

DAC EVALUATION SERIES

Review of gender and evaluation

**Final report to
DAC Network on Development Evaluation**

OCTOBER 2003

© OECD 2005

© Software: 1987–1996, Acrobat is a trademark of ADOBE

Permission to reproduce a portion of this work for non-commercial purposes or classroom use should be obtained through the Centre français d'exploitation du droit de copie (CFC), 20, rue des Grands-Augustins, 75006 Paris, France, tel. (33-1) 44 07 47 70, fax (33-1) 46 34 67 19, for every country except the United States.

In the United States permission should be obtained through the Copyright Clearance Center, Customer Service, (508)750-8400, 222 Rosewood Drive, Danvers, MA 01923 USA, or CCC Online: www.copyright.com.

All other applications for permission to reproduce or translate all or part of this book should be made to OECD Publications, 2, rue André-Pascal, 75775 Paris Cedex 16, France.

DAC EVALUATION SERIES

Review of gender and evaluation

**Final report to
DAC Network on Development Evaluation**

JULIET HUNT AND RIA BROUWERS

OCTOBER 2003

Organisation for Economic Co-operation and Development

Pursuant to Article 1 of the Convention signed in Paris on 14 December 1960, and which came into force on 30 September 1961, the Organisation for Economic Co-operation and Development (OECD) shall promote policies designed:

- to achieve the highest sustainable economic growth and employment and a rising standard of living in member countries, while maintaining financial stability, and thus to contribute to the development of the world economy
- to contribute to sound economic expansion in member as well as non-member countries in the process of economic development, and
- to contribute to the expansion of world trade on a multilateral, non-discriminatory basis in accordance with international obligations.

The original member countries of the OECD are Austria, Belgium, Canada, Denmark, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The following countries became members subsequently through accession at the dates indicated hereafter: Japan (28 April 1964), Finland (28 January 1969), Australia (7 June 1971), New Zealand (29 May 1973), Mexico (18 May 1994), the Czech Republic (21 December 1995), Hungary (7 May 1996), Poland (22 November 1996), Korea (12 December 1996) and the Slovak Republic (14 December 2000). The Commission of the European Communities takes part in the work of the OECD (Article 13 of the OECD Convention).

In order to achieve its aims the OECD has set up a number of specialised committees. One of these is the Development Assistance Committee (DAC), whose members have agreed to secure an expansion of the aggregate volume of resources made available to developing countries and to improve their effectiveness. To this end, members periodically review together both the amount and the nature of their contributions to aid programmes, bilateral and multilateral, and consult each other on all other relevant aspects of their development assistance policies.

The members of the Development Assistance Committee are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, the United States and the Commission of the European Communities.

DAC Network on Development Evaluation

The DAC Network on Development Evaluation is an international forum where bilateral and multilateral development evaluation experts meet periodically to share experience to improve evaluation practice and strengthen its use as an instrument for development co-operation policy.

It operates under the aegis of the DAC and presently consists of 30 representatives from OECD member countries and multilateral development agencies (Australia, Austria, Belgium, Canada, Denmark, the European Commission, Finland, France, Greece, Ireland, Italy, Germany, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, the United States, the World Bank, the Asian Development Bank, the African Development Bank, the Inter-American Development Bank, the European Bank for Reconstruction and Development, the UN Development Programme and the International Monetary Fund), plus two non-DAC observers, Mexico and Korea.

Further information may be obtained from Hans Lundgren [hans.lundgren@oecd.org], head of the Evaluation Section, OECD, Development Cooperation Directorate, 2 rue André-Pascal, 75775 Paris Cedex 16, France. Website: <http://www.oecd.org/dac/evaluation>.

Acknowledgements

AusAID, the Netherlands Ministry of Foreign Affairs and the authors wish to thank staff in donor agencies who took the time to submit reports to the review; and both staff and consultants who assisted with follow-up interviews and discussions by email and phone. Thanks also go to members of the DAC reference group of the Working Party on Aid Evaluation, the Working Party on Gender and the DAC Secretariat, who submitted comments during the review process and on drafts of this report.

Since the review was undertaken the Working Party on Aid Evaluation and the Working Party on Gender have become the DAC Network on Development Evaluation and the DAC Network on Gender Equality respectively and are referred to as such where appropriate in this report.

Contents

Organisation for Economic Co-operation and Development	2
DAC Network on Development Evaluation	3
Acknowledgements	4
Abbreviations	7
Terminology Used	9
Summary and Recommendations	10
Purpose and methodology of the review	10
Key findings and recommendations	10
Comparison with previous DAC studies	16
Conclusion – follow-up action required	17
1 Purpose and Background of the Review	18
Purpose	18
Background	18
2 Database for the review	20
Features of an evaluation	20
Thematic and general evaluations	20
Other material submitted to the review	23
Comparisons with previous reviews	24
3 Review framework	26
Evaluation methodology	26
Institutional approaches and change strategies	30
Benefits for women, gender analysis and development outcomes	31
Comparisons with previous reviews	32
4 Findings on Evaluation Methodology	34
Category of evaluations	34
Focus of evaluations	37
Quality of gender analysis in general evaluations	38
Methods, approaches and tools	39
Expertise and quality	43
Conclusions and lessons	45
5 Findings on Institutional Approaches and Change Strategies	48
Partnerships and dialogue on gender equality	48
Donor policy and country assistance strategies	50
Donor agencies	50
The activity cycle	51
Progress towards gender equality at the community level	55
Conclusions and lessons	56

6 Findings on the Benefits for Women, Gender Analysis and Development Outcomes	58
Evidence of benefits for women and changes in gender relations	58
Evidence of a causal relationship between the attention given to gender issues and benefits for women	60
Evidence that benefits for women or changes in gender relations lead to improved development outcomes	62
Conclusions and lessons	64
7 Progress since Previous Reviews and the Need for Follow-Up Action	67
Comparing findings with previous reviews	67
Disseminating findings	69
Conclusion	69
Annexes	
Annex A Terms of Reference for Review on Gender and Evaluation	71
Annex B Documents Submitted to the Review	76
Annex C Summary of Indicators Used in Evaluations	85
Annex D Gender Audits	88
Annex E Features of Evaluation Methodology that Facilitate Gender Analysis	91
Annex F Examples of Benefits for Women, Gender Analysis and Links to Development Outcomes	93
References	105
Boxes and tables	
Box 1 Fundamental factors for building partnerships on gender equality	49
Table 1 Sectoral focus of general evaluations	22
Table 2 Summary of the framework used for reviewing evaluation documents a	27
Table 3 Category and focus of thematic evaluations a	35
Table 4 Category and focus of general evaluations a	36
Table 5 Assessment of the attention given to gender analysis in general evaluations	38
Table B1 Thematic evaluations on gender equality, gender mainstreaming and women's empowerment	76
Table B2 General evaluations that included references to gender issues or gender analysis	80
Table B3 Tools and other documents on gender-sensitive evaluation methodology	83
Table B4 Evaluation reports from other agencies, and other reports	84
Table C1 Indicators used in thematic reports	85
Table C2 Indicators used in general reports	87
Table E1 AusAID evaluations that gave satisfactory attention to gender issues a	91
Table E2 AusAID evaluations that gave unsatisfactory or highly unsatisfactory attention to gender issues a	92

Abbreviations

ADB	Asian Development Bank
AfDB	African Development Bank
AusAID	Australian Agency for International Development
BMZ	Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (German Federal Ministry for Economic Cooperation and Development)
CCA	Common Country Assessment
CIDA	Canadian International Development Agency
DAC	Development Assistance Committee of the OECD
Danida	Danish International Development Assistance
DFID	Department for International Development, United Kingdom
EC	European Commission
GAD	gender and development
ILO	International Labour Organization
IOB	Operations and Policy Evaluation Department, Netherlands Ministry of Foreign Affairs
IPPF	International Planned Parenthood Federation
JICA	Japan International Cooperation Agency
M&E	monitoring and evaluation
MOU	memorandum of understanding
NGO	non-government organisation
NORAD	Norwegian Agency for Development Cooperation
NZAID	New Zealand Agency for International Development
OCAA	Oxfam Community Aid Abroad
ODA	official development assistance
OECD	Organisation for Economic Co-operation and Development
PRA	participatory rural appraisal
PRSP	Poverty Reduction Strategy Paper
RBM	results-based management
SDC	Swiss Agency for Development and Cooperation
Sida	Swedish International Development Cooperation Agency
SPC	Secretariat of the Pacific Community
SWAps	sector-wide approaches

TORs	terms of reference
UNDAF	United Nations Development Assistance Framework
UNIFEM	United Nations Development Fund for Women
USAID	United States Agency for International Development
WFP	World Food Program
WID	women in development
WP-EV	Working Party on Aid Evaluation
WP-GEN	Working Party on Gender
WSS	water supply and sanitation

Terminology Used

The term “results” is used to refer generally to a change that has come about due to a development intervention. A result may be a short-term change (an output), a medium-term outcome or benefit, or a longer term impact.

The term “activity” is used to refer to a development intervention, which may be a whole program or project. The word “activities” refers to tasks or actions that form a part of a project or program, and that need to be undertaken to achieve outputs.

The terms “institutional” and “institutionalisation” refer to how gender equality policy commitments are integrated into overall agency policy, strategies and procedures for planning, designing, implementing and monitoring activities.

The term “management” refers to the management of a project and program, such as partner policy, activity structures and administrative procedures, which are not specifically related to gender equality.

According to the DAC, “gender equality” refers to:

... equal enjoyment by women and men of socially valued goods, opportunities, resources and rewards. Gender equality does not mean that men and women become the same, but that their life chances are equal. (OECD 1999, pp. 13-15)

This does not assume a particular model of gender equality for all societies and cultures. However, a critical aspect of gender equality is equal participation in decision making (OECD 1999, p. 13-15).

“Gender mainstreaming” is a strategy adopted at the United Nations 4th World Conference on Women in Beijing in 1995, and by the DAC and many of its member agencies, to support the goal of gender equality. It requires gender equality concerns to be considered in the formulation and analysis of all policies, programs and projects through the activity cycle. Specific initiatives to enable both women and men to participate, benefit and make decisions are a key aspect of successful mainstreaming strategies (OECD 1999, p. 15).

“Poverty reduction” has been adopted as an overarching goal by the international development community. There is now recognition that poverty is multidimensional. Poverty reduction requires action in economic, social, natural resource and political spheres to reduce inequality, including gender inequality (OECD 2001, pp. 17-19).

Summary and Recommendations

Much work still needs to be done to ensure that both women and men receive the benefits of development cooperation. More attention needs to be given to gender equality issues in evaluations of development activities, and policy commitments to gender equality need to be systematically implemented by donor agencies.

Since the 1999 DAC gender review the number of thematic evaluations of gender equality, gender mainstreaming and women's empowerment has increased. However, overall progress in incorporating gender perspectives into general evaluations of development assistance has been slow and uneven.

These are the main findings of the review on gender and evaluation, which was undertaken during 2002 and early 2003 and builds on the DAC's previous work in this area.

Purpose and methodology of the review

See Chapter 1 and Annex A

The review's purpose was to distil experiences with evaluating gender equality objectives, in order to improve evaluation practice and development outcomes. It focused on three areas:

- evaluation methodology for assessing changes in gender equality, gender mainstreaming and women's empowerment
- institutional approaches and change strategies that have addressed recurrent obstacles to gender mainstreaming, and
- improvements in overall development outcomes due to gender analysis and gender mainstreaming.

See Chapters 2 and 3 and Annex B

The review's database included 42 **thematic evaluations** specifically designed to evaluate gender equality, gender mainstreaming and women's empowerment efforts, and 43 **general evaluations** that included some gender analysis or references to gender issues. Each evaluation was qualitatively assessed in relation to the three areas of focus, which had been used to select the general evaluations included in the sample.

Key findings and recommendations

See Chapter 4

Evaluation methodology

For thematic evaluations, the review highlighted methodological problems associated with treating gender mainstreaming as a goal rather than as a means to the long-term objective of gender equality. A number of thematic evaluations focused on evaluating institutional mainstreaming, without considering the extent to which this leads to changes in gender

relations. It is important for future thematic evaluations to focus on results, as well as on institutional mainstreaming practices.

Problems also arose when thematic evaluations were based on the assumption that gender mainstreaming leads to gender equality outcomes. It proved to be more useful to, first, find positive or negative gender equality results of activities and, then, examine the factors that promoted good or poor performance.

After reviewing both thematic and general evaluations it was concluded that evaluating gender equality issues or objectives requires, in the first instance, good quality evaluation design and implementation. That is, an evaluation needs to have a clear and simple objective, a transparent design with systematic data collection methods, findings based on evidence, clear evaluation criteria, and gender-sensitive indicators.

More specific findings on evaluation methodologies follow.

- Few general evaluations employed gender-sensitive indicators. Not surprisingly, all those that did had a higher quality of gender analysis and tended to include more information on benefits for women.
- Many evaluations were faced with a lack of sex-disaggregated baseline data and limited monitoring information, particularly on gender relations and benefits, largely because activity designs did not address gender issues. Faced with these limitations, some evaluations used qualitative data collection techniques to assess results. However, many others did not because of inadequate time and resources. Few studies had sound or comprehensive quantitative and qualitative data about progress made for women and men over the period evaluated.
- Overall, the general evaluations were characterised by poor analysis of gender issues, even when some significant sex-disaggregated quantitative and qualitative data had been collected. Discussion of gender issues was often limited to the participation of women in development activities. There was little examination of unintended impacts for women, or any other group such as men and women living in poverty.
- Most of the evaluations that had satisfactory gender analysis used participatory methods of data collection. However, many general evaluations did not use such methods, and sometimes these methods did not lead to adequate attention being given to gender issues.
- Very few evaluations provided information on the use of local expertise. Follow-up interviews with a few evaluators showed that local experts were considered indispensable in thematic evaluations, although they required adequate resourcing and training. Very few evaluations considered involving local evaluators as a means of building their capacity.

- The evaluations in the sample show no evidence that donors are working towards more joint evaluations (either joint donor or joint donor–partner evaluations), as might be expected in line with the changes towards sector-wide approaches and increased emphasis on national ownership.
- Many thematic evaluations assessed the capacity and commitment of partners to work on changing gender relations. However, this was very limited in general evaluations and confined largely to evaluations of programs undertaken by non-government organisations (NGOs).
- In general, the evaluations did not address the links between gender equality and poverty reduction.
- Many general evaluations in the sample did not include terms of reference (TORs). Those that did tended to include vague references to women or gender. Most (about two-thirds) required gender expertise on the evaluation team.
- Many evaluations used the word “gender” as a synonym for “women”. “Men” were missing from most reports. This means that there was little or no focus on gender relations or on the impact of development activities on gender equality, and little focus on how gender relations had influenced women’s capacity to participate or benefit. On the positive side, in those evaluations that did report on gender equality, the focus was usually on changes in women’s decision making at the community level.

The weaknesses identified in the evaluation methodology of both types of evaluation suggest a need for overall improvements in evaluation capacity, as well as improvements specific to undertaking gender analysis in evaluations. This suggests the need to build evaluation capacity further with, for example, specific guidance from agencies about their expectations for including gender equality perspectives in individual evaluations, leadership and support from agency management, and training and courses on the whole process of evaluation, which also have a strong focus on how to address gender equality issues.

Specific areas of evaluation capacity that need to be improved are:

- the capacity to determine appropriate research questions to investigate potential differences in participation, benefits and impacts for women and men
- the capacity to identify and collect sex-disaggregated information using a range of different methods, and to develop gender-sensitive indicators, and
- the capacity to analyse the data collected in a systematic way.

