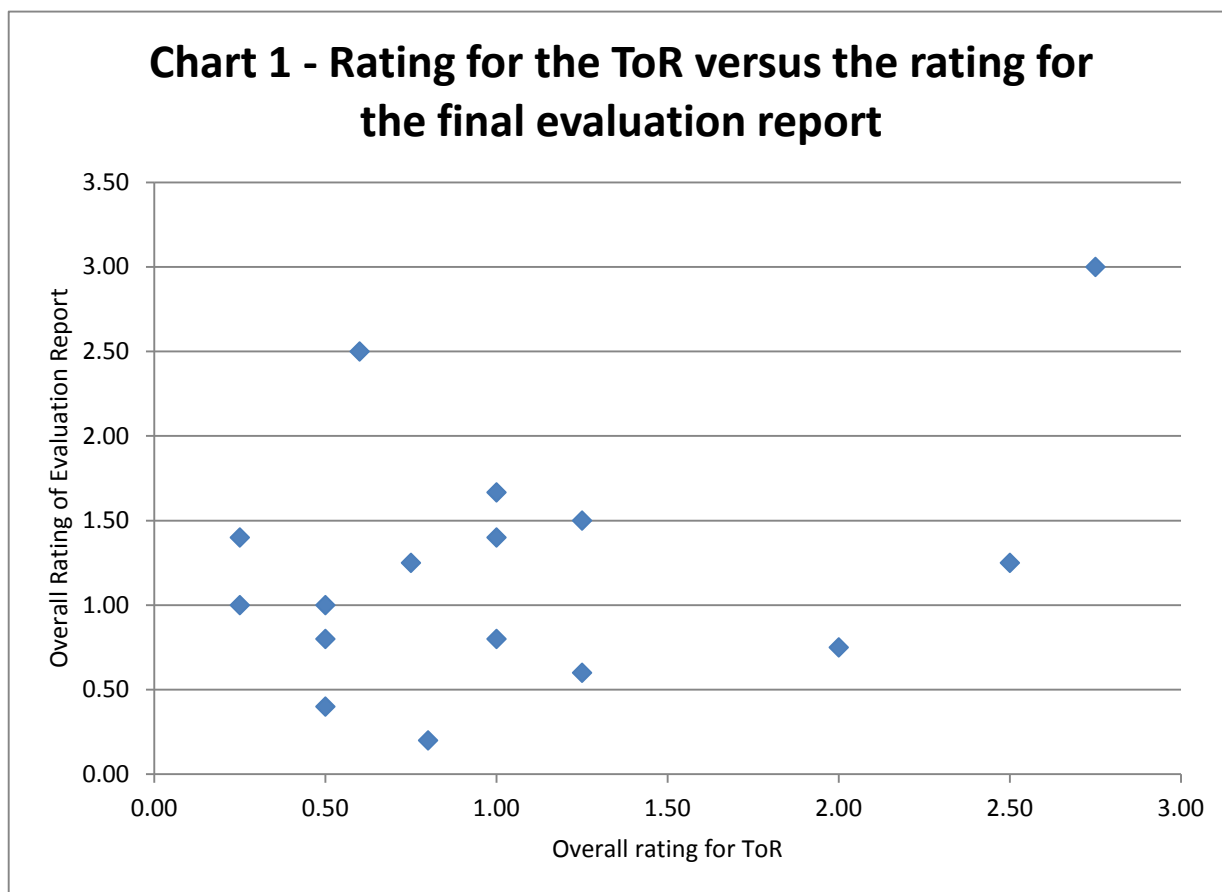
¹⁸ 0=no mention of gender equality issues and 3=robust discussion of gender equality issues.

¹⁹ Given that ratings of 2 and 3 were considered to be satisfactory ratings, to account for rounding, all documents with average ratings of 1.5 or higher were considered to be satisfactory.

people interviewed – see Finding 10), this is not apparent in this sample. See Chart 1. Chart 1 is a ‘scatter plot’ graphing the scores on ToR and final reports. Visually there is no correlation between these two. Had there been, we would have seen the diamonds trending toward a left-right upward line; it is not the case.

There are cases where the report score is higher than the score for the ToR and vice-versa. This **could** indicate that other factors (such as the capacity of the evaluator, whether or not a Gender Equality Specialist or Advisor is involved, etc.) may be more important determinants of successful attention to gender equality in the final report, but a causal relationship cannot be proven in the current meta-evaluation.

However, this should not lead to the conclusion that the quality of the ToR is unimportant. It may simply signal that there is an area for potential gains for improving the quality of final reports.



Finding 3 – None of the sections in the terms of reference documents received an average rating in the satisfactory range.

The analysis of the terms of reference in the sample highlights notes that these documents are weak through all the sections reviewed. Table 2 summarizes the average ratings for each key section of the ToR. Findings from this Table are:

- The project description section was – on average – the strongest section (but this was still in the unsatisfactory). This reflects the fact that this section draws on the official description of the project and in most cases there was some mention of gender equality project description, usually in the official results statements.

- The second strongest section is the section containing the evaluation questions. It is possible that this section scores marginally better given that there is generic standard language in the template for the terms of reference on gender equality results.²⁰
- The standard language on sex-disaggregated data was a footnote that states: “the report will present the findings disaggregated by sex whenever possible.” If the document included this footnote a score of 1 was given. This was not considered strong enough for a ‘2’ rating, given the language of ‘whenever possible’ without elaboration and that this was a footnote rather than located in the text proper. Even so, the average for this section is only 0.65. So not all documents contained even this minor notation.
- Most ToRs did not have a ‘context’ section. In these cases, no score was given for that item and it was not considered in the overall score for each document.

Table 2 – Average Ratings for Key Sections of Terms of Reference						
	Development Context	Project Description	Evaluation Questions	Level of Effort	Requirement for sex-disaggregated data	Overall Rating for Terms of Reference
Average Rating	0.40	1.35	1.30	0.91	0.65	1.04
Number of ‘sections’ reviewed ²¹	5	23	23	23	23	23

- The level of effort section was reviewed to see whether or not there was an explicit requirement for gender equality expertise and/or specific days allocated to a gender equality specialist. A common requirement was phrased as “fully acquainted with CIDA’s RBM and Gender Equality orientation and practices.” This was rated ‘1’, acknowledging that at least there was a mention of gender equality. This was not considered ‘satisfactory’, as familiarity or acquaintance with CIDA’s Gender Equality Policy is not the same as experience and expertise in incorporating gender analysis and attention to gender equality into projects, programs and evaluations.

Finding 4 – The average rating for work plans was less than one, with only one section of the work plans receiving an average rating in the satisfactory range.

Table 3 summarizes the average ratings for each major section of work plans. Findings from this Table are:

- Very few work plans had a section on the development context. Of those that did, none had information on gender equality. This is a missed opportunity as this part of the document could

²⁰ Work plans scoring 2 and 3 on this section often had a common question such as: “What progress is being made toward achievement of results (including specific gender results) at the immediate, intermediate, and ultimate outcome levels?” with a sub-question such as “To what extent has the investment: 1) advanced women’s equal participation with men as decision-makers, 2) promoted the rights of women and girls, and 3) increased women’s access to and control over development resources and benefits?”

²¹ Not all documents had every section. For example only five documents had a ‘development context’ section.

be used to explain how and why gender inequalities are relevant to the context of the initiative and provide background linking general analysis to gender analysis.

- The average rating for the methodology section was less than one. Only two work plans (10%) received a satisfactory rating for this section. Very few evaluations include gender-sensitive methodologies (sex-disaggregated data, assessing and analyzing women’s and men’s responses separately, qualitative analysis, etc.) in their planning. Statements such as ‘gender equality will be addressed as a cross-cutting theme’ that had no further elaboration were scored as 1. Without an explanation of what this meant in the specific context of the evaluation, these types of statements provide little insight and, at worst, appear to be only ‘window dressing’ with no actual impact on the evaluation methodology.
- The highest rated section was the Evaluation Design Matrix. Most work plans contained at least one question related to gender equality, often word for word from the ToR. However, most of the questions were generic with little elaboration or adjustment to the specific project or institution. This raises questions regarding the extent to which there would be an assessment of gender equality results, issues and challenges pertaining to the specific framework of the project or institution.
- Just over half of the work plans included an ‘evaluability assessment’ and only two of these assessments addressed gender equality issues in a meaningful fashion. This marks a large gap in the understanding of gender equality dimensions of evaluations. Usually, assessing changings in gender relations, understanding how women’s perspectives/experiences differ from those of men, accessing information regarding changes at the household level, etc. are difficult challenges for an evaluation and the evaluability assessment offers significant potential to explore how and if these types of issues can and will be explored during the evaluation.

Table 3 – Average Ratings for Key Sections of Work Plans

	Develop-ment Context	Descrip-tion	Evalua-bility Assess-ment	Logic Analysis	Method-ology	Re-sources Allocated for Gender Equality	Eval-uation Design Matrix	Overall Rating for Work Plans
Avg rating	0.00	1.33	0.77	0.56	0.95	0.38	1.67	0.93
# of sections reviewed ²²	3	21	13	9	21	21	18	21

Finding 5 – Even though the overall rating for final reports was low (1.22), the average rating for the ‘findings’ section was in the satisfactory range, indicating that the majority of final reports did include at least one finding on gender equality issues this section. These findings, however, were rarely followed up in the conclusions or recommendations sections (as these two sections had significantly lower ratings).