Recommendation 1
to the DAC

It is recommended that the DAC continue to work with members and partners on building evaluation capacity, focusing particularly on how to address gender equality issues in evaluations.

- Recommendation 2 to agencies *It is recommended that agencies finance staff and partners to attend courses on evaluation that have a strong focus on gender equality perspectives.*
- Recommendation 3 to the DAC *It is recommended that the DAC set up an area on its website that has links to existing gender-sensitive evaluation tools. These should be accompanied by brief descriptions and commentaries to guide users to the tools that best suit their needs. The location of the site should ensure that evaluators can easily access the tools.*
- Recommendation 4 to the DAC *It is recommended that the DAC Network on Development Evaluation and the DAC Network on Gender Equality hold a joint workshop on evaluation capacity building, with a strong focus on how to address gender issues. One area for discussion should be the weaknesses in evaluation capacity found in this review, and agencies' experience of how to address these weaknesses. The workshop should primarily be forward looking and cover ground not adequately addressed in the sample of evaluations included in this review or in previous DAC workshops. For example, the workshop should focus on the evaluation of poverty reduction strategies and outcomes, sector-wide approaches (SWAs) and other program-based approaches. Partner agencies must be involved. The workshop could also be a forum for determining further collaborative work on gender and evaluation, including priorities for future research and evaluation.*

See Chapter 5 **Institutional approaches and change strategies**

The review found that partnerships and dialogue on gender equality were fundamental to addressing gender issues and are essential for promoting and consolidating changes in approach. Among the evaluations reviewed, four factors were consistently regarded as essential for building partnerships and promoting dialogue:

- national or partner ownership of gender equality goals and strategies, including a shared vision on gender equality either at country assistance or activity level
- consensus between donors and partners on objectives that are clearly stated and transparently relevant to the partner government's policies and commitments
- stakeholders (including partner agencies, civil society and particularly women's organisations) involved in advocacy and setting directions, and
- long-term commitment from donor agencies.

Explicit links between poverty reduction and gender equality objectives were rarely made at the policy level, in country assistance strategies and in individual activities, including in activity designs. This is an area requiring urgent attention.

A lack of accountability within donor agencies to gender equality policies, including a failure to integrate gender equality perspectives into agency procedures, was highlighted as a continuing problem. Leadership and commitment were found to be essential to address this obstacle, and overall agency plans were useful for some agencies.

Recurring obstacles in activity design and implementation included a lack of baseline data on gender issues, a lack of participation by women, poor needs analysis, a lack of expertise in gender and social analysis, and a failure to address gender issues in activity objectives. Where gender analysis had been undertaken in design, it was often not linked coherently to the implementation of activities, to overall social analysis or to expected results (including poverty reduction). The lack of an explicit gender strategy integrated into overall design was also identified as a constraint.

Among the evaluations reviewed there was general agreement about successful change strategies for activity design, implementation and monitoring. These included ensuring that gender equality issues are adequately addressed in design, addressing responsibilities for implementing gender equality policy in job descriptions and TORs through the activity cycle, and using gender-sensitive indicators as a minimum standard for design, implementation and monitoring.

Using participatory approaches to activity design and implementation, strengthening women's leadership capacity and working with women's organisations as agents of civil society were identified as effective approaches for making progress towards gender equality at the community level.

Recommendation 5
to the DAC

It is recommended that the DAC consider tasking the DAC Network on Gender Equality to develop, in collaboration with the DAC Network on Development Evaluation, a short policy note or communiqué on institutional approaches based on the findings of this report.¹ This should focus on successful strategies for building partnerships on gender equality, other successful change strategies on which there is a high degree of consensus, and links between poverty reduction and gender equality.

Recommendation 6
to agencies

It is recommended that agencies note the evidence in this report on successful change strategies for building partnerships on gender equality and in other areas, and continue their best efforts to ensure accountability to gender equality policies.

¹ A tipsheet, "Effective strategies for promoting gender equality", has been published by the DAC Network on Gender Equality since this report was written.

See Chapter 6 **Benefits for women, gender analysis and development outcomes**

In general, the quality of gender analysis in the evaluations reviewed was inadequate for agencies to assess differences in the benefits for women and men and the relationship of those benefits to overall development outcomes. Systematic attention appeared rarely to be given to gender issues in activity design, implementation and monitoring, except for those activities directly aimed at promoting gender equality.

There is clear evidence that gender issues must be addressed when activities are being designed and implemented if agencies want to increase the likelihood that both men and women participate in and benefit from activities, and to ensure that nobody is disadvantaged. While this is a necessary condition it will not ensure that benefits will be achieved and be sustainable.

Overall, most of the gender benefits identified met women's practical needs. Evaluations of activities that aimed to promote gender equality and targeted women (such as gender equality funds) showed the strongest evidence of strategic changes in gender relations, most often increased participation of women in decision making. Whether or not strategic changes had been planned, the findings reinforced the importance of the local social and institutional context and of partner commitment to sustaining benefits. Key factors in promoting and sustaining benefits for women and strategic changes in gender relations are stronger women's groups, with greater organisational capacity and the empowering effects of some training opportunities.

There is evidence to support the proposition that benefits for women improve the effectiveness of development initiatives and their outcomes at the activity and community (micro) level. For example, women's participation in local governance structures and in project activities, groups and committees increased the effectiveness of development activities. Where women gained control of income, there is evidence that they tended to use this for basic family needs and schooling for children, which both directly and indirectly reduced poverty at the community level.

However, there is no evidence for this proposition in the database of evaluations that focused on results at the macro level, where partner capacity and commitment and the socio-economic and political context are key determinants of the sustainability of development impacts. This may be because so few agencies investigated such links, or because the attention given to gender issues at any level was limited in most cases. Further research is needed in this area. More evaluations that make links between micro, meso and macro levels of results are also required if agencies want to explore issues relating to overall development effectiveness and outcomes at the macro level.

Future research and evaluation should give priority to the links between benefits for women, gender equality and poverty reduction. This work should be jointly undertaken, at

a country and/or sectoral level, by donor agencies and development partners. It should focus on the work of multiple donors in one country, rather than the work of one donor in dissimilar countries.

See Chapters 1, 2, 3 and 7

Comparison with previous DAC studies

Some areas of this review are not comparable with studies undertaken by the DAC Working Party on Aid Evaluation in 1994 and 1999. For example, this is the first DAC review to include evaluations specifically designed to assess gender equality and women's empowerment policies and efforts. As a result, the current review has gathered evidence about successful strategies for institutionalising a commitment to gender equality. This is also the first DAC review to investigate evidence of links between benefits for women, gender analysis and gender mainstreaming efforts, and overall development outcomes. The findings in these areas reinforce extensive anecdotal evidence about the need to increase the attention given to gender issues through the activity cycle.

This review's one common area of enquiry with its predecessors is its focus on gender issues in general evaluations, although all of the reviews used different sampling methods. Taking this into account, progress in addressing gender issues appears to have been rather slow and uneven. This was also the overall conclusion of the 1999 review, which was based on agencies' own assessments of progress.

Due to different sampling methods, it was difficult to assess whether more evaluation TORs mentioned the need to address gender issues. Previous reviews and DAC advice highlighted the need to identify specific questions and issues in TORs to provide guidance to evaluators. This was still rare in the general evaluations included in the current study.

The quality of the analysis of gender issues remains a problem. It varied markedly both between and within agencies, as it did at the time of the 1999 review, and it is reasonable to conclude that there has been little overall improvement since then. Many general evaluations did not use gender-sensitive indicators, collected little information on gender issues, and did not analyse the information collected. Furthermore, while many general evaluations used standard evaluation criteria² as an overall framework for data collection and analysis, gender equality issues were not systematically integrated into this analysis.

The majority of the current sample of general evaluations focused on activities and outputs, rather than outcomes or impacts. This was also evident in the evaluation reports included in the 1994 and 1999 reviews.

² Standard evaluation criteria as defined by the DAC include relevance, achievement of objectives, efficiency, effectiveness, impact and sustainability.

See Chapter 7

Conclusion – follow-up action required

Many of the findings of this review are not new. There is now a strong basis in evidence for actions that need to be taken to increase the attention given to gender issues through the activity cycle. Agencies need to take follow-up action to this review, both collaboratively and individually, and to ensure that the findings are disseminated to development practitioners.

1 Purpose and Background of the Review

Purpose

This report presents findings from a review on gender and evaluation undertaken during 2002 and early 2003 by AusAID and the Netherlands Ministry of Foreign Affairs. The review's purpose was to distil experiences with evaluating gender equality objectives, in order to improve evaluation practice and development outcomes. There were three areas of focus:

- **evaluation methodology** for assessing changes in gender equality, gender mainstreaming and women's empowerment
- **institutional approaches and change strategies** that have successfully responded to recurrent obstacles to gender mainstreaming, and
- improvements in **overall development outcomes** due to gender analysis and gender mainstreaming.

Background

Previous reviews undertaken by the DAC Working Party on Aid Evaluation

The review builds on previous work undertaken by the DAC on the integration of gender equality perspectives into evaluation methodology. That work included two reviews that focused on gender and evaluation and were undertaken in 1994 and 1999 by the Working Party on Aid Evaluation in collaboration with the Working Party on Gender, a workshop on evaluating gender equality and women's empowerment undertaken in 1999 by these two DAC working parties, and publications of the Working Party on Gender that address gender equality perspectives in evaluation (OECD 1998, 1999).

The 1994 review (OECD 1994) was undertaken as a contribution to the 1995 United Nations 4th World Conference on Women held in Beijing. One of three themes was an assessment of women in development as a crosscutting issue in evaluations. Conclusions were that limited attention had been given to women and gender issues in the evaluation studies reviewed, and that the evaluations themselves focused on outputs rather than impacts because agencies were only just beginning to develop methodologies for impact assessment. On the positive side, there had been a steady increase in the attention given to gender issues in evaluations since 1989. Recommendations were that agencies should give more attention to gender issues in TORs for evaluations, adopt more participatory approaches in activity design, and take action to improve the overall quality of evaluations.

The 1999 review (Woodford-Berger 1999) was designed to assess whether improvements had been made in evaluation design and implementation since the first review (1994).

Sixteen agencies used a reporting guide to assess their own work since 1993, and three DAC members (Canada, Sweden and Australia) undertook desk studies of around 40 evaluation reports each. The conclusion of the 1999 review was that uneven progress had been made. While there was some increase in the quality and quantity of information on gender issues, the variation within and between agencies was considerable. TORs were addressing gender issues more frequently, but the depth and relevance of discussions on gender issues in evaluation reports had not improved much. The review recommended that agencies improve their evaluation methodologies by using more sex-disaggregated data and making more consistent use of participatory approaches, and take action to improve evaluation capacity among both donors and partners. The follow-up workshop to the review held in 1999 in Stockholm made similar recommendations.

The methodologies of the 1994 and 1999 reviews are compared with the current review's methodology in Chapter 3. The findings of these reviews are compared with those of the current review in Chapter 7.

The current review

In 2001 Australia presented to the DAC Secretariat preliminary ideas for possible follow-up to the 1999 workshop. This was followed by a DAC Secretariat Discussion Paper to the Working Party on Aid Evaluation, which identified the three areas of enquiry for the current review (Development Co-operation Directorate 2001). The DAC Secretariat discussion paper envisaged specific outputs for each of the three areas of enquiry. Australia developed draft TORs for the current review (see Annex A) and AusAID contracted a consultant (Juliet Hunt) to begin work on the review in March 2002, after receiving input on the proposed methodology and TORs from the Working Party on Aid Evaluation's reference group for the study.

A progress report (including a summary of methodology, draft database, preliminary findings and request for further reports) was discussed at the Working Party on Aid Evaluation's 36th meeting in May 2002. The Netherlands Government offered financial support for the review following that meeting, and a consultant (Ria Brouwers) was appointed from the Netherlands in August 2002. A further progress report (including an updated database) was circulated at the Working Party on Gender's October meeting, which was attended by the Netherlands consultant. A draft of this report was circulated to the reference group for the review, and comments from members were incorporated. A revised draft report was discussed at the 37th meeting of the Working Party on Aid Evaluation in March 2003, and further comments from members were incorporated in this final version of the report.

The "Summary and Recommendations" (page 10) include key findings and recommend follow-up work based on the findings of the review.

2 Database for the review

Features of an evaluation

An evaluation is defined by the OECD as:

... a systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors. (OECD 2002, pp. 21-22)

In evaluations, a commitment to gender mainstreaming requires that both their content and methodology are gender sensitive. Potential differences in results for males and females need to be systematically investigated, along with the impact of activities on gender relations and gender equality, if agencies wish to have credible information on how development assistance affects people's lives, and to provide lessons for future assistance. Differences in results for women and men may significantly affect the long-term sustainability of development efforts. Equally, the objectives, efficiency and effectiveness of activities may have significant implications for gender equality.

Clearly, the weight given to gender equality compared with other issues in evaluations varies from one activity to another, and depends on a range of factors including evaluation objectives, scope and resources. The DAC has pointed out that questions about gender equality must be adapted for each activity to be evaluated, taking into account what the agency needs to learn from the evaluation for planning purposes. This requires explicit and feasible direction in evaluation TORs (OECD 1998, p. 37).

Thematic and general evaluations

The database for the review includes two types of evaluation of development activities carried out by DAC members and other agencies between 1999 and 2002:

- **thematic evaluations** specifically designed to evaluate gender equality, gender mainstreaming and women's empowerment efforts, and
- **general evaluations** that include some gender analysis or references to gender issues but were not specifically designed to evaluate gender equality policy or initiatives.

For general evaluations, agencies were requested to submit reports from 1999 to 2002 that included significant gender analysis and addressed at least one of three selection criteria. These criteria were defined by the three areas of focus for the review:

- the use of innovative, practical and effective evaluation methodology for evaluating changes in gender equality, gender mainstreaming and women's empowerment

- analysis of institutional blocks to gender mainstreaming and of strategies or experience that successfully dealt with these institutional blocks, and
- evidence or analysis that makes a link between the inclusion of gender analysis or gender perspectives at any stage in the activity cycle and development outcomes.

A total of 85 evaluations were included in the review – 42 thematic evaluations from twelve DAC members and five multilateral agencies (the World Bank, the Asian Development Bank, the United Nations Development Fund for Women (UNIFEM), the International Labour Organization (ILO) and the World Food Program); and 43 general evaluations from ten DAC members (see Annex B, Table B1 and Table B2). Eight DAC members submitted both thematic and general evaluations. Of the total of twenty-two DAC members,¹ fourteen had evaluations included in this review. With few exceptions, only reports produced in English were reviewed.

One feature of both the thematic and general evaluations included in the review is the high degree of diversity regarding the type of assistance evaluated, the number and size of activities evaluated, the scope of the evaluations, their objectives, and the different country contexts in which the evaluations had been undertaken. In most cases the general evaluation reports came from the evaluation sections of donor agencies. In other cases, the general evaluations originated in program sections.

The initial database was identified from various sources. The Working Party on Aid Evaluation's Planned Evaluation Matrix (November 2001) was the first source for thematic evaluations; its Evaluation Inventory (accessed February 2002) was the first source for the general evaluations. Initial research from these two sources indicated that 21 thematic and up to 66 general evaluations might have been available. However, that database differs considerably from the final database (see Annex B) for a number of reasons.

- Some thematic evaluations did not go ahead as planned.
- A number of evaluations were deleted from the database because a large thematic meta-evaluation that distilled findings and lessons from other reports had been undertaken (see World Bank 2001, 2002).
- Agencies were asked to screen general evaluations to ensure that they included some gender analysis and contained material on at least one of the three areas of focus for the review. The methodology assumed that the selection criteria would yield a database of general evaluation reports with significant quality and quantity of gender analysis. This assumption was correct for only 40% of the sample of general evaluations.

¹ DAC members are **Australia**, Austria, **Belgium**, **Canada**, **Denmark**, Finland, France, **Germany**, Ireland, Italy, **Japan**, Luxembourg, the **Netherlands**, **New Zealand**, **Norway**, Portugal, Spain, **Sweden**, **Switzerland**, the **United Kingdom**, the **United States**, the **Commission of European Communities**. (Bold indicates that the review included evaluation reports from those members.)

- Some agencies did not submit reports at all despite a number of requests, and it was necessary to search their websites. However, searches were limited to only a few agencies due to time constraints.
- The requirement that the general evaluations have significant gender analysis was revised to have some content pertaining to gender issues. However, reports that did not meet any of the three selection criteria or that had the most superficial references, such as a few lines only, were culled from the sample.

In summary, the final sample of general evaluations represents evaluations with the highest quality and quantity of gender analysis that were made available to the reviewers by participating agencies, in addition to other reports independently accessed by the consultants from agency websites. This is reflected in the sectoral distribution of general evaluations (see Table 1), which are overwhelmingly in social and community development areas where the importance of assessing differences in impact on males and females is now rarely contested.

Significant effort was devoted by the consultants to ensure that the database was as comprehensive as possible. For example:

- The total number of reports assessed by the consultants is much larger than the list in Annex B. More than 130 reports were screened to select the 85 evaluations in the sample.
- Selected agencies were followed up a number of times to ensure that as many agencies as possible were included in the review.
- The deadline for the submission of reports was progressively extended in an effort to ensure that all reports that included some gender analysis were included in the review.
- At the Working Party on Gender's meeting in October 2002 the Netherlands consultant asked gender specialists to forward any other relevant evaluation reports from their agencies, resulting in more thematic reports being included in the database.

Table 1 **Sectoral focus of general evaluations**

Sector	Project and cluster	Country, program or sector	NGO programs	Total
Agriculture, rural or community development	9 evaluations	2 evaluations	2 evaluations	13
Health	2 evaluations	4 evaluations	1 evaluation	7
Governance	2 evaluations	3 evaluations		5
Education	1 evaluation	3 evaluations		4
Private enterprise	1 evaluation	1 evaluation		2
Post conflict			2 evaluations	2
Land, environment	1 evaluation	1 evaluation		2
Water	2 evaluations			2
Infrastructure	1 evaluation	1 evaluation		2
Mixed		2 evaluations	2 evaluations	4
All	19	17	7	43

Of the total database of evaluations 70% (42 thematic reports and 17 general evaluations) contain significant material on gender analysis. This group of reports also provided significant information on the other two areas of focus – institutional blocks and links between gender analysis/perspectives and overall development outcomes. The total database was used to review evaluation methodology.

For many of the agencies that submitted reports, the database is probably representative of the evaluations that give some attention to gender issues. However, it is not possible to assess whether the database is representative of evaluations undertaken by all DAC members. At the end of November 2002 there were 661 evaluations included on the Evaluation Inventory for the period 1999 to 2002. A key word search of “women” and “gender” yielded 128 reports of both thematic and general evaluations that had some focus on gender issues. This suggests that about 19% of the total number of reports submitted to the inventory pay some attention to women or gender issues.²

Other material submitted to the review

During consultation with interested DAC members on the TORs for the review, it was suggested that material other than evaluation reports be included in the review. The other documents submitted or accessed by the consultants are also characterised by significant diversity in type and focus (see Annex B, Table B3 and Table B4). While all of these documents were reviewed, generally they have not been included in the analysis of findings in Chapters 4, 5 and 6. In the time available, and for the sake of coherent analysis, it was necessary for the consultants to focus on the evaluation reports.

The list of tools included in Table B3 is by no means comprehensive. Many include excellent advice on how to conduct gender-sensitive evaluations. One set of tools received very early in the review process was the *CIDA Evaluation Guide*,³ a work in progress that includes separate “How to Perform Evaluations” guides on various topics. One of these guides focuses on gender equality.⁴ Gender equality issues are also integrated into model TORs and a separate guide to participatory evaluations. CIDA evaluation staff offered to conduct interviews with selected gender specialists and program staff on the use of the

² However, the sample included in the review is not a clear subset of those 128 evaluations included on the inventory. While many reports reviewed do appear on the inventory, a lot of others in the final database do not, as explained when discussing the draft database, and because some evaluations undertaken during 2002 may have been submitted to the inventory, but not to the consultants conducting the review.

³ CIDA (2001e) in Table B3.

⁴ CIDA (2001i) in Table B3.

gender equality evaluation guide, and the main findings from this survey are included in Chapter 4.

Comparisons with previous reviews

The current review was explicitly designed to build on, rather than repeat, the work of previous DAC reviews on gender and evaluation. Whereas the previous reviews were designed to assess the attention given to women in development and (later) gender equality in **general** evaluations, this was only one element of the current review. While it has some similarities in focus with the 1994 and 1999 reviews, the current review departs from earlier reviews in significant ways by:

- including thematic evaluations on gender equality and women's empowerment (which did not exist at the time of previous reviews)
- restricting the sample of general evaluations to those that met pre-determined criteria, specifically that they included some gender analysis
- focusing on evidence about strategies that promoted the institutionalisation of gender equality commitments (the second area of focus for the review), and
- focusing on the links between the benefits for women, gender analysis and gender mainstreaming efforts, and overall development outcomes (the third area of focus for the review).

The focus in the database for this review was designed to yield reports that could be expected to include lessons on institutionalisation and development outcomes, and to include valuable experience with evaluating gender equality objectives. Earlier reviews had been based on representative samples of agencies' evaluations.

The size and type of the database also varied widely between the three reviews. The 1994 study included 1315 evaluation reports from 1989 to 1993 from 16 donors. This was a well-resourced study designed to contribute to the United Nations 4th World Conference on Women. In contrast, the 1999 review used a reporting guide that agencies used to review and assess their own evaluation work and any improvements since 1993. From a total of 31 reporting guide questionnaires distributed to DAC members and other agencies, the 1999 review received 16 completed reporting guides from fifteen DAC members and one observer. Thirteen of these agencies had also participated in the 1994 review. In addition, three agencies (in Sweden, Canada and Australia) undertook internal assessments of a total of 121 evaluation reports (Woodford-Berger 1999, para. 29-48). For comparison, the current

review's database was 85 evaluations from fourteen DAC members and five multilateral agencies. Only twelve DAC members participated in both the 1999 and current reviews.⁵

⁵ The DAC members that participated in both the 1999 and current reviews included Australia, Belgium, Canada, Denmark, the Commission of the European Communities, Japan, the Netherlands, Norway, Sweden, Switzerland, the United Kingdom and the United States. Finland, Portugal and Spain participated in the 1999 review only. Germany and New Zealand participated in the current review, but not the 1999 review.

3 Review framework

Comprehensive and systematic qualitative assessments were undertaken of each evaluation report in the database. A draft data sheet was developed to guide the assessment of each evaluation report in relation to the three areas of focus. For example, areas of enquiry in the data sheet on evaluation methodology focused on:

- innovative approaches to evaluating changes in gender equality, gender mainstreaming and women's empowerment
- issues and weaknesses in evaluation methodology identified in previous reviews, and
- the DAC's existing guidance on how to apply gender concepts to monitoring and evaluation (OECD 1998, pp. 36–39; 1999, p. 24).

The draft data sheet was modified by the AusAID consultant in response to comments on the TORs for the review from the reference group for the study. The data sheet was tested using a small sample of both thematic and general evaluations. It was further modified after the Netherlands consultant joined the study, and again after receiving feedback on the third area of focus from the Working Party on Gender's meeting in October 2002. The design of the final data sheet facilitated the quantification of key findings.

The framework for the review is summarised in Table 2, and is explained in this chapter according to each of the review's three areas of focus.

Evaluation methodology

The data sheet investigated a very wide range of questions on evaluation methodology used in the evaluations.

The **scope and focus** of each evaluation were summarised in a descriptive section of the data sheet that recorded: the type of evaluation conducted; country/region; sector; objectives of the activities evaluated; budget details; phase of the activity cycle when the evaluation was conducted (implementation, completion, ex-post); which of the three selection criteria were addressed; whether/how gender equality objectives were addressed in the design and implementation of the activity; and a summary of the major findings of the evaluation relating to gender equality.

Due to the diversity in the sample, two descriptive features emerged as most critical for managing and analysing the findings. These were the category of evaluation (whether policy, program, project, cluster, sector or NGO program/project) and its focus (whether activity results, the institutionalisation of gender equality policy commitments, or program/project management issues).

Table 2 **Summary of the framework used for reviewing evaluation documents**^a

Area 1 Evaluation methodology	Area 2 Institutional approaches & change strategies	Area 3 Benefits for women, gender analysis & development outcomes
<ol style="list-style-type: none"> 1. Scope & focus of the evaluation 2. Features of evaluation methodology <ul style="list-style-type: none"> – Description of tools, frameworks, processes, sampling methods – Details of participatory & social analysis methods used – Assessment of tools, frameworks, approaches 3. Quality & quantity of information & analysis <ul style="list-style-type: none"> – Use of gender-sensitive indicators – Collection of quantitative & qualitative sex-disaggregated information – Content/depth of gender analysis undertaken (including use of gender analysis concepts) – Attention to partner capacity, commitment & dialogue on gender equality – Attention to resource requirements for gender-sensitive development 4. Evaluation TORs <ul style="list-style-type: none"> – Attention to gender analysis information & analysis requirements (potential gender issues; reference to gender-sensitive indicators) – Attention to gender expertise & use of local consultants – Attention to partner &/or stakeholder participation & consultation 5. Standard evaluation criteria & principles <ul style="list-style-type: none"> – Use of evaluation criteria – Application of gender equality & women's empowerment concepts to evaluation criteria (relevance, achievement of objectives, impact, development results, effectiveness, sustainability, alternatives, efficiency, lessons learned) 	<ol style="list-style-type: none"> 1. Recurrent institutional blocks identified in evaluation reports 2. Successful strategies based on evidence & findings in evaluation reports <ul style="list-style-type: none"> – What has been successful & why? – What has not been successful & why? 3. Comments, other learning & recommendations on strategies identified in evaluation reports 4. All the above identified according to level of institutionalisation: <ul style="list-style-type: none"> – Agency policy – Policy dialogue between donors & partners – Donor country assistance strategies – Donor organisations – Partner organisations – Activity design – Activity implementation – Activity monitoring & evaluation – Progress towards gender equality at the community level 	<ol style="list-style-type: none"> 1. Evidence of benefits to women &/or changes in gender relations <ul style="list-style-type: none"> – Does the evaluation specify how women/men have benefited from activities? – Were there differences in benefits for women & men? – If so, how were these differences addressed or explained? 2. Evidence of a causal relationship between attention given to gender issues and benefits for women <ul style="list-style-type: none"> – Is there evidence that attention given to gender issues in design or implementation leads to benefits for women, or to changes in gender relations? – Is there evidence that activities directly targeting women produce strategic gender equality results for women? – Is there evidence of other factors being important in producing benefits for women? 3. Evidence that benefits for women or changes in gender relations lead to improved development outcomes at: <ul style="list-style-type: none"> – micro level – macro level

^a This framework is a simplified version of the review methodology and includes revisions made to the methodology during the review process.

The **features of the evaluation methodology** recorded included: a summary of the purpose of the evaluation; descriptions of tools, frameworks and processes used to collect and analyse data; details of participatory and social analysis methods employed (including the types of stakeholders consulted); sampling methods; and an assessment of whether any tools or frameworks could be included in a “toolbox”.¹ This section of the data sheet noted whether local partners, evaluators or gender specialists had been involved in the evaluation, and whether there was an explicit intention to strengthen local counterparts. An assessment was also made of whether evaluators were faced with a lack of sex-disaggregated baseline and monitoring data on gender issues, and what methods were used to collect data in such circumstances.

The **quality and quantity of information and analysis** recorded on the data sheet included: the use of gender-sensitive indicators; the collection of sex-disaggregated quantitative or qualitative information; attention given to partner capacity and dialogue on gender equality; and attention given to resource requirements for gender-sensitive approaches to development. A number of questions were asked to assess the quality and depth of analysis of gender equality issues. These included: whether sex-disaggregated information was analysed or merely presented; whether gender analysis was integrated with overall social analysis of the activity and its results; the extent to which international standards for gender analysis were employed in the collection and analysis of data; whether the evaluation assessed the participation of women in the activity; whether the evaluation assessed women’s share of activity benefits; and whether the evaluation assessed the equality of results for women and men.

The international standards for gender analysis included on the data sheet were: gender division of labour; access to and control over resources; understanding of women’s and men’s different needs; understanding of constraints or barriers to male or female participation; strategies for working with females and males; practical needs or strategic interests identified or met; and changes in gender relations, including decision making and leadership.

The data sheet recorded whether the evaluation collected or analysed such information in its assessment of gender equality issues and objectives. It was not assumed that all of the above gender analysis concepts would be relevant for every activity evaluated. The consultants’ assessment was based on the content on gender issues in the evaluation reports, not on the use of gender jargon.

¹ A “toolbox” or discussion note on innovative approaches and tools was envisaged as a possible output from this area of enquiry in the DAC Secretariat discussion paper. See Development Co-operation Directorate (2001).

On **evaluation TORs** the data sheet requested information on: the attention given to sex-disaggregated information and gender analysis requirements; the need for gender or social analysis expertise and the use or strengthening of local consultants; and the attention given to partner and stakeholder participation and consultation (male and female) in the evaluation process.

The use of **standard evaluation criteria and principles** in each evaluation was recorded, as was the application of gender equality and women's empowerment concepts to those criteria (relevance, achievement of objectives, development results and impact, effectiveness, efficiency, sustainability, alternatives, and lessons learned).

Many evaluation reports addressed few of the issues and questions included on the data sheet, although this varied considerably both between and within agencies. For example, in some cases reports did not include TORs and/or there was little detail about the evaluation methodology employed. This applied to both thematic and general evaluations, although most reports on thematic policy evaluations included significant details on their methodology.

Using the information collected on the data sheets, an assessment was made of the quality of gender analysis in **general evaluations**. Three ratings were defined.

- **Satisfactory attention given to gender issues** The reports provided at least one piece of information on how women benefited from development activities, or they made at least one statement about the impact of activities on women. This rating also covers reports that included some information or an assessment of changes in gender relations, or that assessed differences in benefits and impacts for females and males.
- **Unsatisfactory attention given to gender issues** The reports assessed the participation of women in activities in a limited way, but did not provide information on how women benefited or about the impact on women.
- **Highly unsatisfactory attention given to gender issues** The reports did not assess women's participation or the benefits or impact on women, despite references to women or gender issues in the text. This may have been due to scarce sex-disaggregated data and poor activity design and monitoring. Some of these reports noted that more gender analysis was needed.

In all cases, these ratings were interpreted generously. Wherever there was any doubt about which rating to apply (due to diversity in the type, focus and objectives of the evaluations) the consultants gave the report the higher rating.