²² As mentioned above, not all documents included these sections. When the document did include this section, it was rated.

The relatively high score for the ‘findings’ section indicates that the majority of reports included – at a minimum – one finding devoted to gender equality. Only one report received a ‘0’ rating and three reports were rated ‘1’. However, these findings were of an inconsistent quality. Only 11 or less than half were rated as strong.

As well, the findings relating to gender equality were not always followed through in the conclusions or recommendations sections, as can be seen by the significant drop in the ratings for these two sections. The reasons for this failure to address gender equality issues in the conclusions and recommendations sections are not clear.

Consistent with work plan ratings, the methodology section of the final reports was weak (the lowest rating of all the sections in the final reports). Sex-disaggregation of responses was almost universally absent. In one case, the interview guide instructed the interviewer to note whether the interviewee was male or female, but then the analysis was not disaggregated by sex. These poor ratings on methodology also lead to questions about the quality of the findings and whether or not they are fully substantiated by solid methodological approaches.

Table 4 – Average Ratings for Key Sections of Final Reports						
	Methodology	Findings	Conclusions	Recommendations	Lessons Learned	Overall Rating for Final Reports
Average score	0.58	2.20	1.20	1.27	0.67	1.22
# of sections reviewed	26	25	20	26	18	26

Finding 6 – Of the 18 evaluation reports that had a section on either lessons or lessons learned, only five (or less than one-third) mentioned gender equality issues.

Three reports received the highest rating for this section (3), indicating that they included at least one specific lesson on gender equality. Examples of these lessons are included in Box 1. Even though these lessons were given the highest rating, they are not particularly innovative or new lessons. These insights would be common knowledge for Gender Equality Specialists and Advisors.

The little attention given to lessons relating to gender equality is very striking. As noted above, there are considerable challenges in implementing commitments to gender equality in development assistance and much scope to improve overall practice. There is a strong need for improvement and good lessons. Given that less than one-third of the reports included this type of lesson, there appears to be room to improve the quality and increase the number of pertinent lessons related to gender equality in evaluation reports.

Finding 7 – Including a specific section in the evaluation report on gender equality is not equivalent to addressing gender equality across the major sections of the report.

Twenty out of the 26 final reports (77%) had a specific section on gender equality. However, the quality and the detail contained in these sections varied considerably across the reports.

Given that overall rating of the reports was poor, it cannot be concluded that if an evaluation report has a section on gender equality that gender equality has indeed been satisfactorily incorporated into all elements of the evaluation.

Also, one report did not have a specific section on gender equality yet obtained an overall rating in the satisfactory range, indicating that it is possible to address gender equality issues in a cross-cutting fashion without a specific section on gender equality.

Box 1 – Examples of Lessons from Evaluation Reports

- If CIDA does not ensure its cross-cutting themes are adequately addressed in projects it contributes to, chances are they will not be.
- Dedicated assessment and budgeting tools are needed for promoting gender equality...
- A more thorough gender analysis and gender strategy at the design stage would have clarified the specific gender-related results the project was seeking and how these would be achieved. The project would have benefited by more support and practical advice on gender issues and how to address these. What was missing was a very applied and practical analysis and strategy at the outset; it is difficult to retrofit such a strategy later.

4.2 Findings from Interviews and Surveys

Finding 8: In general, the interviews and the survey confirmed the low quality of attention paid to gender equality in evaluations.

Inputs from headquarters Gender Equality Specialists confirmed from their perspective that, in general, evaluation documents pay poor attention to gender equality.²³ Five of the seven Gender Equality Specialists could not identify a specific CIDA evaluation that they thought was well-done with regards to integrating gender equality. (This was partly a result of not being involved in of evaluations).

Local Gender Equality Advisors had a slightly more positive opinion of gender equality integration in evaluations (out of the 12 Advisors answering this question, one thought that in general evaluations were poor, six thought they were good, two thought the ones they had seen were very good and three advisors didn't answer or said that they had not seen enough evaluations to offer an assessment). Seven out of the 12 local Gender Equality Advisors surveyed could provide an example of an evaluation that they thought had done a good job of integrating gender equality, however these were not always program-led evaluations.²⁴ This perception may stem from the fact that local advisors admittedly require training in gender equality and evaluation (see Finding 12 below).

The program officers interviewed said that the evaluations they managed were poor to adequate in terms of attention to gender equality issues. For four out of the five evaluations considered, the program manager and team leader still had questions about 'how much attention to gender equality is enough'.

²³ Two Specialists did not comment as they believed they had not seen enough program-led evaluations to comment.

²⁴ They often mentioned program-wide evaluations, for example. Only one good example mentioned was from the sample used for this meta-review.

In the opinion of ESU staff, most evaluations are characterized by an absence of due consideration to gender equality. They could not cite a single good example.

Box 2 outlines specific comments from the survey and interviews on the overall quality of attention to gender equality in DFATD evaluations.

Finding 9: Program staff and ESU staff see the involvement of gender equality specialists in the evaluation process as important, but – in general – these specialists are only infrequently consulted.

Headquarters Gender Equality Specialists reported that they are rarely involved in discussions of ToR, work plans or evaluation documents. They said that they often do not know when evaluations are being conducted and are only sporadically involved. When they are asked to comment on a document, they

often do not feel part of the process and are somewhat marginalized. One Specialist noted that she did not know what happened to her comments (whether they were used or not). They all agreed that opportunities to be more proactive were limited by their current workloads. It is difficult for them to actively investigate which evaluations are being undertaken and then request to be involved. In the current context, they just find it hard to keep up with the responsive nature of their work and the high number of requests for comments from various sources that they constantly receive.

Field-based Gender Equality Advisors appear to be involved on a slightly more regular basis (of the 11 Advisors answering this question, four said they were not involved in evaluations, three said they were involved from time to time and four said they were regularly involved). They also reported that they worked with a variety of people to ensure the integration of gender equality in evaluations (see Box 3 for an explanation of the various DFATD gender equality positions and responsibilities):

- Headquarters Gender Equality Specialists (noted by 3 Advisors);²⁵
- Other local DFATD Gender Equality Advisors (noted by 3 Advisors);
- DFATD field gender equality focal points (noted by 3 Advisors);
- Local gender equality specialists contracted for specific projects (noted by 3 Advisors);
- Field DFATD gender equality focal points (noted by 2 Advisors);
- Other donor gender equality specialists (mentioned by 2 Advisors); and
- Headquarters country program gender equality advisor/consultant (mentioned by 1 Advisor).

Box 2: Comments on the quality of gender equality integration in evaluations

“The report wasn’t specific enough for us to use. It was very general with no examples.”

- Headquarters Gender Equality Specialist

“There is a lack of understanding of what gender equality involves.”

- Headquarters Gender Equality Specialist

“The evaluator did not identify what worked well and not so well. The gender equality section was very generic. They should have covered more than women’s participation.”

- Program Manager

“In general, gender equality is only seen as women’s involvement, with little attention to changing mentalities or institutional reforms that will address gender gaps (in rights, access and control over resources, etc.)

- Local Gender Equality Advisor

²⁵ As noted in the text, this question was answered by 11 advisors. Most named more than one person they worked with on a regular basis, so the total does not equal 11.

One Gender Equality Advisor noted that the timing of her involvement is often problematic. She is generally called on to comment on a draft evaluation report, without any knowledge of the previous steps in

the evaluation process. Thus she has to look for the ToR, search out project documentation (logic model, performance measurement framework, etc.) and analyze how the evaluation has dealt with gender equality issues in this specific case.

Only three of the headquarters Gender Equality Specialists said that they worked with local Gender Equality Advisors on evaluations. All three of these Specialists were from the former TSSD unit in the former Geographic Programs Branch. Gender Equality Specialists in other former CIDA Branches (multilateral programs and Partnerships with Canadians) did not have regular contact with the local Advisors. However, even those Specialists with geographic responsibilities did not always find it easy to work with the local Advisors – generally and on evaluations -- with the relationship varying from program to program. One Specialist noted that she had regular ‘gender equality team meetings’ that involved local Advisors and field and headquarters gender equality focal points. On the other hand, another advisor noted: “Most programs do not want the headquarters Specialists to be in direct contact with the local advisors without prior permission of the program. There is not an open communication in most programs and most GE local advisors are part-time consultants who are engaged on a request basis only.”

Finding 10: The most frequent challenges or factors cited as influencing whether or not gender equality issues are incorporated into the evaluation process were a) the capacity of the evaluator, b) the extent to which gender equality issues were identified in the evaluation terms of reference, and c) a general lack of understanding of gender equality issues on the part of partners and/or DFATD staff.²⁶

The interviewees identified numerous challenges or obstacles that worked against quality integration of gender equality in evaluations. Table 5 provides examples of specific challenges that were identified and Box 4 includes a sample of comments on the matter.