Taking into account the DAC's definition of evaluation (see p.20) the policy commitments of both the DAC and donor agencies to gender mainstreaming throughout the activity cycle, and the selection criteria for general evaluations included in the review,

the rating of “satisfactory attention given to gender issues” in general evaluation reports may be seen as a minimum requirement. Ideally, such a rating would also require that any differences in the results of an activity for men and women be investigated. With few general evaluations focusing on gender relations, as distinct from women (see Chapter 4), this would have been unrealistic.

If the DAC undertakes any further reviews of the attention given to gender issues in evaluations, the definition of “satisfactory” used in this review could provide a useful benchmark for measuring progress, taking into account the fact that the current sample includes only those general evaluations where agencies had an expressed intent to assess gender issues in some way. Without such a benchmark, accurate comparisons of progress between reviews are difficult to make (see the section “Comparisons with previous reviews” on page 32).

A questionnaire was developed and tested for follow-up on methods and tools used by evaluators in **thematic evaluations**. This questionnaire was used as a general guide only and was modified extensively for each agency according to the scope and objectives of its evaluations and depending on whether follow-up occurred by phone or by email. Both staff and consultants engaged in conducting evaluations for the following agencies were contacted:

- Asian Development Bank (ADB) – 1 interview
- AusAID – 5 interviews and/or questionnaires
- Belgium – 1 interview
- CIDA – 13 questionnaires distributed and collated by staff from CIDA’s performance review branch
- Oxfam Community Aid Abroad (OCAA) – 6 questionnaires
- Sida – 2 interviews, and
- World Bank – 2 interviews.²

Institutional approaches and change strategies

For the second area of focus, the data sheet investigated:

- institutional obstacles to gender mainstreaming (widely interpreted as obstacles to having gender issues addressed)

² See Table B1 in Annex B for the reference details of all evaluation reports except that of the OCAA, which is referenced in Table B4.

- successful strategies for addressing institutional and other obstacles to gender equality, and
- recommendations and other comments made in the evaluation reports about actions required to ensure that gender issues are properly addressed.

A matrix was used to record findings from each evaluation. It focused on the following levels of institutionalisation in relation to the obstacles, strategies and recommendations mentioned above: agency policy; policy dialogue between donors and partners; donor country assistance strategies; donor organisations; partner organisations; activity design; activity implementation; activity monitoring and evaluation; and progress towards gender equality at the community level.

There is an extensive literature on institutional obstacles to gender mainstreaming, and there are many frameworks and strategies to address these. However, good practice case studies on the institutionalisation of gender equality approaches are often based on intention rather than outcome, and few are based on evidence from activity evaluations. While previous DAC reviews addressed the issue of institutionalisation, the value added by the current review is in the evidence collated from the evaluations on what worked. This is a new perspective on this area of work. Findings are presented in Chapter 5.

Benefits for women, gender analysis and development outcomes

The original research question for this area of focus was to assess whether gender analysis and gender mainstreaming improve overall development outcomes. This area of the review presented the most significant methodological challenges, which are outlined in Chapter 6.

The review's TORs indicated that, for this area of enquiry to yield significant findings, donor agencies would need to identify evaluations that linked the inclusion of gender perspectives in an activity's design and implementation with successful or improved development outcomes. Very few evaluations explicitly addressed these links at the macro level of development outcomes. As noted in the TORs (see Annex A) the size of the database for this area of enquiry was unknown when the review commenced. Being a meta-review, the consultants were unable to conduct any research using primary sources to more closely examine the links between gender analysis and mainstreaming and development outcomes. Furthermore, some members were interested in achievements related to the Millennium Development Goals.³ There were very few references in the evaluations to these goals, and no findings in this area are reported in Chapter 6.

³ For details on the Millennium Development Goals, see <http://unstats.un.org/unsd/mi/pdf/mdglist.pdf>.

As a result of these difficulties, changes were made to the review framework, and questions on the data sheet were reformulated as the review progressed. This was to ensure that all relevant information relating to links between the benefits for women and men, the attention given to gender issues, and development outcomes could be gleaned from reports.⁴ This refinement of the review methodology enabled some interesting and useful findings to be extracted from the evaluation reports (see Chapter 6). It also pointed to the need for more research in this area.

The final research questions for this area focused on:

- evidence of benefits for women and/or changes in gender relations, including whether the evaluation specified how women and men benefited, and whether there were differences in the benefits for women and men
- evidence of a causal relationship between the attention given to gender issues and the benefits for women, and
- evidence that benefits for women or changes in gender relations led to improved development outcomes at either the micro or macro level.

Comparisons with previous reviews

The major area of similarity in the frameworks of the current review and the 1994 and 1999 reviews is the focus on evaluation methodology used in general evaluations, which is the subject of Chapter 4. However, direct comparisons with the 1999 review on evaluation methodology are difficult. This is because findings from the 1999 review are quantified according to agency (whereas this review quantifies findings according to evaluation report) and because findings from the 1999 review were inconclusive in some areas.

One example illustrates this point. It is unclear whether participatory methods for including beneficiaries in general evaluations are more common now than in 1999. Chapter 4 reports that 40% of general evaluations included participatory approaches with beneficiaries. For the remainder, many were not expected to include consultation or participation with beneficiaries because of their limited scope and focus, and for others there was simply inadequate information in the evaluation reports to assess whether a participatory approach had been taken. Another confounding factor is that both reviews were based on a sample of reports coming from only about half of the DAC members. Of the twelve agencies that addressed this issue in the 1999 review, nine reported that they

⁴ This area of the review was discussed at the October 2002 meeting of the Working Party on Gender and concern was expressed about the preliminary findings reported to the meeting. Although some changes had already been made to the methodology for this area, the data sheet for the review was refined further following this meeting in response to the concerns raised.

used participatory approaches in most evaluations, leading Woodford-Berger (1999, para. 88) to conclude that the 1999 review “did not provide conclusive evidence on these issues”.

4 Findings on Evaluation Methodology

All of the thematic and general evaluations in the sample were reviewed to identify lessons about the specific methodology the evaluators used and, in the case of general evaluations, to learn about the extent to which gender issues were integrated into the overall evaluation methodology.

For the purpose of managing the review, both the thematic and general evaluations were divided into different categories – for example, policy, cluster or project (see Table 3 and Table 4). Where these categories yielded significant differences in findings, these are explicitly referred to in the following analysis; otherwise the findings cover the whole sample.

Category of evaluations

Thematic evaluations

The 42 thematic evaluations were categorised as individual project evaluations, evaluations of gender equality funds, clustered projects evaluations and studies, and evaluations of gender equality policy (see Table 3).

The evaluations of **individual projects** and of **gender equality funds** dealt with activities focused on women and/or gender equality exclusively. Their usual aim was to compare performance and achievements with the stated objectives and/or to make recommendations for the future. While most of the evaluation reports specified the aims of the evaluations, they seldom made clear why the evaluation was undertaken, or how it would be used and by whom.

The **cluster** evaluations either examined the results of a group of activities for women and/or gender relationships or they looked particularly at the dynamics of gender mainstreaming, in order to enhance understanding of the issues by activity and policy staff and to help them improve their work in these areas.

Policy evaluations were broader than the cluster studies. Often they had the double function of making judgements about gender equality features in the agency's programs and of yielding lessons for the future. Some of the evaluations were set up primarily to assess the results of gender equality policy;¹ several others had the character of an audit,

¹ See Netherlands (1998), Sida (2002a), World Bank (2002) and ADB (2001) in Table B1.

Table 3 **Category and focus of thematic evaluations**^a

Focus ^b	Project	Gender equality fund	Cluster	Policy	Total
Primarily results	CIDA (1999a)	CIDA (2000b)	Belgium (2001)		13
	CIDA (1999b)	CIDA (1999c)	JICA (2001)		
	NZAID (2002a)	USAID (1999c)	JICA (2000)		
	Sida (2000a)		USAID (2001b)		
	USAID (2001a)		USAID (1999d)		
Primarily institutionalisation of gender equality policy commitments	CIDA (2002a)		BMZ (2001b)	AusAID (2001) ^c	16
	NZAID (2002b)		DFID (2000a)	CIDA (2001a) ^c	
	SDC (2000)		DFID (2000b)	ILO (2002) ^c	
	Sida (2002c)		DFID (2000c)	NORAD (1999b)	
			DFID (1999a)	UNIFEM (2002) WFP (2002)	
Both results & institutionalisation			BMZ (2000)	ADB (2001) Netherlands (1998) Sida (2002a, 2003) World Bank (2001a, 2002)	4
Management issues	CIDA (1999d) NORAD (2001) NORAD (2000) NORAD (1999a) USAID (1999a) USAID (1999b)	CIDA (2000a) Sida (2001a) Sida (2000b)			9
Total	15	6	11	10	42

^a For reference details of evaluation reports, see Annex B. ^b The focus on results was minimal in a number of evaluations. For explanations of the terms "institutionalisation" and "management", see "Terminology Used" (p. 9). ^c Gender audit.

with groups of staff assessing the processes and procedures for gender mainstreaming applied in the organisation.²

General evaluations

The 43 general evaluations included 19 in the project and cluster category, 17 in the country, program or sector category, and 7 in the NGO programs category (see Table 4). The sizes and scope of the evaluations varied considerably, ranging from those that addressed only one activity to those that covered a few hundred activities in a particular sector.³ The reasons why the evaluations addressed gender issues also varied. Sometimes

² See AusAID (2001), CIDA (2001a) and ILO (2002) in Table B1.

³ See BMZ (2001a), Danida (2000a) and DFID (2001b) in Table B2. The large sector- or program-wide studies generated subreports, which were not part of this gender review. Whatever specific information they may have about a gender focus in their data collection is in this review only to the extent that it is included in the main report.

Table 4 **Category and focus of general evaluations**^a

Focus ^b	Project & cluster	Country, program or sector	NGO programs	Total
Primarily results	AusAID (2002b)	AusAID (2000e)	AusAID (2000a)	18
	AusAID (2000b)	CIDA (2001b)	AusAID (2000g)	
	AusAID (2000c)	SDC (1999b)		
	AusAID (2000d)			
	AusAID (2000f)			
	AusAID (1999a)			
	AusAID (1999b)			
	AusAID (1999c)			
	AusAID (1999c)			
	CIDA (2000h)			
	SDC (1999a)			
	SDC (1999d)			
	SDC (1999e)			
Primarily management issues	CIDA (2000c)	NORAD (1999c)	CIDA (2000f)	7
	SDC (1999c)		CIDA (1999e)	
			CIDA (1999f)	
			CIDA (1999g)	
Both results & management issues	CIDA (2000d)	BMZ (20001a)	CIDA (2000e)	18
	CIDA (2000g)	Danida (2002a)		
	Sida (2001b)	Danida (2002b)		
	Sida (2000c)	Danida (2001a)		
		Danida (2001b)		
		Danida (2000a)		
		Danida (2000c)		
		Danida (1999)		
		DFID (2002)		
		DFID (2001b)		
		EC (2001a)		
		EC (2001b)		
		Netherlands (2002)		
Total	19	17	7	43

^a For reference details of evaluation reports, see Annex B. ^b For an explanation of the term "management", see "Terminology Used" (p. 9).

the activity had an explicit objective on the participation of women and/or gender equality; sometimes gender issues were taken into account because gender is seen as a crosscutting policy issue in many agencies.

The amount of information on methodology included in the reports of general evaluations also varied. Several provided little or no information on aspects of methodology, so that it is unknown how the studies measured benefits for women and men, and what indicators and criteria were used to assess progress towards gender equality. Others were fairly complete, including TORs, information about the involvement of stakeholders and beneficiaries in the evaluation process, and information on data collection methods, indicators and criteria. Most of the following information comes from the last group,

although efforts have been made to derive information about the methodology used by the first group as well, on the basis of the findings in the reports.

Focus of evaluations

The 1999 gender review found particular weaknesses in the assessment of development results. Has this changed? How wide was the focus of the evaluations in the current review? Did they focus primarily on institutional and managerial arrangements, did they address mainly results for beneficiaries, or did they cover all of these aspects?

Thematic evaluations

The focus of the thematic evaluations is summarised in Table 3. Two-thirds (10) of the 15 **project** evaluations dealt with institutional and managerial aspects. This was not only the case when the object of evaluation was an institution. It also applied for activities intended to improve the lives of women and men. This meant that opportunities to assess the results of the activities for ultimate beneficiaries were underused. In the evaluations that did address benefits, the information provided was mostly about activities (for example, number of women trained) and seldom about results.

The focus of the **fund and cluster** evaluations was more balanced. Half focused on the implementation and results of the interventions. Those focusing on institutional or managerial aspects were explicitly designed to examine the gender specificity in policy and planning, procedures, resources, monitoring and evaluation processes.

The **policy** evaluations presented a varied picture. Four gender policy evaluations focused on both results and institutional aspects.

- The core focus of the Sida (2002a) evaluation was whether the mainstreaming strategy had influenced gender equality results, with emphasis on interventions in three selected countries.
- The World Bank (2001a) study had an institutional focus and looked at the gender analysis of the Bank's analytical work and gender integration in country assistance strategies. The World Bank (2002) study focused on results, taking the country as the unit of evaluation (12 countries had been selected), and evaluated the extent to which the Bank had helped to reduce gender disparity in health and education sectors, had helped to increase the participation of women in economic activities, and had influenced institutional changes that support the advancement of women.
- The Netherlands (1998) evaluation also focused on results for women, analysing two country programs for this reason, in addition to assessing the policy and institutional structure at headquarters and embassies for their contribution to these results.

- The ADB (2001) evaluation addressed both institutional aspects and the results for beneficiaries in three countries.

The focus of all other evaluations in the policy category in the sample was primarily on institutionalisation, which means that they assessed the extent to which gender equality concerns were integrated into policy, strategy and procedures for planning and implementation. Some of the evaluations were gender policy audits.

General evaluations

The focus of the general evaluations is summarised in Table 4. It shows that the majority of evaluations in the project and cluster category focused on results, although this does not necessarily mean that they focused on results related to gender equality, or on the benefits for women. The majority of the NGO evaluations focused on management issues, while the evaluations in the country, program or sector category tended to focus on both results and management issues.

Quality of gender analysis in general evaluations

Seventeen evaluations (40%) in the sample of general evaluations gave satisfactory attention to gender issues (see Table 5). These evaluations assessed the participation of women in interventions. They also provided at least one piece of information on how women had benefited from the activities evaluated, or they made at least one statement about the impact of activities on women. A few also included some information or assessment of changes in gender relations. (Annex E illustrates some of the features of evaluations that facilitated gender analysis.)

Eleven evaluations (26%) have unsatisfactory gender analysis. These evaluations assessed the participation of women in interventions in a limited way. They did not provide information on how women benefited, nor about the impact on women.

Table 5 **Assessment of the attention given to gender analysis in general evaluations**

Satisfactory attention given to gender issues	Unsatisfactory attention given to gender issues	Highly unsatisfactory attention given to gender issues
6 AusAID	4 AusAID	2 AusAID
1 BMZ	3 CIDA	4 CIDA
3 CIDA	1 Danida	1 Danida
5 Danida	1 Netherlands	2 EC
2 DFID	1 NORAD	4 SDC
	1 SDC	2 Sida
17	11	15

Note: Definitions of satisfactory, unsatisfactory and highly unsatisfactory are included on page 29. The total sample of general evaluations is 43.

Fifteen evaluations (35%) are highly unsatisfactory in their treatment of gender issues. These evaluations did not assess women's participation, their benefits or the impact on women. In some cases, the reports made superficial reference to women or gender and presented little or no data even though their stated intention was to assess the benefits for women. They sometimes described agency policy commitments to gender equality, or suggested that more gender analysis was needed.⁴

Inevitably, as with any rating system based on qualitative assessments, there were evaluations that were difficult to rate because of their diversity in category, focus and objectives. As already noted in Chapter 3, for those cases, the consultants erred on the side of generosity and gave a higher rating to the evaluations.