The top three challenges cited when asked specifically about challenges were:

- **Experience and capacity of the evaluators.** There was strong consensus across the four groups of people interviewed and surveyed that the capacity and knowledge of evaluators is an extremely important factor. Specialists, Advisors, ESU staff and Program Staff all noted that in many cases the evaluators were only able to provide general observations on gender equality

²⁶ An attempt to quantify the frequency of responses to this question is presented in Table 5. The text for this section does not quantify specific responses, given the difficulty of quantifying themes raised throughout the interviews/surveys.

Box 3 – Gender Equality Positions in DFATD/Development

Headquarters Gender Equality Specialists: DFATD staff working full-time on gender equality issues, primarily providing support to geographic programs.

Field Gender Equality Advisors: Locally contracted gender equality specialists to advise on and support country-specific programming.

Field DFATD gender equality focal points: DFATD staff working in field offices who have a coordinating responsibility for gender equality as one part of their job description.

Headquarters country program gender equality advisory/consultant: Some countries programs hire a consultant at headquarters to support their work on gender equality. This person is not staff and may or may not have a direct relationship with the field-based Gender Equality Advisors.

issues. They noted that evaluators often had a narrow understanding of gender equality results and analysis, often seeing women’s participation as the only subject of interest. Several Specialists and ESU staff noted weaknesses in evaluators’ skill in bringing a gender equality lens to methodology (including qualitative methods, sex-disaggregated data, interviewing women and men beneficiaries separately, etc.). Evaluators also lacked an understanding of DFATD/Development’s expectations and practices regarding gender equality. For example, they rarely included an assessment of the ‘gender equality code’ given to the initiative in the evaluation. (These comments were confirmed by the review of evaluation documents. See above section 4.1.)

As noted earlier, if gender equality experience is mentioned in evaluation terms of reference it is often phrased in terms of ‘knowledge of’ CIDA’s Gender Equality Policy. This is not the same as capacity and experience. ESU staff noted that there is a general shortage of evaluators with strong experience in bringing a gender perspective to evaluations and that the screening criteria on the standing offer arrangement (used to pre-select evaluators for CIDA/DFATD evaluations) had not been particularly robust on gender equality expertise.

Several Specialists and ESU staff noted that estimated budgets for evaluations seldom included the hiring of a dedicated gender equality specialist.

It is interesting to note that in the ‘good practice example’ outlined in Section 7 below, the evaluation team had a dedicated gender equality specialist. This appears to have made a major contribution to the overall quality of the evaluation.

- **Attention to gender equality issues in the evaluation terms of reference.** This challenge was noted, in particular, by Gender Equality Advisors and Specialists. They said that it was much harder to argue for attention to gender equality issues in the final report, if this had not been set out in the terms of reference. In some cases, gender equality issues are added once major decisions (such as team composition or the overall focus of the evaluation) are already taken, and only in a cursory manner (e.g. inclusion of words without analysis or follow-up in subsequent text). While the benefits of improving the quality of gender analysis to inform decision-making were clear, one Specialist noted that during an evaluation, the evaluator was reticent to improve attention

Box 4 - Comments on challenges integrating gender equality into evaluations

“The first challenge is rooted in the project design. We can only evaluate something that exists. Where gender equality is effectively mainstreamed, the challenge is less.”

- Local Gender Equality Advisor

“There are so many issues that the evaluation has to look at, so gender equality often comes in low on the priority list. So it is important to make sure that it is on the list of issues to be examined.”

- Program Manager

“If a gender expert is not in the group [managing and conducting the evaluation] throughout the process, the issue will not get the level of analysis it deserves.”

- Local Gender Equality Advisor

“Evaluators often lack capacity. Often their findings are limited to ‘gender is weak’ or ‘very little attention paid to gender equality’.”

- Headquarters Gender Equality Specialist

“The trend toward quantitative methods in evaluation makes it difficult to capture changes relating to gender equality. More emphasis on qualitative methods would be helpful.”

- Headquarters Gender Equality Specialist

to gender equality issues in the final report, arguing that the request for additional information went beyond the terms of reference and the agreed level of effort for the evaluation process. This denotes both a lack of planning for the collection of relevant gender-sensitive data and information, and a lack of understanding of what it means to integrate gender equality as a cross-cutting theme in evaluation.

When discussing the quality of attention to gender equality in evaluations, Program Managers all stressed the importance of the evaluation terms of reference for setting the agenda and scope of issues to be explored. If gender equality questions and issues are not included here, they thought there was little chance of them being included in the evaluation process. For them, this was a crucial step in the process. They noted the importance of modifying the general evaluation questions/template so that it was relevant to the specific project being evaluated.²⁷

This importance of 'getting the terms of reference right' would, at first glance appear to be contradicted by Finding 2 that noted no statistical relationship between a good ToR and a good final report. However, this is not necessarily the case. Given that both ToR and final reports in generally scored quite poorly, it may be possible to contribute to improved final reports through improved ToR. As noted above, Gender Equality Advisors and Specialists indicated that explicit attention to gender equality issues in the evaluation ToR increased their ability to comment and justify their involvement.

- **Lack of understanding of gender equality on the part of partners and/or DFATD staff.** In many cases, it is not clear to partners or DFATD staff how and why gender equality issues are related to the overall results of the initiative; and by extension it is not clear how gender equality 'fits' into the evaluation. Gender Equality Specialists and Advisors linked their marginalization from the evaluation process to the fact that program managers have a poor understanding and appreciation of gender equality as a central cross-cutting theme in all initiatives and evaluations.

On the other hand, Program staff mentioned the lack of understanding by partners was an important obstacle.

Other challenges raised in the interviews and surveys include:

- Poor original project design. This was highlighted by the local Gender Equality Advisors in particular. Almost half of them pointed to the lack of gender analysis, explicit gender equality results and gender equality indicators as contributing to poor attention to gender equality in the evaluations.
- Lack of guidance on gender equality issues in evaluations. Several Specialists pointed out that the current DFATD guidance on gender equality in evaluations is out of date and lacks a results-based focus.

²⁷ However, the responsibility for consulting gender equality specialists or advisors during the development of the ToR to render evaluation questions more relevant and specific, and for conducting regular dialogue with evaluators to ensure that gender equality is addressed throughout the evaluation process, belongs to the program officer or manager who commissioned the evaluation.

Table 5 – Integrating gender equality into evaluations: Challenges cited				
Challenge	Number of times mentioned ²⁸			
	By local Gender Equality Advisors (12)	By headquarters Gender Equality Specialists (8)	By program officers interviewed (6)	By ESU staff ²⁹
Capacity of evaluators (lack of gender equality expertise and/or lack of gender equality specialist on the evaluation team)	6	4	1	✓
Poor original project design (including lack of baselines)	4	2	-	-
Poor elaboration of gender equality issues and questions in evaluation terms of reference	3	4	1	-
Difficulties inherent in evaluating changes related to gender equality / challenges using gender-sensitive evaluation methodologies	1	-	-	✓
Failure of the Project Team Leader to involve gender equality specialist/advisor	3	3	-	-
General lack of understanding of gender equality issues on the project team (including DFATD staff and/or partners)	4	2	4	-

Finding 11: DFATD staff interviewed noted that there is currently no common DFATD/Development mechanism to share lessons and good practices relating to gender equality in evaluation. However, there was no consensus on the type of mechanism that would be both viable (given workloads) and effective/useful.

Almost half of the Local Gender Equality Advisors (five) noted that the in-country donor group on gender equality is an important mechanism to exchange information and share best practices - but two of these acknowledged that the deliberations of and actions taken by the groups could be improved. Other mechanisms used by local Gender Equality Advisors include staff meetings (mentioned by three Advisors of the 11 Advisors answering this question), debriefing sessions with evaluators (mentioned by one), annual reporting (mentioned by one) and internal capacity building workshops (mentioned by one).

The Gender Equality Specialists consistently indicated that there was no current mechanism that facilitated this task. All thought that something would be useful but were skeptical that an effective

²⁸ People often mentioned more than one challenge. This table covers answers to this specific question, not challenges raised throughout the interview/survey.

²⁹ The two staff were interviewed together, so only an indication of issues raised is noted rather than a frequency.

online mechanism could be developed, given the difficulties in maintaining and finding specific lessons/good practices in a database.

The Development Evaluation Division does produce an annual report on Lessons. Gender equality lessons could be incorporated into this report. According to ESU staff, the challenge, to date, has been the absence of relevant and useful lessons generated by evaluations that could be included. This has been confirmed by this meta-review (see Finding 6).

Finding 12: Headquarters Gender Equality Specialists and local Gender Equality Advisors would benefit from increased understanding of the evaluation process within DFATD/Development and how to ensure strong attention to gender equality in evaluations.

The Gender Equality Specialists and Advisors noted that general gender equality analysis only took them so far when commenting on evaluation documents. Both groups generally felt confident providing comments at a general level, but less confident in assessing specific evaluation methodologies.

Several headquarters Gender Equality Specialists noted that the major focus of their work was on initial project design with only limited involvement in implementation, monitoring and evaluation. Only two Gender Equality Specialists indicated that they were confident they had the skills needed to integrate gender equality into evaluations. One Gender Equality Specialist noted that she had a strong research background and this boosted her confidence when reviewing evaluation methodologies as well as content. The remainder of the Specialists indicated that they would like to build their capacity through:

- Taking a course on monitoring and evaluation; or
- Participating in an evaluation team from start to finish (so they could observe the entire process); or
- Receiving a briefing and/or receiving more information on DFATD's approach to evaluations, what happens to evaluation findings once an evaluation is completed, the role and functions of the Evaluation Committee, the most useful advice to provide to evaluators and program officers managing evaluations, the role and function of the ESU, and how best to systematize gender equality inputs into evaluations.