Also, the ratings in Table 5 should not be seen as representative of agencies' overall evaluations, because of the selection criteria used for the review sample. Agencies were requested to submit evaluations that addressed gender issues; evaluations that paid no attention to gender issues were culled from the sample.

Methods, approaches and tools

How were the evaluations conducted? Were they desk reviews, were field studies carried out, and were stakeholders and beneficiaries involved? Were frameworks developed, criteria and indicators applied, and special tools used or developed?

Thematic evaluations

The evaluations in the **project** category and in the **gender equality fund** category were usually conducted in a standard way. That is, desk studies were combined with brief field visits (of 1 to 2 weeks) to interview stakeholders, mainly project and counterpart staff, and to meet with direct primary beneficiaries. In general, the studies were not carried out very systematically; evaluation criteria such as relevance, effectiveness, efficiency, impact and sustainability (standards mentioned in the DAC definition of evaluation) were applied in an *ad hoc* way. Only 4 of 15 project evaluations made explicit use of gender-sensitive indicators. (See Annex C for a summary of the types of indicator used.) Others may have used indicators in the evaluation process, but these were not documented. Beneficiaries participated in a little over one-third of the evaluations, participation usually meaning that beneficiaries were consulted. A more active form of participation was found in one CIDA (1999a) project evaluation, where partner agencies had direct input into recommendations based on problems and achievements identified by beneficiaries and other stakeholders.

⁴ In some cases the evaluations were constrained by scarce sex-disaggregated data due to poor activity design and monitoring, or the evaluators noted that there was inadequate time to deal with all evaluation issues.

The methods used in the 11 evaluations in the **cluster** category varied considerably. Those initiated to study experiences with mainstreaming were mainly desk studies, in incidental cases complemented by interviews with staff. On the other hand, the Belgium (2001) evaluation was designed to be very participatory, with a high level of involvement of primary stakeholders. Many of these evaluations deviated from what they set out to achieve, mainly because of the complexity of their designs, which were based on the assumption of causality between changes in gender relations and poverty reduction. The USAID (1999d) evaluation of the education sector in multiple countries is a good example of a well-designed study, which systematically applied gender-sensitive criteria and indicators, contrary to many of the other studies in this category.

The bigger **policy** evaluations are the most interesting in terms of methods and tools. The World Bank, Sida, ADB and Netherlands evaluations were all systematically laid out, with analytical frameworks either for the whole evaluation or for parts of it. Gender-sensitive indicators and evaluation criteria were applied throughout. All of these evaluations included desk reviews, staff surveys, stakeholder consultations and participatory assessments of benefits in the countries involved. For the sake of participation, a range of tools were applied, varying from well-known participatory rural appraisal (PRA) tools to separate village studies (that is, outside the context of project evaluations) for the purpose of understanding the perceptions of the population on developments in their area. The participation of stakeholders and beneficiaries, mainly in the form of consultation, was part of the design of all four studies. The World Bank explained that this approach was a response to the lack of available sex-disaggregated data for assessment, making it necessary to collect data directly from beneficiaries during its evaluation. The success of a participatory approach depended largely on the available time and resources to carry out the studies; frustration was encountered when the evaluations had to fit into a regular two-week country field visit.

The qualification “interesting” used for policy evaluations does not automatically imply that the methods yielded good evaluation results. The design of the Sida (2002a) evaluation was too ambitious and complex to address its TORs.⁵ It was based on the idea (“hypothesis”) that gender mainstreaming would effect gender equality and that the activities selected would provide the lessons of how this worked. When the expected effects were hardly found (most of the activities had only nominally implemented the mainstreaming strategy, and several were mainly of an institutional nature, making it difficult to assess the direct results for women and men), the evaluation team sought lessons on processes for introducing and implementing gender equality mainstreaming, and ideas on how Sida’s gender equality policy could be strengthened and better integrated

⁵ See Sida (2003) in Table B1.

with the overall poverty reduction strategy. In hindsight, an alternative sample of activities might have better served the objectives of the evaluation.

Gender auditing is a new method of establishing accountability for gender concerns (see also Annex D). The features distinguishing a gender audit from an evaluation are its focus on mainstreaming policy commitments and its involvement of a large group of agency staff. Two of the three gender audits in the review's database (see Table 3) are examples of participatory self-assessments. The ILO audit was designed to promote learning by participants and to facilitate their ownership of the findings, in the expectation that this would lead to direct action on their part to improve gender mainstreaming. This is a major strength of the participatory approach to gender auditing, although it is too soon to assess whether gender audits will fulfil these expectations. A weakness of most gender audit methodologies is that links to development results are limited or absent.

Concerns about the lack of a gender equality focus in sector-wide approaches to development initiatives and in poverty reduction strategies led to the BMZ (2001b) and DFID (2000c) evaluations that examine the processes involved and the critical factors for women's participation in them. These evaluations are promising examples for promoting the gender equality focus in these new areas of the development agenda, provided the lessons are heeded.

General evaluations

With a significant number of reports on general evaluations providing little information about their evaluation methodology, it is impossible to quantify how many of these evaluations applied certain methods. Most of the evaluations rated as giving satisfactory attention to gender issues (Table 5) collected both quantitative and qualitative data. The quantitative data focused mainly on inputs and outputs such as the participation of women in training, the number of women in groups, the number of women's groups, or the number of loans provided. The qualitative data tended to focus on the benefits and/or impact.

Participatory methods were used to collect the qualitative data. At least 17 of the general evaluations (40%) involved the beneficiaries, while some others involved project staff and stakeholders. Most of the evaluations that had satisfactory gender analysis used participatory methods at some level. However, participatory methods were also applied by evaluations that had unsatisfactory gender analysis, indicating that these methods did not always lead to adequate attention being given to gender issues. Very few reports indicated whether the evaluations used separate data collection processes for women and men to ensure that women's perceptions of benefits and impacts were heard.

Only 7 evaluations (16% of the sample of general evaluations) explicitly reported that they used gender-sensitive indicators. (See Annex C for a summary of the types of indicator used.) This figure may under-report the use of gender-sensitive indicators in general evaluations, because the larger country, program and sector evaluations did not provide sufficient detail on methodology to assess whether indicators had been used. Not surprisingly, all reports that did include gender-sensitive indicators were rated as having satisfactory gender analysis. In many cases, these indicators were formulated at the evaluation stage (rather than during activity design or implementation), and therefore were not linked to monitoring and information systems. Moreover, it appears that most were not formulated in a collaborative manner between donors and partners. The indicators included:

- process indicators that focused on the institutionalisation of gender equality commitments through the activity cycle and in organisations
- indicators to measure the participation of women in development activities, and
- indicators to measure the benefits to women and changes in gender relations, including empowerment indicators.

While it is acknowledged that empowerment indicators may be challenging to develop, gender-sensitive indicators for any type of result were rare.

In most of the general evaluations, international standards of gender analysis⁶ were applied in rather *ad hoc* ways, and often key information relating to the activity or sector was missed. The satisfactory evaluations had more consistently applied standards that were relevant to the activity being evaluated.

Overall, the evaluations are characterised by poor analysis, even when some significant quantitative and qualitative data had been collected. Discussion of gender issues was often limited to considering the participation of women. There was little or no examination of unintended impacts for women, or any other group such as men and women living in poverty. However, to be fair, such examinations were beyond the scope of some institutional reviews included in the sample.

Some general evaluation reports in the sample included TORs. These TORs generally made only a vague reference to women or “gender”, usually among a long list of objectives or questions that the evaluators were required to address. Some referred to the agency’s gender policy; others required an assessment of whether “gender aspects have

⁶ As noted in Chapter 3, the consultants assessed the use and application of the following international standards of gender analysis in evaluations: gender division of labour, access to and control over resources, understanding of women’s and men’s different needs, understanding of constraints and barriers to male and female participation, strategies for working with females and males, practical needs or strategic interests identified or met, and changes in gender relations, including in decision making and leadership.

been integrated”. Most required some gender expertise on the evaluation team, but about one-third did not.

Special gender evaluation tool

As already noted in Chapter 2, CIDA has a dedicated tool to support evaluators, partners and staff to focus on gender equality, as part of the *CIDA Evaluation Guide*. For this review, CIDA offered to examine the application and use of the gender equality evaluation guide. It cautioned that the gender equality evaluation guide was published only 18 months prior to the examination, with minimal publicity. The results of a limited number of interviews with gender equality specialists (4) and program staff (9) showed that only two of the nine program staff had used the guide, four had not used it and three did not know it existed. Questions in the interviews about how to improve the focus on gender relations in evaluations evoked suggestions such as ensure a proper baseline on gender equality in project design, ensure gender expertise and understanding on the team, and ensure a better understanding of evaluation itself.

Expertise and quality

What kind of expertise was called for in the evaluations, and what was the level of involvement of partner agencies? Did evaluations assess the capacity of partners to work on changing gender relations? What was the quality of the basic material that evaluators had to work with? What was the quality of the evaluations submitted?

Information about those involved in the evaluations and about their **level of expertise** was far from complete. The available data suggested that local and international consultants were hired in the case of thematic evaluations carried out in the field. Some agencies were very explicit about ensuring an equal or larger share of local expertise in the teams. The use of local partners/consultants was less obvious in general evaluations. In rare cases, donors conducted “joint reviews” – reviews undertaken with representation of the local partner or with a consultant selected by the partner. There was no indication that donors were working towards more joint evaluation practices, either joint donor or joint donor–partner, as might be expected in line with the changes in development practices, such as sector-wide approaches and emphasis on national ownership.

The evaluation reports did not provide information about the experiences with **local expertise** in evaluations. Hence, some interviews conducted for this review addressed this issue. Answers showed that local expertise was considered indispensable from the point of view of both the quality of the evaluation and its reliability and accuracy. However, having people from different backgrounds work together on an evaluation required careful and skilful preparation and training in collecting appropriate information on gender issues and

impacts, and on participatory methods. In turn this preparation and training required adequate resources of time and money, which needed careful budgeting. The World Bank and the Netherlands also considered the involvement of local evaluators to be a capacity-building exercise. The ADB found that local sector specialists were better performers than gender specialists without sectoral knowledge.

Many thematic evaluations did assess the **capacity of the partner** to work on changing gender relations. All of the evaluations in the gender equality fund category did, almost all in the cluster category did, as did two-thirds of those in the project category, and a large proportion in the policy category, including all of the overall policy studies. In contrast, only one-third of the general evaluations looked at partner capacity or commitment, and these did so in only a limited way. NGO institutional evaluations were most likely to consider this issue, although some others did at least recognise (though not analyse) the crucial importance of the commitment and capacity of government and/or partners for a successful change in gender relationships.

The **quality of the basic material** with which evaluators had to work was often poor. Over one-quarter of the general evaluations explicitly identified a lack of sex-disaggregated information as a constraint to assessing the participation of women in activities evaluated, and to assessing results. The thematic evaluations had similar problems – a lack of baseline data, limited monitoring information and an absence of sex-disaggregated data on gender relations and on benefits. Faced with these problems, evaluators collected qualitative data to assess results. The larger evaluations often had the opportunity and capacity to collect such data themselves. However, many others did not, due to inadequate time and resources, which meant they were unable to produce concrete information. Overall, few evaluations presented sound quantitative and qualitative data about progress made for women and men over the period evaluated; many were limited to general remarks, unsupported by evidence.

The **quality of the evaluations** reviewed in terms of their methods varied, although they had some common shortcomings. Despite the shift in approaches at the policy level from “women in development” to “gender equality” or “gender and development”, the evaluations focused overwhelmingly on women. This was so in the gender sections of the general evaluation reports, but also in the thematic evaluations. The word “gender” was frequently used as a synonym for “women”. Few studies consistently qualified “gender” and used it correctly in connection with relationships, equality, roles or responsibilities. In most reports men were not mentioned. This means that there was little or no focus on gender relations or on the impact of development activities on gender equality, nor was there any focus on how gender relations influenced women’s capacity to participate or benefit. This points to an apparent contradiction: while donors have adopted the gender language and have gender equality as their official policy, the implementation process seems not to be oriented towards gender equality. This may well be one of the reasons for

much confusing terminology in evaluation reports. On the positive side, for those studies that did report on gender equality, the focus was usually on changes in women's decision making, most often at the community level.

Conclusions and lessons

The lessons regarding the features of evaluation methodology that ensure adequate attention is given to gender issues are not new, and they highlight the importance of both general evaluation and gender analysis capacity.

Objectives of evaluations

The objective of “learning lessons” was stated repeatedly in all types of evaluation. This assumed that:

- there were lessons to be learned in practice (frequently, this was less so than expected)
- the evaluation team had the ability to pass on these lessons (many evaluation reports showed that they lacked this capacity), and
- there was someone listening and taking the lessons to heart (a minority of the evaluation reports made explicit who would be the users of the evaluation).

Scope and focus

Gender mainstreaming is merely a means to an end, not an end in itself. However, many evaluations focused only on evaluating institutional mainstreaming, without considering the extent to which this led to the desired changes in gender relations. In some evaluations, this may have been the result of limited resources, or a decision to focus on what the agency could control or influence in the context of attempts to increase agency accountability to gender equality policy. It is important for future thematic evaluations to focus on results, as well as on institutional mainstreaming practices.

Also problematic was designing an evaluation on the assumption that mainstreaming leads to effective gender equality outcomes. It proved to be more useful to find positive or negative results of activities first, and then to examine the factors that promoted good or poor performance.

Methods, approaches and tools

The review does not point to any one particular method, approach or tool that can be used for effectively evaluating gender equality, but the following lessons can be drawn.

- Evaluating gender equality issues or objectives requires, in the first instance, good quality evaluation design and implementation. This means that there is a clear and simple objective, and a transparent design, with findings based on evidence, clear evaluation criteria and gender-sensitive indicators. Without these basic prerequisites, it

is very difficult to collect credible and useful information on gender equality. Some of the thematic studies developed good frameworks that can be adapted to other evaluation studies.⁷

- The design of the evaluation needs to match the objectives. If learning lessons is the first objective, the selection of activities to be evaluated needs to be handled very carefully and the study should be adapted if necessary.
- Consistent quantitative and qualitative sex-disaggregated data are crucial for a good evaluation. Baseline data (disaggregated by sex, as well as by age, ethnicity and other key socio-economic variables) need to be collected during the design phase. Without a baseline, and sex-disaggregated information collected through implementation, it is difficult to measure change and potential differences in results for women and men. If data are not available at the time of the evaluation, efforts should be made to collect it in an efficient way.
- The whole range of PRA techniques, including focus group discussions, will bring out satisfactory information about gender relations only if these are used separately with women and men. The findings resulting from PRA techniques can be cross-checked with findings resulting from the use of other methods. This will limit any bias from using only one method and will solve the problem of the small numbers of people involved in PRA data collection methods.
- No matter how effective the methods, approaches or tools might be, they require human and financial resources to implement them. They require a capacity to undertake gender analysis that is appropriate to the scope and objectives of the evaluation. Specifically, evaluators need the capacity to determine appropriate research questions to investigate potential differences in participation, benefits and impacts for women and men; the capacity to formulate gender-sensitive indicators; the capacity to identify and collect sex-disaggregated data; and the capacity to analyse the data collected in relation to the activity being evaluated. They also need sufficient time and resources to do a good job. If expertise, time and money are lacking, even the best evaluation design will prove to be ineffective.

Expertise and quality

The quality of the evaluations is very mixed. There are good and sound evaluations, but there are also a lot of poor evaluations. The latter group had a high level of improvisation, characterised by the absence of evaluation criteria, gender-sensitive indicators and a framework for the evaluation. Sometimes, evaluation designs were too complex to be carried through, leading to disagreement among members of the team, descriptive instead

⁷ See World Bank (2002), ADB (2001) and Netherlands (1998) in Table B1.

of evaluative reports, and highly impressionistic findings that describe the views of the researchers, which cannot be traced back to data presented.

Apparently, there is considerable ignorance about evaluation as an instrument to assess an activity or policy, both among evaluators and agency staff. This was reinforced by interviews with a small sample of CIDA staff. Those interviews indicated that CIDA officers could benefit from courses in how to manage evaluations that include integrating a gender equality perspective at all stages in the process of evaluation.

The likelihood of adequately addressing gender issues in general evaluations is increased where attention is given to a “critical mass” of evaluation features. No one feature (such as the collection of qualitative data, the employment of local consultants, or substantive references in the TORs) is sufficient.

5 Findings on Institutional Approaches and Change Strategies

Institutional approaches and change strategies have been used to respond to recurrent obstacles to gender mainstreaming. While there is an extensive literature on the institutionalisation of gender equality approaches,¹ the value added by this current review is that it distils evidence based on agency evaluations. The word “evidence” is important here. Many of the evaluations in the database contain suggestions or recommendations for ways to increase the attention given to gender issues. This chapter does not report on those suggestions and recommendations. Unless otherwise stated, the approaches and strategies discussed are drawn from evaluators’ assessments of what has worked in practice.

The different types of evaluation tended to focus on different levels of institutionalisation. The thematic evaluations in the policy and cluster categories tended to focus more on policy dialogue, country assistance strategies and obstacles relating to donor and partner institutions, whereas those in the project category tended to focus on obstacles in the activity cycle. Comparatively few evaluations focused on the progress made towards gender equality at the community level.

There is a very high degree of consistency in the findings about the major obstacles and about the strategies that increased the attention given to gender issues at various levels. Unless otherwise stated, only those change strategies identified as successful in three or more evaluations are noted here. To avoid repetition, the obstacles and successful change strategies are discussed under the following headings:

- Partnerships and dialogue on gender equality
- Donor policy and country assistance strategies
- Donor agencies
- The activity cycle
- Progress towards gender equality at the community level.

Partnerships and dialogue on gender equality

A major obstacle identified in a number of evaluation reports is the lack of attention given to assessing and strengthening partner capacity and commitment to gender-sensitive and participatory design and implementation.² Others identified the lack of ownership of

¹ A review of the literature in this area was beyond the scope of this review.

² See AusAID (2001), ILO (2002), CIDA (2001d), CIDA (1999c), Sida (2001a), Sida (2000b) in Table B1 and Table B2.

gender equality goals and strategies at the activity level as a key problem.³ This also applied to NGOs.⁴ As the ADB pointed out, the capacity to implement gender provisions in design is often assumed.⁵

Regular and consistent dialogue between donor agencies and partners was highlighted as a key factor resulting in a shared vision and clear objectives. In the education sector, all of the factors mentioned in Box 1 were fundamental to developing successful programs to expand access to universal primary education.⁶

Box 1 **Fundamental factors for building partnerships on gender equality**

From the evaluations reviewed, a number of recurring factors emerged across agencies and different types of assistance as fundamentally important to successful change strategies. These factors were found to be important for successful *policy dialogue*,⁷ for incorporating gender equality issues into *country assistance strategies*,⁸ and for successfully addressing gender equality issues and women's needs in *activity design, implementation and monitoring*.⁹

The fundamental success factors are:

- *National and/or partner institution ownership*, which needs to be translated into a shared vision between donor and partner agencies for how gender equality is relevant to country strategies and individual activities in the prevailing socio-economic and political context.
- *Explicit consensus on clear and simple objectives* (focused on meeting women's and girls' needs, or on gender equality) that are transparently relevant to partner government policies and commitments, and that result in agreed decisions for investments and activities.
- *The involvement of stakeholders from civil society* in advocacy and in setting directions, which appears to play a key role in facilitating the incorporation of gender equality objectives into country assistance strategies. Strengthening civil society was also mentioned as a strategy for overcoming problems with partner institution commitment or with gaps between commitments and practice. The importance of national machineries for women, NGOs, women's organisations and other local gender equality advocates was highlighted by some evaluations in this regard.¹⁰
- *Long-term commitments*, which are important for making progress towards gender

³ See Danida (2002a), Sida (2002a) and AusAID (2000d) in Table B1 and Table B2.

⁴ See CIDA (2000c), CIDA (1999f) and ADB (2001, p. 20) in Table B1 and Table B2.

⁵ ADB (2001, pp. iii, 22) in Table B1.

⁶ See USAID (1999d, pp. 14-15) in Table B1.

⁷ See Sida (2002a), Sida (2002c), CIDA (2001d), CIDA (2000b), WFP (2002), CIDA (1999c), USAID (1999d), World Bank (2001a), SPC (2001) in Annex B.

⁸ See World Bank (2001a, 2002), UNIFEM (2002) and NORAD (1999b) in Table B1.

⁹ See Sida (2002a), Sida (2000b), CIDA (2002a), CIDA (2002b), CIDA (1999c), ADB (2001), AusAID (1999b), DFID (1999a), JICA (2000), AusAID (2001), NZAID (2002b), World Bank (2002) and AusAID (2000b) in Table B1 and Table B2.

¹⁰ See Sida (2002a), Sida (2001b) and NORAD (1999b) in Table B1 and Table B2.

equality at both macro and micro levels.

Donor policy and country assistance strategies

Those evaluation reports that focused on obstacles to implementing gender equality policy noted that donor agencies have crowded policy agendas, which results in a failure to prioritise gender equality objectives in country strategies. This goes hand in hand with a lack of guidance on how to “operationalise” policy. The evaluations called for stronger and more explicit links to be made between gender equality and poverty reduction policies, and for advice on how these should be addressed in country strategies and interventions.¹¹ No clearly successful strategies emerged to address these issues in the evaluations reviewed.

Donor agencies

Although the evaluations listed a number of different obstacles within donor agencies, the main obstacle about which there was consensus was their lack of accountability for gender equality commitments that are not embedded in general agency procedures, instruments and overall policy implementation monitoring systems.¹² For example, one of AusAID’s general evaluations noted that three activities were rated highly on AusAID’s primary management and reporting tool, despite gender issues being very poorly addressed in two of them.¹³

Only two evaluations noted that having agency-level targets and plans helped to address some of the issues in their lack of accountability,¹⁴ and one pointed to the need for such a plan.¹⁵ Others noted that leadership and commitment were key factors in creating an enabling corporate culture.¹⁶

Some evaluations noted that training was often ineffective at assisting staff to integrate gender equality concerns into their immediate tasks.¹⁷ A related problem was a lack of

¹¹ See Sida (2002a), WFP (2002), World Bank (2002), DFID (1999a), NORAD (1999b), AusAID (2001) and ILO (2002) in Table B1, as well as Chapter 6.

¹² See BMZ (2000, p. 3), AusAID (2001, p. vi), Sida (2002c), WFP (2002, p. 27), NORAD (1999b, pp. 2-4), CIDA (2002a) and World Bank (2001a, pp. 24, 31) in Table B1.

¹³ AusAID (1999c, p. 42) in Table B2.

¹⁴ WFP (2002) and ILO (2002) in Table B1.

¹⁵ AusAID (2001) in Table B1.

¹⁶ See NORAD (1999b), UNIFEM (2002) and Sida (2002a) in Table B1.

¹⁷ See AusAID (2001), WFP (2002), NORAD (1999b) and World Bank (2001a) in Table B1.

conceptual clarity – for example, treating gender mainstreaming as a goal rather than as a strategy, or confusing women-focused approaches and gender equality approaches.¹⁸ The likelihood of training being applied to tasks increased when it was sector-specific, hands-on and/or directly linked to individual activities and contexts.¹⁹

Four evaluations mentioned the lack of gender analysis tools as a constraint, but each had a different perspective on the matter. One Sida evaluation mentioned that there were no tools and little expertise for addressing gender issues in SWApS; the ILO noted that there was a lack of effective or accessible tools (partly due to poor institutional memory); and AusAID's policy evaluation identified a need for sector-specific tools.²⁰ Sida's policy evaluation identified the problem as a lack of use of existing tools.²¹ No evaluations identified tools as having assisted with addressing gender issues, unless the development of the tool had been used as a training or learning exercise.²²

The activity cycle

Design and implementation – obstacles

There is consensus in the evaluation reports that the lack of attention given to gender issues in activity designs is a critical constraint to gender-sensitive implementation. A lack of participation by women in activity design, poor needs analysis, the lack of baseline data on key gender differences, and a failure to address gender issues at the level of activity objectives were all highlighted.²³

A more specific obstacle identified by some evaluations arose when gender analysis was undertaken in design but not linked in any meaningful way to the activity to be implemented or to the expected results.²⁴ This made it exceedingly difficult to seriously address gender issues during implementation, and increased the likelihood that the

¹⁸ See AusAID (2001), ILO (2002), WFP (2002) and UNIFEM (2002) in Table B1.

¹⁹ See Sida (2002a), ILO (2002), EC (2001b), AusAID (1999b) and DFID (1999a) in Table B1 and Table B2.

²⁰ See Sida (2002c), ILO (2002) and AusAID (2001, p. vii) in Table B1.

²¹ Sida (2002a) in Table B1. Anecdotal evidence also suggested that existing tools were rarely used (Hunt 2000, p. 13).

²² See NORAD (1999b) in Table B1.

²³ See AusAID (2000g), EC (2001b), DFID (2002a), World Bank (2001a), JICA (2000), Danida (2001b), AusAID (1999b), AusAID (2000f), EC (2001a), AusAID (2000g), ADB (2001) and WFP (2002) in Table B1 and Table B2.

²⁴ See CIDA (2001d), CIDA (2000d), Danida (2001b), AusAID (2001), Sida (2002a), CIDA (2001a), ADB (2001), CIDA (2002b) and ILO (2002) in Table B1 and Table B2.

attention given to women would be marginalised to individual components.²⁵ Other evaluations identified the lack of a coherent approach to addressing gender issues in an activity, particularly the lack of an explicit gender strategy, as a key obstacle.²⁶

Some evaluations noted the lack of expertise in gender and social analysis as a constraint during implementation.²⁷ The lack of integration between gender analysis and the overall social analysis undertaken for design was also identified as a constraint.²⁸ Findings on evaluation methodology (see chapter 4) suggest that this problem is extensive. Without integrated analysis, agency staff are far less likely to be able to make concrete links between poverty reduction and gender equality objectives, and to strategically apply the fundamental success factors identified in Box 1. A DFID evaluation of gender mainstreaming in three countries highlights this issue. The evaluator pointed out that DFID devotes far greater resources to gender mainstreaming in the Pakistan program, where the scale of the problem is great, the opportunities for working in partnership limited and the likelihood of sustainable short-term change minimal, than in the southern African program, where the opportunities for working in partnership are great and the consequent possibilities for sustainable change considerable.²⁹

Design and implementation – change strategies

Not surprisingly, there is considerable consensus that addressing gender issues in design is an important strategy for ensuring that they are properly addressed in implementation. While many evaluations were vague about the details, others specified key success factors: involving both women and men in design and implementation, ensuring that there is adequate time for field-based studies, having adequate baseline data, and ensuring that gender issues are incorporated into activity objectives.³⁰

Other evaluations noted that an explicit gender strategy for an activity is critical, but most qualified this by saying that the strategy has to be of high quality, pragmatic and, of

²⁵ See BMZ (2000) in Table B1.

²⁶ See AusAID (1999a), Danida (2001b), Danida (2000c), CIDA (2002b) and JICA (2000) in Table B1 and Table B2.

²⁷ See DFID (1999a), Sida (2002c) and AusAID (2001) in Table B1.

²⁸ See World Bank (2001a), Sida (2002a), WFP (2002) and EC (2001b) in Table B1 and Table B2.

²⁹ DFID (1999a, para. 2.3.11).

³⁰ See ADB (2001), AusAID (2001), AusAID (2000b), AusAID (2000g), AusAID (1999a), AusAID (1999b), BMZ (2000), CIDA (2001a), DFID (1999a), EC (2001b), JICA (2000), JICA (2001) and Sida (2002a) in Table B1 and Table B2.

course, implemented.³¹ While Sida's policy evaluation did not explicitly identify a gender strategy as a success factor in addressing gender issues, it did note the importance of key elements that should be included in any strategy. These included exploring what gender equality means for each situation and development activity, as a step towards identifying expected goals, results and indicators.³²

There is an emerging consensus that the likelihood of gender issues being addressed increases if explicit responsibilities for achieving gender equality objectives are included in TORs, the scope of services and job descriptions. This applies for all personnel engaged in the design, implementation and monitoring of an activity.³³ The importance of having gender-aware, competent and committed staff and stakeholders in the field to implement and manage activities was noted by many evaluators as a key to successfully addressing gender issues.³⁴ One DFID evaluation highlighted the benefits of having designated staff with time, resources and support to fulfil their responsibilities.³⁵ Others identified having gender expertise in the partner country as a key success factor where gender issues had been addressed.³⁶

Monitoring

As expected from the findings in chapter 4, a number of evaluations identified weak monitoring processes as a serious constraint. Specific issues highlighted included:

- the lack of attention given to or the capacity for monitoring impacts³⁷
- the lack of baseline data, which prevented the assessment of impacts³⁸
- the lack of gender-sensitive indicators and sex-disaggregated information,³⁹ and

³¹ See CIDA (2002a), CIDA (2002b), CIDA (2001a), AusAID (2000d) and AusAID (1999a) in Table B1 and Table B2.

³² Sida (2002a, pp. xv, 98) in Table B1.

³³ See ADB (2001), AusAID (2001), CIDA (2001a), CIDA (2002b), EC (2001b), JICA (2001) and Sida (2002a) in Table B1 and Table B2.

³⁴ See ILO (2002), JICA (2000), CIDA (2002b), DFID (1999a), WFP (2002), AusAID (2000d), EC (2001b) and Danida (2002a) in Table B1 and Table B2.

³⁵ DFID (1999a) in Table B1.

³⁶ See Netherlands (1998), ADB (2001), JICA (2000, pp. 285), AusAID (2000c), AusAID (1999c) and Sida (2002a) in Table B1 and Table B2.

³⁷ See AusAID (1999c), BMZ (2000), Sida (2000b), NZAID (2002b), EC (2001b) and NORAD (2000) in Table B1 and Table B2.

³⁸ See AusAID (2000c), JICA (2000), JICA (2001) and WFP (2002).

³⁹ See AusAID (1999b), AusAID (2000a), AusAID (2000b), BMZ (2000), CIDA (2002b), ILO (2002), World Bank (2001a), AusAID (2000f), AusAID (1999a), CIDA (1999c) and CIDA (2001c) in Table B1 and Table B2.

- the lack of qualitative gender equality indicators.⁴⁰

⁴⁰ See ADB (2001), CIDA (2000d), EC (2001a), Sida (2002a) and WFP (2002).