Only three local Gender Equality Advisors replied that they were confident of their skills in the area of gender equality and evaluation. Three other advisors noted that they were confident to some extent but would appreciate more training and five thought that more knowledge and information would be beneficial.³⁰ One headquarters Gender Equality Specialist noted: "Many of the local Gender Equality consultants are from an academic background (and continue to be lecturers in universities) or are former managers of the GE Fund. So there are gaps in their skills around evaluation in general and also in understanding how GE needs to be covered in evaluations specifically."

Finding 13: There was some support for an expanded role for the ESU in strengthening attention to gender equality issues.

The key areas identified where ESU initiatives would be useful include:

- improving guidance on gender equality in evaluations (such as updating the tool from 2001);
- ensuring that evaluators on new Standing Offer lists have strong gender equality experience and capacities;

³⁰ One local Gender Equality Advisor did not answer this question. Therefore the total answering this question is 11.

- officially designating a ‘point person’ who could provide specific expertise on gender equality in evaluation on request to the Gender Equality Specialists (for example, if they had a specific query on sampling or evaluation methodologies);
- providing an initial briefing to the evaluator, program manager and Gender Equality Specialist on how gender equality issues could be included in a specific evaluation; and
- advocating for improved training/capacity for DFATD staff on gender equality issues in evaluation (in both the gender equality courses and the Monitoring and Evaluation course).

One specific request from two Gender Equality Specialists was for the ESU to provide a briefing on how evaluations are conducted at DFATD. This could include everything from how evaluators are selected, to what happens in the Evaluation Committee. Three Specialists noted that the evaluation process appeared to be opaque and they would appreciate understanding more of the context and process of evaluations (from the start of the process to how the findings are used).

However, all of the Program Managers interviewed saw no value added in an expanded role for the ESU during the actual evaluation process. One said that it was more important for an increased role for the gender equality specialists.

If the ESU is to play an expanded role, there may be a need to strengthen the unit’s capacity on gender equality issues in evaluation. There is currently one ESU staff member with considerable experience on gender equality and evaluations, who acts as an informal resource on these issues. The other ESU staff interviewed noted that they did not feel confident in their skills in the area of gender equality in evaluations. They appreciated having a colleague with specific expertise in this area and often relied on her for technical support and advice. Even with this person’s expertise, it may be useful to take steps to improve the Unit’s overall capacity, as the Unit is the gateway for DFATD and other donors’ requests regarding various aspects and types of evaluations, not just for decentralized evaluations. Specific areas where more guidance and capacity would be appreciated include elaboration of specific, precise questions in ToR and more guidance on sampling techniques (in order to offer advice on appropriate methodologies).

Finding 14: There were differing perceptions of the capacity of program managers to incorporate gender equality into the evaluations they managed. On one hand, the Program Managers interviewed all felt confident (with one officer qualifying this to “kind of” confident) identifying the requirements vis-à-vis gender equality throughout the evaluation process. Two (out of six) said that they would need supplementary support from the headquarters Gender Equality Specialist. (It should be remembered that this may not be a representative sample.)

On the other hand, ESU staff said that they consistently had to explain to program officers why questions relating to gender equality were relevant to the evaluation. “Seven out of ten times, they ask us: why do we need this?” Local Gender Equality Advisors also felt that the capacities of DFATD staff on gender equality could – in general – be improved, and this in turn would lead to improvements in the quality of evaluations.

Finding 15: Overt resistance to incorporating gender equality in evaluations is not seen as an issue by the Gender Equality Specialists, Program Managers or the local Gender Equality Advisors.

Only two Local Gender Equality Advisors (out of eight answering this question) said that they experienced resistance or reticence on the part of DFATD staff, partners and/or consultants when they promoted integration of gender equality in evaluations. However, four Advisors (or half of those answering this question) answered a ‘qualified no.’ They said that although there was generally no overt

resistance, they did experience consistent and recurring challenges and this could be seen as a form of resistance. They noted that specific measures (such as consistency and accountability on the part of DFATD staff) were required to support improved attention to gender equality. One Advisor responded: “no to direct open resistance, but there is a need for more ownership and more engagement.”

Only one headquarters Gender Equality Specialist noted that she felt resistance on the part of DFATD staff. Another Specialist thought what was missing was the “how” to ensure attention to gender equality (the capacity), especially when there are competing priorities within the evaluation and it was not clear how gender equality fit into the mix. A third Specialist noted that she didn’t feel resistance for the ‘most part,’ however “there seems to be a reluctance to involve specialists by some programs/staff, so if we don’t find out about it [the evaluation], it could be considered an oversight or a form of resistance. Programs are often sensitive about being evaluated so tend not to share information unless they have to.”

Only one program manager noted some resistance from project implementers to include questions relating to gender equality in the evaluation matrix. They didn’t see the relevance – in general – of gender equality to the initiative. This was overcome through gender analysis and highlighting the key GE issues involved, and through policy dialogue. All the other program managers indicated that they had not encountered any resistance at all.

Finding 16: Improving the quality of reporting on gender equality performance through evaluation requires efforts at two levels: a) front-end attention to gender equality in project/program planning (generally and specifically for monitoring and evaluation); and b) increased attention to gender equality integration throughout the evaluation process and especially during data collection and analysis.

Local Gender Equality Advisors, in particular, made strong links between the general attention DFATD pays to gender equality issues in project planning and the extent to which these issues are part of evaluation processes and content.

Headquarters and program staff, on the other hand, emphasized specific steps that could be taken to address evaluations. These included improving guidance materials, sharing good practices and involving specialists/advisors. Program officers asked for more clarity on how much attention to gender equality in evaluation was enough when gender equality was a cross-cutting theme and not the major objective of the project.

Table 6 provides a summary of some of the repeated responses to the question what would improve integration of gender equality in evaluations. Given the different sample sizes it is difficult to draw many conclusions from the numbers, other than the ones outlined above on what was emphasized by the different groups.

Table 6 – Suggestions from the Interviews and Surveys on how to Improve Integration of Gender Equality in Evaluations				
Suggestions to improve gender equality in evaluations	Number of times mentioned³¹			
	Local Gender Equality Advisors	Headquarters Gender Equality Specialists	Program Staff Interviewed	ESU Staff³²
Improve initial project planning (including initial data collection and sex-disaggregated baselines)	7	1	1	-
Ensure good evaluation ToR	2	-	3	✓
Improve the overall awareness and capacity of DFATD staff to work with gender equality perspectives and gender-equality results	7	-	-	-
Take into account that changes in gender relations take time (often longer than a project)	1	-	-	-
Ensure/encourage regular involvement of gender equality specialists / advisors generally and in the evaluation process	1	2	1	✓
Improve guidance materials on gender equality and evaluation, including documenting and sharing good practices and lessons	1	5	3	✓
Improve the capacity of evaluators on gender equality issues / ensure the selection criteria for evaluators includes more robust attention to gender equality capacities	-	2	1	✓

It should be pointed out that many of these suggestions are beyond the mandate of the Development Evaluation Division to address.

³¹ Only suggestions that were mentioned by more than one interviewee are included. Interviewees often mentioned more than one way to improve attention to gender equality in evaluations. Also, these are responses to this specific question, not a tally of various issues mentioned throughout the interview/survey.

³² Again, given that only two staff were interviewed, this column only indicates which issues were mentioned, not their frequency.

5. Conclusions

Conclusion 1: Despite a long-term commitment to gender equality as both a development goal and a cross-cutting theme, integration of gender equality in DFATD program-led evaluations is weak.

When a robust review framework (consistent with DFATD/Development's Gender Equality Policy and the international state of the art literature on gender equality in evaluations) is used, all three major evaluation documents in the sample provided by the ESU (terms of reference, work plans, and evaluation reports) – on average – fail to achieve satisfactory ratings. This conclusion was confirmed by the interviews and survey.

Given that all three documents are weak, attention is required at all stages of the evaluation process, from the initial conception of the evaluation and the determination of the key issues to be covered, to the evaluation design, to the development of findings, conclusions, recommendations, and lessons.

The above, however, is consistent with the general state of attention to gender equality in evaluations globally. DFATD/Development does not appear to be substantially behind other international development organizations, as other studies have produced similar findings.

Conclusion 2: The extent of gender equality integration in evaluations is linked to overall gender equality integration in DFATD planning and staff capacity.

The stronger the attention to gender equality issues in project design, the more likely these issues will be reflected in program-led evaluations. For example, in summarizing the project description in work plans and final reports, evaluators draw on the project's LFA and PMF. If gender equality results and indicators are part of these project documents, then it is likely that they will be on evaluator's radar as important elements to incorporate into evaluation design and addressed during the course of the evaluation.

Project/program monitoring and evaluation plans also provide a potential entry point to strengthen gender equality in evaluations. It is interesting to note, however, that few DFATD staff interviewed noted these plans as a key mechanism. This may indicate that they are rarely used to their potential.