AusAID's policy review also noted the lack of resources for collecting data where gender equality objectives were not included in logframes and in monitoring and evaluation plans.⁴¹ The strategy used successfully to address this obstacle was of course to develop and use gender-sensitive indicators. Some evaluations noted that this increased the attention given to gender issues during implementation.⁴²

Progress towards gender equality at the community level

Obstacles

The lack of effective targeting of women to ensure that they participate in key training and other project activities was identified as a problem by three AusAID evaluations,⁴³ whereas three other evaluations focused on the lack of participatory processes in project implementation as a constraint.⁴⁴ Mandatory requirements or even just the expectation that women should participate sometimes increased the overall burden on women without yielding any economic returns, and led to women dropping out of project activities. The assumption that women will provide unpaid labour to participate was also identified as a problem.⁴⁵

Other evaluations highlighted the difficulty of women moving from participating in an activity to being involved in decision making and to controlling resources.⁴⁶ While some evaluators noted that "cultural constraints" inhibited women from participating and benefiting,⁴⁷ few reports identified the specific constraints, which is an essential first step to addressing them.⁴⁸ The need for dialogue and the involvement of men to achieve benefits for women or changes in gender equality was rarely identified as an issue.⁴⁹

⁴¹ AusAID (2001, pp. 15-16) in Table B1. Only one evaluation (ILO 2002) indicated that the general understanding of indicators was poor, as distinct from gender-sensitive indicators.

⁴² See ADB (2001), ILO (2002), CIDA (2002b) and JICA (2000) in Table B1.

⁴³ See AusAID (2000a, 2000d, 2000f) in Table B2.

⁴⁴ See AusAID (1999a), Sida (2002a) and World Bank (2001a) in Table B1 and Table B2.

⁴⁵ See World Bank (2002), CIDA (2002b) and ADB (2001) in Table B1.

⁴⁶ See AusAID (1999a), CIDA (2000c), WFP (2002) and World Bank (2002) in Table B1 and Table B2.

⁴⁷ See AusAID (2001), ADB (2001) and AusAID (1999b) in Table B1 and Table B2.

⁴⁸ See, for example, USAID (2001b) in Table B1.

⁴⁹ See Sida (2002a), CIDA (1999c) and CIDA (2000g) in Table B1 and Table B2.

Successful change strategies⁵⁰

Participatory approaches were highlighted as successful for bringing about changes in gender relations at the community level. Providing leadership training, forming women's groups, involving women in decision-making bodies, and ensuring that women had access to project resources were all mentioned as having been successful, along with networking and support for women's advocacy with local men.⁵¹

A number of evaluations noted that the use of female staff and extension agents helped to effectively target women and engage them in activities.⁵² The importance of supporting women's organisations as agents of civil society featured strongly in the success of some programs.⁵³

Conclusions and lessons

There is a very high degree of consensus about the institutional obstacles to addressing gender issues and about what change strategies have been successful in the field. The change strategies mentioned are neither new nor a comprehensive list of those tried. They reflect the experience of program planners, implementers and gender specialists about actions that can improve the implementation of gender equality policy in development activities.

The following lessons can be drawn from the findings on institutional approaches to gender mainstreaming and successful change strategies.

- More attention needs to be given to the capacity of development partners to implement gender-sensitive approaches and to commit to gender equality objectives across all levels of institutionalisation.
- Building partnerships on gender equality through dialogue needs to be seen as a fundamental aspect of any change strategy. There is a high degree of consensus about the importance of four factors to support this process: national or partner ownership, shared clear objectives, stakeholder involvement, and long-term commitment.

⁵⁰ See Annex F for examples of how women have benefited from the activities evaluated. Chapter 6 examines evidence of a causal relationship between the attention given to gender issues and the benefits identified.

⁵¹ See AusAID (1999b), AusAID (1999c), AusAID (2000c), CIDA (2000b), CIDA (2000e), CIDA (2000g), EC (2001b), JICA (2000), Sida (2000a) and WFP (2002, p. 9) in Table B1 and Table B2.

⁵² See AusAID (1999b), CIDA (1999e), Danida (2000c), AusAID (2000b), AusAID (2000c) and JICA (2001) in Table B1 and Table B2.

⁵³ See USAID (2001b), USAID (1999c), Sida (2001a) and CIDA (2000a) in Table B1.

- There is an urgent need to make explicit links between gender equality and poverty reduction objectives at the policy level, in country assistance strategies and in individual activities.
- Within donor agencies there needs to be greater accountability to gender equality policy and for mainstreaming gender equality perspectives in general procedures. Although there are no simple ways to increase this accountability, leadership and commitment are essential, and overall agency plans have been useful for some agencies.
- Including responsibilities for implementing gender equality policy in the job descriptions and the TORs of key personnel can help to ensure that attention is given to gender issues in activity design, implementation and monitoring. While few evaluations were specific about how these responsibilities should be addressed, it can be assumed that vague references to policy will be less effective than specific requirements relating directly to the activity and the tasks of personnel.
- The use of gender-sensitive indicators needs to be seen as a minimum standard requirement in activity design, implementation and monitoring. This will increase the attention given to gender issues and provide more information on activity results (see also Chapters 4 and 6).
- Using participatory approaches to activity design and implementation, more effectively targeting women, and strengthening their leadership capacity can hasten progress towards gender equality at the community level. (Chapter 6 points to the need to underpin these strategies with sound analysis of gender relations and other factors in the socio-economic and political context.) Supporting women's organisations as agents of civil society is important at both the macro and micro levels for making progress towards sustainable changes in gender relations.

6 Findings on the Benefits for Women, Gender Analysis and Development Outcomes

Determining whether gender analysis or gender mainstreaming had affected overall development outcomes posed challenges because most evaluations included in this review were not tasked to investigate links between gender analysis and benefits and overall development outcomes.

This type of attribution is not self-evident and requires a level of sophistication in gender analysis that most of these evaluations did not have. (It also assumes a high level of attribution from development activities in general.) Furthermore, most evaluations in the review's database reported on activities that had poor gender analysis, and the evaluators were faced with a scarcity of data on participation in activities, let alone results. Systematic approaches to gender analysis and mainstreaming in the activities evaluated appear to have been rare. (Annex F provides examples that illustrate the extent of this problem.) Where there was some evidence of attention given to gender issues, it was usually difficult to determine whether this occurred in design or implementation.

In retrospect, it is clear that the database for this review was inadequate to answer the research question posed for its third area of focus,¹ although it does suggest areas for future research and evaluation. Because of the difficulties encountered, the research question was changed during the review to capture:

- evidence of benefits for women and changes in gender relations
- evidence of a causal relationship between the attention given to gender issues in activity design or implementation and the benefits for women, and
- evidence that benefits for women or changes in gender relations led to improved development outcomes.

Examples from specific evaluations are presented in Annex F.

Evidence of benefits for women and changes in gender relations

Around 60% of the total sample of thematic and general evaluations reported on development benefits or impacts, but many of the general evaluations gave only minimal information (see Tables 3, 4 and 5 in Chapter 4). One-quarter of the evaluations in the database is of women-specific activities or gender equality funds. Few evaluations

¹ The review's TORs (see Annex A) foresaw some of these difficulties and noted that the size of the database for this area of focus was unknown.

specifically investigated differences in the benefits for men and women, or reported on changes in gender relations.

Overall, most of the benefits identified addressed women's practical needs. However, this varied according to the type of evaluation. Thematic policy and general evaluations generally identified practical benefits for women. Evaluations of activities that aimed to promote gender equality and targeted women tended to identify strategic changes in gender relations.

Usually, in the evaluations that identified practical benefits for women, the evaluators also noted that these benefits were limited to a few areas of the activity. Some also mentioned that there were no strategic changes in gender relations, or identified negative results alongside the benefits.²

Qualifications about benefits typically occurred in evaluations that employed more systematic gender analysis. Some evaluations illustrated how meeting practical needs might also have contributed to strategic changes in gender relations and the empowerment of women.³ Sida's policy evaluation highlighted the strengthened identity and organisational capacity of women's groups in relation to men as key factors;⁴ others focused on the empowering effects of training.

Several reports addressed the benefits in girls' education at the macro level. Evidence of increased enrolments varied considerably from one country and evaluation to another, and serious questions about the sustainability of increased girls' enrolments were raised due to the impact of increased enrolments of boys and girls on the quality of education.⁵

In the database for the review there were 21 evaluations of activities that specifically targeted women or gender equality. Half of these identified strategic changes in gender relations, but many of these activities were on a small scale. Five were evaluations of gender equality funds, which support local government, national machineries for women or NGO activities. These evaluations are an interesting subset of the database because they covered activities aimed at creating an enabling social and political environment for gender equality. The participation of women in political affairs, particularly at the local level, was an outcome of half the gender equality funds.⁶ Other activities resulted in

² For example, BMZ (2000), Danida (2000c) and AusAID (2000f) in Table B1 and Table B2.

³ See Belgium (2001c), JICA (2001, p. 54) and AusAID (2000g, p. 35) in Table B1 and Table B2.

⁴ Sida (2002a, p. xiii) in Table B1.

⁵ See World Bank (2002), USAID (1999d) and CIDA (1999a) in Table B1.

⁶ See CIDA (1999c), CIDA (2000b) and USAID (1999c) in Table B1.

increased public discussion on gender issues and violence against women.⁷ USAID's Women's Empowerment Program in Nepal is an example of an activity that successfully focused on individual empowerment at the household level.⁸ Evidence of women's increased participation in decision making at various levels is a commonly cited result of most of these activities.

Evidence of a causal relationship between the attention given to gender issues and benefits for women

In those evaluations that identified benefits for women, many provided no explanation or analysis of how or why these benefits came about (see Chapter 4, page 38, for discussion on the quality of gender analysis in general evaluations). The evaluations that did discuss how the benefits arose can be divided into three broad groups.

The first group of evaluations concluded that women benefited because attention was given to gender issues in activity design or implementation, even where this was limited. That attention was in the form of having gender or social expertise available in design or implementation, using participatory design or implementation strategies that involved women, providing training for women, or ensuring that women had access to other resources provided by the program or project (see Annex F).⁹

A number of evaluations cited the targeting of women as a key strategy for ensuring that women benefited. The strongest evidence came from cluster evaluations where one or two activities showed benefits for women because gender issues were addressed in some way, whereas other activities that gave little or no attention to women showed no benefits.¹⁰

Evidence of a causal relationship between the attention given to gender issues and women's participation was found in the Netherlands policy review, but women were still far behind men in terms of equality and benefited mainly in economic activities. Sida's policy review found that the two activities that had explicit gender equality objectives had the most evidence of positive gender equality impacts.¹¹

⁷ See Sida (2001a), NZAID (2002a) and CIDA (2000a) in Table B1.

⁸ USAID (2001a) in Table B1.

⁹ See AusAID (2000d), Danida (2000c), CIDA (2000g), CIDA (2000e), CIDA (1999e), NORAD (1999c) and WFP (2002) in Table B1 and Table B2.

¹⁰ See JICA (2000), JICA (2001), AusAID (1999c) and AusAID (1999b) in Table B1 and Table B2.

¹¹ See Netherlands (1998) and Sida (2002a) in Table B1. Specific features of activity design and implementation that assist with institutionalising gender equality objectives are discussed in Chapter 5.

Only one evaluation in the database – an NGO-managed microfinance and co-operative development project in Guatemala – was designed to assess whether the introduction of a gender strategy had increased results for women. It showed that women were more likely to access project resources (training and credit) and to participate as members and decision makers in co-operatives as a result of the gender strategy. Sensitisation of men was identified as essential for ensuring increased participation by women.¹²

The second group of evaluations concluded that women did not benefit equally, or suffered adverse impacts, because gender analysis was not undertaken and little or no attention was given to their needs during activity design or implementation.¹³ In some cases, benefits were seen as coincidental or unintended because no provision was made to ensure that women participated or that their needs were met.¹⁴

The third group of evaluations (mainly policy, sectoral or program evaluations at the macro level) concluded that the local cultural, institutional and policy context of activities was a far more important determinant of whether women benefited, and whether the benefits would be sustained, than any specific actions on addressing gender issues during activity design or implementation.¹⁵ Sida's policy evaluation also highlighted the importance of public discussion of gender equality to influence factors such as political support and cultural resistance. DFID's review of support for health sector reform went furthest in making this point:

Inclusion of participation and gender equity in project design has had little impact on their successful adoption in public sector health systems, **except where prevailing policies and practices of partner governments have been favourable.**¹⁶

The World Bank's evaluation presents examples that fall into each of the three groups of evaluations. It cited many instances where gender-blind activity design resulted in an adverse impact on women's economic activities. It seems clear that at the activity level, results for women tend to be better if gender considerations are integrated into activity design and implementation. However, there is also evidence of the critical importance of institutional and other contextual factors in education and economic assistance. There were some instances where the World Bank's assistance may not have given attention to gender issues, but partner institutions did, and that is the critical point:

¹² See CIDA (2002b, p. 5) and CIDA (2002a) in Table B1.

¹³ See BMZ (2000), AusAID (2000g), AusAID (2000f), AusAID (2000d), AusAID (1999a) in Table B1 and Table B2.

¹⁴ See CIDA (2001c) and Danida (2001a) in Table B1 and Table B2.

¹⁵ See DFID (1999a), BMZ (2001a), Danida (2000c), AusAID (2000d), EC (2001b) and Sida (2002a) in Table B1 and Table B2.

¹⁶ DFID (2001b, p. 4) in Table B2; emphasis added.

In Poland and Vietnam, both men and women were able to benefit equitably from largely gender blind Bank assistance because implementing agencies stepped in with targeting mechanisms for women, or because both women and men were able to access benefits.¹⁷

Findings from activities specifically aimed at promoting gender equality and changes in gender relations reinforce the importance of the local social and institutional context, and partner capacity. Where there was some sound evidence that gender relations were changing, key factors appear to have been:

- the relevance and consistency of activities with partner government priorities and policies for the advancement of women
- the ownership of activities by local stakeholders who participate in decisions on the strategic change areas to be targeted, and
- the capacity within local implementing organisations to ensure the sustainability of outcomes.¹⁸

However, it was still rare for evaluations to acknowledge that men also have a role to play in changing gender relations.

Evidence that benefits for women or changes in gender relations lead to improved development outcomes¹⁹

The Netherlands policy evaluation found that, although women were participating more in 1995 than in 1985, and benefiting from small-scale economic activities, there was:

... no real change in the dynamics of rural areas ... The small scale of the interventions in general and the predominantly small-scale nature of the activities of women do not appear to stimulate economic growth in a broader sense.²⁰

The World Bank policy evaluation explored the links between investments in the human capital of women (health and education) and poverty reduction for women:

Improving women's health and increasing enrolments of girls—although essential to sustaining growth—are not adequate in and of themselves to result in poverty reduction for women. The Bank focused effectively on increasing girls' enrolments, but failed to focus attention on issues related to employment for girls entering the labour force after secondary

¹⁷ World Bank (2002, pp. 14-16) in Table B1.

¹⁸ See Sida (2001a, 2000b), CIDA (1999c, 2000a, 2000b) and USAID (1999c) in Table B1.

¹⁹ This section summarises findings from thematic policy evaluations, from activities that demonstrate changes in gender relations, and findings in different sectors in all types of evaluations.