If gender equality specialists and advisors are consulted and regularly involved throughout the project cycle, then they are more likely to be part of the team commenting and contributing to evaluation ToR and providing quality control on work plans and drafts of final reports. However, if the reverse is true, then there is a major missed opportunity to strengthen gender equality in evaluations.

The implication of this conclusion is that although specific steps to improve gender equality in evaluations can be taken by the ESU and others, significant progress will also depend on broader efforts and investments in gender equality by DFATD. Progress in evaluations is linked to progress and improvements in how the Department moves forward on implementing policy commitments. These investments include the elements contained in the Gender Equality Action Plan relating to capacity, accountability and engagement. Work on gender equality in evaluation cannot be isolated from the overall progress of the Department on the implementation of the Gender Equality Policy.

Conclusion 3: There is still significant uncertainty regarding 'how much is enough' when it comes to incorporating gender equality as a cross-cutting element in evaluations.

As noted in the Findings section, program officers noted that they often face a challenge when attempting to keep gender equality issues on the agenda in an evaluation where this is not the primary focus. They requested more assistance and guidance on this issue.

The confusion surrounding this issue is also reflected by the reporting on the DFATD/Development's Gender Equality Action Plan (GEAP). It was reported to the consultant by a Gender Equality Specialist that according to GEAP reports well-over two-thirds of evaluations "integrated GE". Clearly this figure differs significantly from this meta-review of results. This demonstrates that different yardsticks are being used and that there are different expectations of what constitutes successful integration of gender equality in evaluation.

Updated guidance on gender equality in evaluations that clearly outline criteria and expectations could help to resolve this uncertainty and confusion.

Conclusion 4: There are important capacity gaps regarding gender equality issues in evaluation within DFATD/Development.

Almost everyone interviewed and surveyed for this review noted that they could benefit from improved skills relating to gender equality in evaluation. However, not all staff identified the same required skills. Therefore there is not a 'one size fits all' solution to these capacity gaps, with Gender Equality Specialists and Advisors having different needs than Program Staff and staff of the ESU.

There are different opportunities to address these capacity gaps:

- changes to DFATD's official training courses on gender equality and monitoring and evaluation;
- development of briefings by the ESU for specific target groups on gender equality in evaluations (the development of online materials was suggested by several Specialists);
- development of tools and guidance materials for specific target groups;
- specific coaching and mentoring;
- participation of Specialists/Advisors on evaluation teams; and
- briefings by the ESU on the decentralized development evaluation function at DFATD.

Despite the finding that a mechanism to share evaluation lessons on gender equality does not exist, this is not a priority for short-term investment. First, the review of evaluation documents notes that very few evaluation reports have lessons related to gender equality. Second, there was no consensus on what this mechanism could look like and how it would work. Generally, interviewees stressed the need for something that was easy to use and searchable, but they could not easily envision this type of tool. In fact, there were concerns that the data base (or whatever was developed) would not be efficient and just add to workloads, rather than facilitate information sharing.

The capacity of evaluators was an ongoing issue. In general, the DFATD staff interviewed/surveyed believed that these evaluators lack experience in bringing a gender equality analysis to evaluations and skills in gender-sensitive methodologies. Addressing this capacity gap is more challenging. Given DFATD's inability to provide training to consultants, it is difficult to see how the experience of the current group of evaluators could be broadened. Options include increasing evaluation budgets to include a gender equality specialist in the evaluation team, strengthening the criteria for the selection of evaluators so that there is a greater emphasis on proven experience in gender equality and encouraging professional evaluation associations to take up this issue.

Conclusion 5: Although this review has revealed and confirmed a number of factors that appear to influence the integration of gender equality in evaluations, there is no clear indication which of these factors is the most determinant.

In summary, the various factors influencing the integration of gender equality in evaluations include:

- experience and capacity of evaluators;
- the extent to which gender equality issues, questions and methodologies are included in terms of reference;
- the extent to which and when in the process gender equality specialists/advisors are consulted and involved;
- the extent to which the project/program design includes a gender-sensitive evaluation plan, including an explicit gender analysis, gender equality results and indicators, and a sex-disaggregated baseline;
- the level of awareness and capacity of DFATD staff to work with gender perspectives and results across all programming areas, especially when managing evaluations;
- difficulties evaluating changes in gender inequalities and relations given limited time frames and limits on evaluation methodologies that often do not reach to the household/beneficiary level;
- the cost of including a gender equality specialist on the evaluation team; and
- the lack of current DFATD/Development guidance on gender equality issues in evaluation.

Unfortunately, the triangulation of findings does not produce a consistent ranking of these factors. Therefore it is difficult to conclude, with certainty, which are **the most** important overall or the most influential.

Conclusion 6: There are concrete and specific short- and medium term steps that DFATD can take to improve integration of gender equality in evaluations.

People interviewed and surveyed suggested various options for priorities and tools to improve attention to gender equality in evaluation. Some of these ideas are low-cost, low-effort initiatives, others are more ambitious and wide-ranging. These suggestions were used to develop the recommendations in the next section.

6. Recommendations

Recommendations for DFATD management:

1. Continue to invest in improving capacity, accountability and engagement on gender equality. Improved staff capacity and improved quality of project/initiative design will improve not only the level of integration of gender equality in evaluations, but the quality of reporting on gender equality performance.
2. Ensure that Department training course on monitoring and evaluation includes specific and robust attention to gender equality and that gender equality training courses include attention to evaluation issues and methodologies.

Recommendations for the Development Evaluation Division:

3. Update specific written guidance on gender equality in evaluation. The current guidance dates from 2001. Although it contains relevant insights and information it is out of date with DFATD's results-based operating context.
4. Develop a briefing that could be used at the start of each evaluation, bringing together program managers, Gender Equality Specialists/Advisor and the evaluator. This briefing could outline DFATD expectations for attention to gender equality in evaluations, share best practices and

provide tailored advice for each specific evaluation. It could also serve to involve the Specialists/Advisors at an early stage in the process and highlight the importance of their ongoing involvement.

5. Strongly encourage program officers managing evaluations to consult with the gender equality specialists (field or headquarters).
6. Develop a briefing for gender equality specialists on the evaluation process at DFATD. Given that specialists noted that they often felt marginalized from the practice of evaluation, it would be useful for them to better understand the regular steps in an evaluation process, how evaluators are chosen, OECD/DAC evaluation criteria, the standard ESU templates for evaluation ToR, how to develop gender-sensitive evaluation questions, the difference between monitoring and evaluation, different types of evaluations (mid-term, summative, etc.), how the Evaluation Committee functions, what happens after an evaluation has been completed, how consultants are evaluated, etc.
7. Ensure that new standing offer arrangements (and other pre-screening processes) for evaluators include robust criteria relating to gender equality experience and expertise. Expectations should not just be 'knowledge of DFATD's policy on gender equality.' There is a need to set the bar higher in terms of the capacities required to ensure integration of gender equality.
8. Encourage evaluators to upgrade their skills on gender equality. Professional associations, such as the Canadian Evaluation Society, could be approached and encouraged to build the capacity of members in using gender-sensitive methodologies and integrate gender equality issues in their work. The ESU could sponsor workshops at regular evaluators' meetings or conferences.
9. Work with gender equality specialists/advisors to strengthen ToR so that the generic template is modified and made specific for each evaluation. The less generic the language, the greater the guidance on the specific gender equality issues that should be addressed in the evaluation.
10. Identify and keep a sample of model ToR, work plans and final reports that adequately integrate gender equality and can be shared with program managers, evaluators and gender equality specialists.
11. Network with other donors and multilateral partners (such as UNEG) to share lessons and build capacity on gender equality in evaluation.

Recommendations to Gender Equality Specialists/Advisors (as feasible and as current resources permit):

12. Take advantage of capacity building opportunities to upgrade skills relating to gender equality and evaluation.
13. Identify opportunities and lobby Program management to become members of evaluation teams to increase knowledge of evaluation processes and methodologies.
14. Dedicate part of regular staff meetings to talk about gender equality dimensions of evaluations (informal sharing of good practices, opportunities for building skills, upcoming evaluations, etc.).

7. Good Practice

One evaluation stands out as a good practice: the evaluation of the District-Wide Assistance Program (DWAP) in Ghana.

The Terms of Reference received an overall rating of 2.8 (out of a possible 3). The evaluation questions included extensive and detailed points relating to gender equality including gender equality results,

participation issues, constraints/obstacles faced by women and girls, review of the strategy for gender equality results and the understanding of the need to promote gender equality among stakeholders. The CIDA working group overseeing the evaluation included the local Gender Equality Advisor. The ToR set out an explicit requirement for an expert on gender equality with a strong background in participatory assessment techniques and participatory gender analysis. The findings were to be sex-disaggregated and the ToR also explicitly requested conclusions, recommendations and lessons learned regarding gender equality.

The work plan outlined a clear and specific approach to explore gender equality issues. The methodology included mixed methods, both quantitative and qualitative. Separate focus groups were to be held with women and men. As mandated by the ToR, a gender specialist was part of the evaluation team with a role in the evaluation mission and writing the case studies. The evaluability assessment explicitly included a discussion of gender equality issues. The evaluation matrix included detailed questions on the achievement of gender equality results, the sustainability of these results, constraints to the participation of women and girls, adequate monitoring of gender equality results, etc. The work plan received an overall rating of 2.8.

The final report received an overall rating of 3.0, with every section receiving the highest rating possible. The report noted that the methodology proposed in the work plan had been implemented and an additional element had been added: focus groups with gender equality experts. There was a specific and detailed section on gender equality as well as attention to these issues throughout the report. Gender equality issues were also highlighted in the Executive Summary. There were clear and evidence-based findings and conclusions on gender equality. Five out of the 22 recommendations were related to gender equality and there were specific lessons learned.

One of the Ghana local Gender Equality Advisors noted that she was able to use the findings and recommendations from the DWAP evaluation in follow-up programming: “The recommendation from the DWAP evaluation to develop a gender toolkit for the District Assemblies resulted in me providing technical support to the Institute of Local Government Studies (ILGS) of the Ministry of Local Government and Rural Development (MLGRD) to develop a toolkit to assist the District Assemblies (DAs) to mainstream gender in their processes.”

Annexes

Annex 1 – Documentation

CIDA Documentation

CIDA Evaluation Guide. Evaluation Division, Performance & Knowledge Management Branch. 2004.

CIDA's Gender Equality Action Plan (2010-2013)

CIDA Gender Equality Policy. 1999.

Gender Analysis Assessment Guidance Note (EDRMS 6572945)

Gender Equality Coding Guidance (EDRMS 6572946)

How to Perform Evaluations: Gender Equality. Performance Review Branch. May 2001.

Other

Alfama, Eva & Marta Cruells (nd) *How can evaluation contribute to the Gender Mainstreaming strategy?* Paper presented at the 3rd European Conference on Policies and Gender (ECPG) ECPR Standing Group on Gender and Politics.

Batliwala, Srilatha (2011). *Strengthening Monitoring and Evaluation for Women's Rights: Twelve Insights for Donors.* Association for Women's Rights in Development (AWID).

Batliwala, Srilatha & Alexandra Pittman (2010). *Capturing Change in Women's Realities: A Critical Overview of Current Monitoring & Evaluation Frameworks and Approaches.* Association of Women's Rights in Development (AWID).

Brouwers, Ria (2013). *Revisiting gender mainstreaming in international development. Goodbye to an illusionary strategy.* Working Paper 556. Erasmus Universiteit Rotterdam.

Bustelo, Maria (2003). "Evaluation of Gender Mainstreaming: Ideas from a Meta-Evaluation Study." **Evaluation** 9: 383 – 403.

De Waal, Maretha (2006). "Evaluating gender mainstreaming in development projects," **Development in Practice** 16:02, 209-214.

Eriksen, Janie; Shavanti Ready & Janice Muir (2011). "Human Rights and gender equality in evaluation," *Evaluation for Equitable Development Results*, UNICEF Evaluation Working Papers.

Evaluation Cooperation Group (ECG)(2013). *Gender Equality and Development Evaluation Units: Lessons from Evaluations of Development Support of Selected Multilateral and Bilateral Agencies.* ECG Paper # 5. IFAD.

Espinosa, Julia (2013) "Moving towards gender-sensitive evaluation? Practices and challenges in international development evaluation," **Evaluation** 9(2): 7-82.

FAO/Office of Evaluation (2013). *Guidelines for Quality Assurance of Gender Equality mainstreaming into FAO Evaluations.* OED Tools. September.

Faúndez Meléndez, Alejandra (2012). "Moving Towards a Gender Equality and Human Rights Perspective in Evaluation," in S. Kushner & E. Rotondo (eds.) *Evaluation Voices from Latin America.* **New Directions for Evaluation.** 134: 39-47.

Freeman, Ted; Britha Mikkelsen et al. (2003). *Reflection on Experiences of Evaluation Gender Equality*. Sida Studies in Evaluation.

Hay, Katherine (2011). "Strengthening Equity-focused evaluations through insights from feminist theory and approaches," *Evaluation for Equitable Development Results*, UNICEF Evaluation Working Papers.

International Labour Organization (ILO) – Evaluation Unit (2012). *Integrating Gender Equality in Monitoring and Evaluation of Projects*. September.

Operations Evaluation Department (OPEV) (2011). *Mainstreaming Gender Equality: A Road to Results or a Road to Nowhere? An Evaluation Synthesis*. Working Paper. African Development Bank.

Podems, Donna (2010). "Feminist Evaluation and Gender Approaches: There's a Difference?" **Journal of MultiDisciplinary Evaluation**, 6(14): August.

Sanz, Belen (2010). Module 1: Evaluation and equity; Unit: Human Rights and Gender Equality in Evaluations. On-line e-Learning programme on Equity-Focused Evaluations.

SDC (nd) *Gender and Evaluation*. Sheet 12.

United Nations Evaluation Group (UNEG) (2011). *Integrating Human Rights and Gender Equality in Evaluation – Towards UNEG Guidance*. UNEG/G (2011)2.

World Bank (2005). *Gender Issues in Monitoring and Evaluation in Rural Development: A Tool Kit*.

World Bank, FAO & IFAD (2009). "Module 6: Gender Issues in Monitoring and Evaluation." In **Gender and Agriculture Sourcebook**.

Annex 2 – List of Evaluation Reports

	Project Number	Title and project	ToR	Work plan	Report
1.	A033159	CEJA evaluation: Improvement of Justice Systems in Latin America, Centro de Estudios de Justicia de las Americas (CEJA) / Justice Studies Centre of the Americas (JSCA)	✓	X	✓
2.	A033803	Canadian Comprehensive Auditing Foundation	✓	✓	✓
3.	S063335	PWS&D; PRESBYTERIAN CHURCH (PCC) - 2006-09	✓	✓	✓
4.	Z020641	Decentralized Management Skills Training (End of phase I) - Ukraine	X	X	✓
5.	A034387	Ethiopia Strategy Support Phase II (ESSP II)	✓	✓	✓
6.	S-064204	OPPORTUNITY INTERNATIONAL CANADA (OIC):“Innovative Rural Expansion of Microfinance Services”	✓	✓	✓
7.	S-063317	Programme FPEJL 2007-2011 - La Fondation Jules et Paul-Émile Léger	✓	✓	✓
8.		Organisational Assessment of CANADEM	✓	X	✓
9.	S064757	Organizational Assessment of the International Council of AIDS Service Organizations (ICASO)	X	✓	X
10.	S063757	WVC Governance, Ecosystems and Livelihoods (GEL) Program	✓	✓	✓
11.	A-032705	Evaluation Organisationnelle Conjointe - Ministère de la Santé (Haiti)	✓	X	X
12.	A-031899	Évaluation d'impact District Wide Assistance Project (DWAP)	✓	✓	✓
13.	A032157	Food and Agriculture Products Quality Development and Control Project (FAPQDCP)	✓	✓	✓
14.	A032275	Agricultural Policy Support Facility (APSF)	✓	✓	✓
15.	A033022	Collège des Amériques	✓	✓	✓
16.	A032049	Peace-Building from Below (Nepal)	✓	✓	✓
17.	A032299	China - Migrant Labour Rights	✓	x	x

	Project Number	Title and project	ToR	Work plan	Report
18.	Z020653	Civil Service Human Resources Management Reform (UCS-HRM) (Ukraine)	✓	✓	✓
19.	A031550	Canada-China Legislative Cooperation Project (CCLCP)	X	✓	✓
20.	A033962	Évaluation mi-parcours - projet Pro-Huerta (Haiti)	✓	✓	✓
21.	A021598	Proyecto de Mejoramiento de la Educación Básica (PROMEB)	✓	✓	✓
22.	A031101	South East Asia Regional Cooperation in Human Development SEARCH	✓	X	✓
23.	S63403	COADY-Build Lead, Knowledge.& Capacity 07/12	✓	✓	✓
24.	A033845	CIDA-Tanzania Program – Tanzania Education Network	X	✓	✓
25.	A032558	Observatoire du Sahara Sahel (OSS)	✓	✓	✓
26.	M012105 M012821 M013403	Evaluation on Micronutrient Initiative (MI)	✓	X	✓
27.	A033576	Local Initiatives for Tomorrow - (LIFT2) with CARE	✓	X	✓
28.	S062413	Adventist Dev. & Relief Agency (ADRA)	X	✓	✓
29.	S063440	Agroforestry Practices for Resource Poor Livelihoods	X	✓	✓
		Availability	23 ToR available	21 Work plans available	26 reports available

X = report was not available

Annex 3 – Blank Document Review Grids

Grid to review Terms of Reference

Grid to review Terms of Reference			
Evaluation			
Section	Rating (0-3)	Comments on Rating	Good Practices
Context			
Project description			
Evaluation questions			
Level of effort (explicit requirement for gender equality expertise)			
Requirement for sex-disaggregated data			
Overall Rating			

Grid to review Work Plans

Grid to review Work Plans			
Evaluation			
Section of the Work Plan	Rating	Comments on Rating	Good Practices
Development context			
Program/project description			
Evaluability assessment			
Analysis of the logic of the intervention			
Methodology			
WBS – resources allocated			
Evaluation design matrix (if included)			
Overall Rating			

Grid to review Evaluation Reports

Grid to review Evaluation Reports			
Evaluation			
Section of the Report	Rating	Comments on Rating	Good Practices
Methodology (including sex-disaggregated data)			
Specific section on gender equality issues			
Findings			
Conclusions			
Recommendations			
Lessons			
Overall rating			

Annex 4 – Interview Guides & Survey

1. *Semi-structured interview(s) with DFATD Evaluation Services Unit Staff*

Background

This meta-evaluation examines gender equality integration in decentralized or program-led evaluations (those managed by program staff, not the Evaluation Unit). The objectives are to:

- Review and assess the relevance, quality of information and level of integration of gender equality in program-led evaluations; and
- Provide recommendations for and input into a strategy to address the gaps or problems identified.

The purpose of this interview is to obtain the inputs of DFATD Evaluation Services Unit Staff in supporting decentralized evaluations and the implementation of CIDA's Gender Equality Policy.

The consultant for this meta-evaluation is Beth Woroniuk. She has worked on gender equality issues for over 25 years for a wide range of clients and has carried out evaluations as both team leader and subject matter specialist.

Questions

- In your opinion, do evaluations adequately report on gender equality performance? If so, what percentage of evaluations would you estimate are adequate or very good with regards to reporting on gender equality performance? What criteria are you using to determine adequacy?
- What would an evaluation that fully addresses gender equality include?
- Are you aware of any good examples of gender equality integration in evaluations (either DFATD or other donor evaluations)? If so, please provide specific examples.
- If there are good examples, what are the contributing factors?
- What challenges, if any, do you face when attempting to provide quality assurance of gender equality integration in each of the following: evaluation planning (ToR), design (work plan) and reports?
- Are you confident of your skills in the area of gender equality in evaluations? If not, what would you need to improve your skills?
- Do program officers ask you for clarification or information on how to integrate gender equality in their evaluation? If so, how often (i.e. seldom (5 times or less a year), on occasion (between 5 and 20 times a year, often (more than 20 times a year))?
- Do program officers ask you to explain why gender equality issues are relevant to their evaluation? If so, how often (i.e. seldom (5 times or less a year), on occasion (between 5 and 20 times a year, often (more than 20 times a year))
- Do you provide specific advice to program officers on the improvements required vis-à-vis gender equality integration in consultant work plans and/or draft reports? If so, how often:

seldom (less than 5% of evaluations); on occasion (between 5 %and 25% of evaluations); and often (between 25% and 100% of evaluations)?

- In what percentage of cases are changes made based on your advice? Are those changes generally satisfactory?
- To the best of your knowledge, are there any specific consequences for accepting poor reporting on gender equality performance (DFATD staff) or for poor reporting on gender equality performance (consultants)?
- Is there a mechanism in place to access or share findings, conclusions, recommendations and/or lesson related to gender equality performance among DFATD staff and/or the donor community? If so, is that mechanism effective? If not, would there be a value-added in creating should such a mechanism? What would be its value-added?
- When you reflect on the integration of other cross-cutting themes into evaluations – how does gender equality compare? Are there lessons/insights that apply across all three cross-cutting themes? What makes working on gender equality distinct?
- Overall, how would you rate the skill level of evaluators in the area of gender equality? Should DFATD take measures to strengthen evaluators’ skills? If so, what should those measures entail (e.g. developing tip sheets, pre-design meetings with consultants to explain DFATD expectations vis-à-vis reporting on gender equality performance?)
- Do you use particular guidance materials to support the integration of gender equality issues in evaluations (tip-sheets, etc.)? If so, what are they? Are they adequate and sufficient?
- Given your overall experience, what would further the integration of gender equality issues in evaluations?

2. Semi-structured interview(s) with headquarters gender equality specialists

Background

This meta-evaluation examines gender equality integration in decentralized or program-led evaluations (those managed by program staff, not the Evaluation Unit). The objectives are to:

- Review and assess the relevance, quality of information and level of integration of gender equality in program-led evaluations; and
- Provide recommendations for and input into a strategy to address the gaps or problems identified.

The purpose of this interview is to document the experience of DFATD Gender Equality Specialists with the integration of gender equality analysis and results in evaluations and the perspectives of these Specialists on options for improvements.

The consultant for this meta-evaluation is Beth Woroniuk. She has worked on gender equality issues for over 25 years for a wide range of clients and has carried out evaluations as both team leader and subject matter specialist.

Questions

- How would you rate the overall quality of evaluation findings, conclusions, recommendations and lessons with regards to gender equality integration (i.e. poor, adequate, very good)?
- Do you have a specific example of an evaluation that you supported that you thought was well-done regarding gender equality?
- Are you systematically involved in the evaluation process (i.e. ToR, work plan, discussions with consultants, review of draft reports)? If you are not systematically involved, when are you consulted most?
- Are there opportunities for you to be proactive if not consulted throughout the process? If so, what are those opportunities?
- Based on your experience, what are the main challenges to assuring an adequate integration of gender equality in evaluations?
- Do you face any resistance/reticence on the part of the programs, partners and/or consultants when promoting a proper integration of gender equality in evaluations?
- Are you confident of your skills in the area of gender equality in evaluations? If not, what do you need to improve your skills?
- Do you use findings, conclusions, recommendations or lessons in gender equality stemming from evaluations in your advisory work? If so, can you give a concrete example?
- Is there a mechanism in place to access or share findings, conclusions, recommendations and/or lesson related to gender equality performance with DFATD staff (especially other HQ and local gender equality specialists) and/or the donor community? If so, is that mechanism effective? If not, would there be a value-added in creating should such a mechanism? What would be its value-added?

- Do you work in collaboration with local Support Unit gender equality advisors in ensuring proper integration of gender equality in evaluations?
- When you are not consulted during the evaluation process, are local Support Unit gender equality advisors consulted instead?
- Overall, how would you rate local Support Unit advisors skills in the area of gender equality in evaluations (i.e. poor, adequate, very good) If poor, why is it so?
- Would it be useful to strengthen the skills of local gender equality advisors in gender equality in evaluations?
- What would support improved reporting on gender equality performance?

3. Semi-structured interview(s) with program staff who have managed evaluations in the sample

This meta-evaluation examines the integration of gender equality in decentralized or program-led evaluations (those managed by program staff, not the Evaluation Unit). The objectives are to:

- Review and assess the relevance, quality of information and level of integration of gender equality in program-led evaluations; and
- Provide recommendations for and input into a strategy to address the gaps or problems identified.

The purpose of this interview is to explore the perspective of DFATD staff in incorporating attention to gender equality issues throughout an evaluation that they have managed.

The consultant for this meta-evaluation is Beth Woroniuk. She has worked on gender equality issues for over 25 years for a wide range of clients and has carried out evaluations as both team leader and subject matter specialist.

Questions

On the specific evaluation:

- How would you rate the reporting on gender equality performance in your evaluation (i.e. poor, adequate, very good)?
- How did you highlight gender equality dimensions in the evaluation? (e.g. ToR, exchanges with evaluators, comments on draft reports, etc.)?
- Did you consult the HQ and/or local GE specialist at each stage of the evaluation process (i.e. ToR, work plan, reports)? If so, please provide their names. If not, why didn't you?
- Were efforts made to include a Canadian and/or a local gender equality specialist (other than the support unit advisor) on the evaluation team? If not, why not?
- Was there any resistance/reticence on the part of DFATD management, partners or consultants to integrating gender equality in the evaluation? If so, how did that resistance/reticence manifest itself?
- Were gender equality related findings, conclusions and recommendations presented in the final report? If so, were they used to improve either the intervention evaluated or inform other interventions?
- Did you assess the consultant's performance in terms of gender equality integration in the evaluation? If so, what was the consultant's overall rating (i.e. poor, adequate, very good)? If performance was not rated, why not?
- What do you see as the strengths, if any, regarding gender equality integration in the evaluation you managed?

Generally

- Have you received training in gender equality in evaluations? If not, would it have been useful at the preliminary meeting with the ESU advisor to have a short briefing on the key issues pertaining to gender equality in the evaluation process (ensuring gender equality expertise in

the evaluation team, use of local gender equality advisors, expectations regarding gender equality integration in the work plan, etc.?

- Did you feel confident identifying the requirements vis-à-vis gender equality issues throughout the evaluation process (i.e. when developing the ToR, reviewing the work plan and evaluation reports)?
- Is there a mechanism by which you can access or share lessons or best practices in gender equality with DFATD staff and/or the donor community? If so, is that mechanism effective? If not, would there be a value-added in creating should such a mechanism?
- What do you see as the biggest challenges, if any, in ensuring a proper integration of gender equality issues in evaluations?
- Should the ESU play a bigger role in strengthening reporting on gender equality performance in evaluations? If yes, how so?
- What would support improved integration of gender equality issues in evaluations?

4. Survey for local Gender Equality Advisors

Background

This meta-evaluation examines the integration of gender equality in decentralized or program-led evaluations (those managed by program staff, not the Evaluation Unit). The objectives are to:

- Review and assess the relevance, quality of information and level of integration of gender equality in program-led evaluations; and
- Provide recommendations for and input into a strategy to address the gaps or problems identified.

The purpose of this interview is to explore the perspective of DFATD staff in incorporating attention to gender equality issues throughout an evaluation that they have managed.

The consultant for this meta-evaluation is Beth Woroniuk. She has worked on gender equality issues for over 25 years for a wide range of clients and has carried out evaluations as both team leader and subject matter specialist.

Questions

- In the last 3 years, what has been your involvement in the evaluation process? Are you systematically involved in the evaluation process (i.e. ToR, work plan, discussions with consultants, review of draft reports)? If you are not systematically involved, when are you consulted most?
- How would you rate the overall quality of evaluation findings, conclusions, recommendations and lessons with regards to gender equality integration (i.e. poor, adequate, very good)?
- Do you have a specific example of an evaluation that you supported that you thought was well-done regarding gender equality?
- Are there opportunities for you to be proactive if not consulted throughout the process? If so, what are those opportunities?
- Based on your experience, what are the main challenges to assuring an adequate integration of gender equality in evaluations?
- Do you face any resistance/reticence on the part of the programs, partners and/or consultants when promoting a proper integration of gender equality in evaluations?
- Are you confident of your skills in the area of gender equality in evaluations? If not, what do you need to improve your skills?
- Do you use findings, conclusions, recommendations or lessons in gender equality stemming from evaluations in your advisory work? If so, can you give a concrete example?
- Is there a mechanism in place to access or share findings, conclusions, recommendations and/or lesson related to gender equality performance with DFATD staff (especially other gender equality specialists) and/or the donor community? If so, is that mechanism effective? If not, would there be a value-added in creating should such a mechanism? What would be its value-added?
- Do you work in collaboration with other gender equality advisors in ensuring proper integration of gender equality in evaluations? (Who?)

- Overall, what would support improved reporting on gender equality performance?

Annex 5 – People Interviewed

Program Managers

Jon Silva (Food and Agriculture Products Quality Development and Control Program)

Hilda Nugent and Estela Gonzalez (PROMEB)

Nicole Robillard (LIFT 2)

Steve Podesto (Ukraine)

Christine Mageau (ADRA & PWSD)

Gender Equality Specialists (Headquarters)

Elizabeth Anctil (Canadian partnerships)

Valerie Bilodeau (geographic programs)

Ashley Crossely (Canadian partnerships)

Claire Fishlock (geographic programs)

Soraya Hassanali (multilateral programs)

Shawn Hayes (geographic programs)

Pamela Scholey (past specialist – geographic programs)

Isabelle Solon-Halal (past specialist)

ESU Staff

Pierre Tremblay

Sandra Gagnon

Human Development and Gender Equality Unit in Strategic Policy and Performance Branch

Duy Ai Kien

Annex 6 – Terms of Reference

Integration of Gender Equality in Program-led Evaluations

1. Introduction

CIDA-funded evaluations serve as a practical management tool for assessing and reporting on performance. They enable the Agency, its partners and the donor community to learn from experience and strengthen aid effectiveness.

In 2008, the Evaluation Directorate (now called the Evaluation Division) conducted the *Evaluation of CIDA's Implementation of its Policy on Gender Equality*, and in 2010 a *Gender Equality Action Plan (2010-2013)* was approved by senior management to address the implementation gaps identified in the report.

2. Rationale and Purpose:

CIDA's Policy on Gender Equality (GE) requires that all policies, programs and projects integrate a GE perspective. The 2008 evaluation of the GE Policy pointed to an increasing trend to include GE in formative mid-term evaluation ToR, and to a high percentage of assessments of GE in summative end-of-project evaluations.

Unfortunately, the evaluation of GE integration in ToR and reports was based on counting the times the word "gender" was mentioned in those documents. A useful assessment of GE integration requires a stronger methodological approach and analysis. Notwithstanding the

STRATEGY	ACTIONS	IMPLEMENTING STEPS	TIMELINE	RESPONSIBILITY
----------	---------	--------------------	----------	----------------

above, section 2.3 of the *Gender Equality Action Plan* sets out the following actions and implementation steps for the Evaluation Directorate/SPPB:

2.3 Increase feedback from project and program evaluations on GE results	2.3.1 Take a more consistent approach to assessing GE quality performance in evaluations done by programming branches and Evaluation Directorate	Review and revise the key Agency evaluation frameworks (e.g. Evaluation Framework for Results and Key Success Factors) to ensure sufficient attention is given to gender equality results	2011	Evaluation Directorate, SPPB
		Develop a standard set of criteria or evaluation questions for the way in which CIDA evaluations and/or joint multi-donor evaluations or frameworks will address GE (e.g. requirements in ToR on specific questions for GE results, requirements for GE experience on the evaluation team) and disseminate these throughout the Agency	2012	Evaluation Directorate, SPPB
		All evaluations undertaken by CIDA will include GE key issues and questions in the ToR	2013	DGs, programming branches

The context for “increasing feedback from project and program evaluations on gender equality results” has changed since the development of the *Action Plan*, especially with regards to the status of evaluation guidelines and tools, including ToR. For example, the suite of evaluation guidelines are currently under revision, new evaluation tools are being developed, and the template for Program-led evaluations ToR has been revised to ensure a better integration of gender equality.

Given the current context, the Evaluation Division must develop an updated strategy to ensure a more effective and consistent approach to assessing and reporting on GE performance.

As a first step, the Evaluation Division will commission a review of the relevance, quality of information and level of integration of GE in program-led evaluations. Findings, conclusions, recommendations and lessons stemming from the review will enable the Division to develop an appropriate strategy for addressing the gaps or problems identified.

This consultancy is limited to the first step in the process; the development of a strategy does not form part of this mandate. It is hoped that the results of the review will also serve to inform better reporting on GE performance in Program Evaluations.

3. Objective

To retain the services of a Consultant to support the Evaluation Division in respecting its commitments under section 2.3.1 of the *Gender Equality Action Plan*.

4. Description of Services

The Consultant will carry out the present mandate. She or he will report directly to the Evaluation Manager and have overall responsibility for:

- Reviewing and commenting on these Terms of Reference, if need be;
- Producing a work plan that explains in detail the process and methodology which will lead to the production of an evidence-based report on the current situation with regards to the relevance, quality of information, and level of integration of GE in Program-led

evaluations; Note: Findings will be based on a representative sample of program-led evaluations conducted in fiscal years 2010-2012, and interviews with selected CIDA staff.

- Conducting the assignment according to the approved work plan;
- Collecting credible, relevant and valid information, using a variety of data collection methods and sources of information;
- Reporting regularly on progress to the Evaluation Manager;
- Producing a report presenting findings, conclusions, recommendations and lessons to guide the development of a strategy to ensure an effective and consistent approach to assessing and reporting on GE performance ;
- Ensuring the timely delivery of products as established in the contractual agreement.

All written material submitted to CIDA as per the requirements described under each Task Authorization, must be at a level of quality and standard consistent with senior professional services (i.e., CIDA does not have to edit and/or re-write the document).

5. Deliverables and Schedule

The Consultant must submit the deliverables as per the following schedule:

- Draft work plan – two (2) weeks after the signature of the requisition;
- Final work plan – ten (10) working days after receiving CIDA’s written comments;
- Draft Report – by July 31st, 2013;
- Final Report – by September 30th, 2013.

Unless otherwise stated, the Consultant will, to the extent possible:

- i. use both sides of the page when producing documents, reports, etc.;
- ii. use recycled paper to print and produce reports and other documents.

6. Evaluation Manager’s Roles and Responsibilities

The evaluation manager will review all documents and either accept the deliverables or request modifications through an e-mail notice within a reasonable period.

The evaluation manager will provide the Consultant with the following documents:

- A list of all CIDA branch-led evaluations for fiscal years 2010-2012 with completed files (i.e. which include ToR, work plan, final evaluation report).
- An electronic copy of the evaluations selected by the Consultant.
- Names of potential interviewees.

7. Location of Work

All work must be conducted at the Consultant’s office located in Canada or at CIDA HQ as required.

8. Cost

The professional fee and related expenses - including GST or HST - for the contract will not exceed \$26,000.

9. Method of Payment

- ❖ A first payment of 30% of the total value of professional fees will be paid after the submission of the final work plan.
- ❖ A second payment of 60% of the total value of professional fees will be paid after the submission of the draft final report.
- ❖ The balance of professional fees (10%) will be paid after receiving the final report.

ANNEX: CONSULTANT PROFILE:

The consultant must possess the following Expertise and Experience:

- Extensive expertise and experience in addressing gender equality as a crosscutting theme in development programs and projects and assessing the achievement of gender equality results in evaluations,

The consultant must possess the following skill levels in English:

- Oral =Advanced: Ability to give detailed explanations and descriptions; ability to handle hypothetical questions; ability to support an opinion, defend a point of view, or justify an action; ability to counsel and give advice; ability to handle complex work-related situations.
- Reading = Advanced: Ability to understand texts dealing with a wide variety of work-related topics; ability to understand most complex details, inference and fine points of meanings; ability to read with good comprehension specialized or less familiar material.
- Writing = Advanced: Ability to write explanations or descriptions in a variety of informal and formal work-related situations; ability to write texts in which the ideas are developed and presented in which vocabulary, grammar and spelling are generally appropriate and require few corrections.

The consultant must possess the following skill levels in French:

Reading = Advanced: Ability to understand texts dealing with a wide variety of work-related topics; ability to understand most complex details, inference and fine points of meanings; ability to read with good comprehension specialized or less familiar material.