²⁰ Netherlands (1998, p. 243) in Table B1.

education ... In addition to investing in women's human capital, supporting the increased participation of women in economic activities is a critical dimension for poverty reduction and the enhancement of overall welfare of women.²¹

The issue of sustainability is critical when considering whether benefits for women lead to improved development outcomes. BMZ's policy evaluation noted that the sustainability of interventions was questionable because gender issues were not addressed in health and water activities, in HIV/AIDS and family planning activities (where it is essential to have men's involvement and acceptance of new measures) and in an agricultural activity that required women to do extra work (which they would do only if they received additional benefits).²²

Both Sida's and the ADB's policy evaluations concluded that more research is needed on the links between gender equality and development outcomes, particularly poverty reduction.²³ The ADB evaluation stressed that research findings need to be published in national languages. One difficulty encountered by the Sida evaluators was that links between gender equality and poverty reduction goals were rarely explicit in activity design. However, the evaluation team found these linkages "in the voices of the poor" – for example, women who saw equality as a necessary condition to escape from poverty. Other evaluations also highlighted the need to clarify links between poverty reduction and gender equality mandates (see also chapter 5).²⁴

The evaluations of activities aimed at promoting gender equality by strengthening local women's organisations to set their own mainstreaming agenda showed the strongest evidence of strategic changes in gender relations. In these cases the implementers assumed that it is a woman's right to participate and benefit from development. Generally, these evaluations did not seek to link gender equality efforts with the instrumentalist view that such efforts ought to improve economic growth, and few explored links with poverty reduction. However, some explored how these activities improved governance processes, and a few general evaluations also made similar links.

The strongest evidence came from four evaluations that explored the impacts of targeting women in post-conflict situations. For example, USAID assistance to women's associations in Rwanda was effective at directly addressing poverty (because women's associations were successful at targeting the most vulnerable, including households headed by females) and had a direct effect on reducing social tensions and promoting unity. WFP

²¹ World Bank (2002, pp. 13-14) in Table B1.

²² BMZ (2000) in Table B1.

²³ Sida (2002a, pp. ix-x, 41) and ADB (2001, p. iv) in Table B1.

²⁴ See World Bank (2002, p. 29), AusAID (2001, pp. vi-vii) and WFP (2002) in Table B1.

and CIDA evaluations and a broader study of post-conflict situations by USAID found similar impacts from working through women's groups.²⁵

Sida's evaluation of the support given to women through Panchayati Raj in India found that women's involvement in local government resulted in an extension of local government programs to the most needy.²⁶ Other evaluations reported that women's involvement in local governance structures or in project activities, groups and committees increased the effectiveness of the projects.²⁷ However, it should be noted that most of these interventions were of a small scale and half were implemented through NGOs or small grant schemes.

In general, the evaluations that included water supply activities did not find much evidence of women's involvement in their implementation. However, where women were involved, the World Bank concluded that their involvement improved the use and sustainability of the physical facilities.²⁸

Control over income, spending and benefits was not thoroughly investigated in most evaluations of income-generation activities. However, those that did investigate these matters found that women were more likely to send children to school²⁹ and to spend their income on basic family needs.³⁰ These outcomes have both direct and indirect impacts on poverty reduction at the micro level.

Conclusions and lessons

If agencies want to learn about how to increase benefits for women and to promote gender equality according to their policy mandates, the quality of gender analysis in evaluations will need to improve. In general, gender analysis has been inadequate for agencies to assess differences in the benefits for women and men, and to learn lessons about what facilitates equal and sustainable benefits and impacts.

²⁵ USAID (1999c, pp. 3, 5), WFP (2002, para. 151), CIDA (2000e) and USAID (2001b) in Table B1 and Table B2.

²⁶ Sida (2000a, p. 18) in Table B1.

²⁷ See AusAID (2000g, p. 35), CIDA (2000g), AusAID (1999b), AusAID (2000d) and JICA (2000) in Table B1 and Table B2.

²⁸ World Bank (2002, p. 16) in Table B1.

²⁹ See Belgium (2001b), USAID (1999d), CIDA (1999d) and USAID (2001a) in Table B1.

³⁰ See USAID (2001a), JICA (2000, p. 283) and AusAID (2000d, pp. ix, 43) in Table B1 and Table B2.

The following lessons can be drawn from evidence of the benefits for women, gender analysis and development outcomes.

- Further research is needed to identify the factors and types of intervention that are more likely to promote strategic changes in gender relations in different social, political and institutional environments. From the evaluations it appears that the practical needs of women were more often addressed. This sometimes led to strategic changes in gender relations, particularly greater involvement of women in decision making. In some cases these changes were not planned. In other cases, women received some benefits from gender-blind activities, but only where institutional and social factors were favourable, and where partner institutions took action to ensure that women benefited.
- Gender issues need to be given systematic attention in activity design, implementation and monitoring, as it is still rare except in activities that are directly aimed at promoting gender equality.
- Giving greater attention to gender issues in activity design and implementation is essential if agencies want to increase the likelihood that both men and women will participate and benefit, and to ensure that neither group is disadvantaged. However, this is a necessary but insufficient condition to ensure that benefits are achieved and sustainable.
- Partner capacity and commitment to promote gender equality need to be assessed, as they are critical determinants of whether and how women can participate and benefit, and the likely sustainability of benefits. Very few evaluations assessed partner capacity or commitment. These factors should be analysed to provide guidance on the different types of intervention on gender equality that might be successful in different country and partner contexts, and the gender strategies that would be appropriate within those contexts.
- Given that very few agencies have successfully used gender analysis in activity design and implementation, it is essential that gender analysis be integrated into other forms of social analysis, particularly poverty and sectoral analyses.
- Lessons from the evaluations of activities specifically aimed at promoting gender equality have broad applicability to all activities that aim to benefit women. Policy dialogue with partners to agree on shared goals and strategies is essential to achieve relevance and ownership of gender strategies, and to achieve sustainable changes in gender relations. This needs to be followed up by activities that strengthen in-country institutions to undertake gender-sensitive development.
- There is some evidence to support the proposition that benefits for women improve the effectiveness of development activities and assist to achieve outcomes at the community level. However, there is no evidence for this in the database at the macro level. This

may be because so few agencies investigated these links, or because only limited attention was given to gender issues at any level. More evaluations that make links between micro, meso and macro levels are needed if agencies want to explore issues relating to development effectiveness and the achievement of overall development outcomes.

The links between benefits for women, gender equality and poverty reduction are an area for future joint agency and partner evaluation at a country and/or sectoral level. To learn lessons about successful strategies for sustainable and equitable development, the findings of this review point to the need to look at the work of multiple donors in one country rather than the work of one donor in dissimilar countries.

Identifying the predominant changes in gender relations in individual countries or across several countries over a period of time, and what women and men saw as having caused those changes, is another area for future research. The impact that development activities have on gender relations could be examined, in addition to the types of activity needed to benefit women and promote gender equality. Again, a joint agency and partner approach is recommended. It is acknowledged that attribution here is a highly complex issue. Nevertheless, it is important to investigate the role that development co-operation can play in changing gender relations, in the context of other social, political and economic changes. A multiyear study may yield the most useful findings, although this would require more resources and a longer term commitment from participating agencies. A pilot study involving a number of donors and partners should be considered.

It is essential to employ participatory methods for both areas of research. This implies longer evaluation and research timeframes (more than the standard two weeks in the partner country, which was the case for many evaluations), and adequate resourcing and training of local evaluators and researchers (see Chapter 4). The design of evaluations should take into account the need to build evaluation capacity in partner countries.

7 Progress since Previous Reviews and the Need for Follow-Up Action

Given the limited comparability of the current review with previous DAC reviews (see Chapters 2 and 3), this chapter discusses whether the attention given to gender issues in general evaluations has increased since the DAC reviews in 1999 and 1994. It also highlights the importance of ensuring that the findings of this review are disseminated to development practitioners.

Comparing findings with previous reviews¹

Taking into account the different criteria used for selecting evaluations to be reviewed, it appears that progress in addressing gender issues in general evaluations has been rather slow and uneven. This was also the overall conclusion of the 1999 review, which was based on agencies' own assessments of progress (Woodford-Berger 1999, para. 8).

The **attention given to gender issues in TORs** might have increased in terms of the number of evaluation TORs that included some mention of the need to address gender issues. However, given the different methods of selecting evaluations in the three reviews, it is impossible to assert strongly that this was the case. It is not surprising that TORs for the evaluations in the current review required evaluators to address gender issues, because agencies were specifically requested to provide evaluation reports that included some gender analysis.

Previous reviews (and DAC guidance) highlighted the need to identify specific questions and issues in TORs so as to facilitate the collection and analysis of gender-sensitive data relevant to the activity being evaluated (OECD 1994; Woodford-Berger 1999, para. 10-11; OECD 1998, p. 37). This was still rare in the general evaluations included in the current review. In most cases, references to gender issues in TORs were vague, with little detail about the quantity and quality of data and analysis expected. (The tables in Annex E provide examples from one agency, which is broadly representative of the total sample.)

The 1999 review found little improvement in the **use of sex-disaggregated data** since the 1994 review and highlighted the need for gender-sensitive indicators and evaluation criteria. The **quality of analysis of gender issues** (including the depth and relevance of discussion on gender issues in evaluation reports) was also identified as a problem that had not improved since 1994 (Woodford-Berger 1999, para. 10-11; Keller 2000).

¹ Caution is needed in comparing findings due to the different sampling methods used by the three DAC reviews (OECD 1994; Woodford-Berger 1999). See pages 24 and 32 for discussion of these differences.

In the current review, some general evaluations presented quality gender analysis supported by adequate sex-disaggregated information and the use of gender-sensitive indicators; other evaluations had no apparent indicators, little information, and even less analysis. The variation was considerable both between and within agencies, as it was at the time of the 1999 review (Woodford-Berger 1999, para. 8). To assess whether there had been any improvement in this area, it is important to keep in mind that the general evaluations in the current review can be reasonably assumed to be the best that agencies could provide in terms of their coverage of gender analysis. Furthermore, all of the evaluations that included TORs had required evaluators to address gender issues in some way. It is thus reasonable to conclude that there had been little overall improvement since the 1999 review in the use of sex-disaggregated data and the quality of gender analysis in general evaluation reports.

Another issue raised in the 1999 review (Woodford-Berger 1999, para. 11, 14; Keller 2000, pp. 6-7) related to the incorporation of gender perspectives into **standard evaluation criteria** and the need for guidance on how this could be done. While many general evaluations in the current review did use standard evaluation criteria as an overall framework for data collection and analysis, gender equality issues were not systematically integrated into this analysis. It was common for the consideration of gender issues to be marginalised to a separate section of the evaluation report, with no explicit links made to evaluation criteria. There appears to have been little overall improvement in the application of gender perspectives to evaluation criteria in general evaluations.

In the 1994 and 1999 reviews, the evaluations that did report on gender issues tended to **focus on activities and outputs, rather than on effects, outcomes or impacts** (Woodford-Berger 1999, para. 9). This was also evident for the majority of general evaluation reports in the current review, and it is difficult to assess whether there was any improvement in this area. Indeed, this tendency was a key factor in the difficulties faced in the current review in reporting on links between gender analysis and mainstreaming, and overall development outcomes (see “Benefits for women, gender analysis and development outcomes” in Chapter 3 and the introduction to Chapter 6).

The 1999 review reported some increase in the attention given to the **competence and composition of evaluation teams** (Woodford-Berger 1999, para. 11). The current review collected data on the inclusion of local expertise and expertise in gender and social analysis in evaluation teams. However, because many general evaluation reports did not provide details of the composition of evaluation teams, it is not possible to compare findings with those of previous reviews. Nevertheless, the findings do point very clearly to areas where evaluation capacity needs to be strengthened, regardless of the composition of teams (see “Conclusions and lessons” in Chapter 4).

Disseminating findings

It is important for the findings of this review to be disseminated to development practitioners. The audience for these findings is broad. It includes both agency staff and consultants/contractors who play key roles in designing, implementing, monitoring and evaluating development activities.

Evaluators, gender specialists, sectoral specialists and development managers at various levels all need to learn the lessons outlined in this review, if evaluation practice is to be improved and successful change strategies implemented so that both women and men benefit from development activities. Evaluators are frequently consultants who may have little access to DAC workshops or courses focused on evaluation capacity building (see Recommendation 4 in “Summary and Recommendations”). While senior development managers will be targeted by a note on institutional approaches to achieving gender equality (see Recommendation 5),² development practitioners directly involved in designing, implementing and monitoring activities are more likely to access findings on successful change strategies through other means.

Consideration should also be given to using a mix of dissemination strategies appropriate to different audiences, such as short summaries of key findings. For example, key findings on evaluation methodology could be included on the DAC website along with links to existing gender-sensitive evaluation tools (see Recommendation 3). A summary of findings on successful strategies for institutionalisation, and on the links between the benefits for women and effective development at community level, could be included on the tipsheets website of the DAC Network on Gender Equality, as part of the process of developing a communiqué (see Recommendation 5).³

Conclusion

It is important to put the findings of this review into context. Changing gender relations is a complex undertaking, with a variety of economic, political, social and cultural obstacles. As a result, institutional change is incremental, and the role that development co-operation can play in changing gender relations needs careful thought, planning, collaboration with partners, and future investigation through well-designed research and evaluation.

² A tipsheet, “Effective strategies for promoting gender equality” (see <http://www.oecd.org/dac/gender>), has been published by the OECD DAC Network on Gender Equality since this report was written.

³ A tipsheet, “Effective strategies for promoting gender equality” (see <http://www.oecd.org/dac/gender>), has been published by the OECD DAC Network on Gender Equality since this report was written.

However, it is also important to note that many of the findings of this review are not new. This applies to both the weaknesses in evaluation methodology identified in Chapter 4, and the strategies needed to institutionalise gender equality policy, which are highlighted in Chapters 5 and 6. The findings provide a strong basis for recommending actions that need to be taken to increase the attention given to gender issues through the activity cycle. They reinforce the need for agencies to take follow-up action to this review, both collaboratively and individually, and to ensure that the findings are disseminated to development practitioners.

Annex A Terms of Reference for Review on Gender and Evaluation*

Background

1. A Discussion Paper on “Options for Future Work on Gender Equality by the DAC Working Party on Aid Evaluation”, presented to the 35th meeting of the Working Party, outlined 3 options for future work on using evaluation tools and processes more effectively to enhance donor activities in relation to gender equality and women’s empowerment. Australia agreed to develop the options further in consultation with interested Members of WP-EV, WP-GEN and the Secretariat. The 3 options build on previous work of the DAC dating from 1991, which is outlined in the Discussion Paper, along with the rationale, objective and outputs expected from each option.
2. The 3 options outlined in the Discussion Paper are:
 - i) Reviewing institutional approaches among donors for integrating gender equality and women’s empowerment in development co-operation activities and management systems
 - ii) Improving the gender dimension in evaluation methodologies and processes
 - iii) Assessing whether engendering development activities improves the achievement of overall results
3. The 2nd option received the most support from Members, and was considered the one most closely linked to the mandate of WP-EV. It was suggested that options 1 and 2 could be combined. Options 1 and 3 received equivalent support.¹

STEP 1: Integrating the 3 options in a review of selected recent evaluations

4. The Discussion Paper proposed that work on options 2 and 3 should begin with a review of selected recent evaluations focused on gender equality and women’s empowerment, with the final outcome from option 2 being a guidance note or “toolbox” of innovative and effective evaluation approaches drawn directly from the review. AusAID proposes to contract a consultant to undertake this review, using a methodological

* This copy of the TORs does not include the original database (Annexes 1 and 2) or the work schedule referred to in paragraph 14, which are now out of date.

¹ Extract from Working Party on Aid Evaluation “Summary Record of the 35th Meeting held on 14-15 November 2001” DCD/DAC/EV/M(2001)2/PROV, point 8.

framework which integrates elements from each of the 3 options. This approach is responsive to Members' support for all options.

5. WP-EV has undertaken 2 previous reviews of attention to gender issues in evaluations. The most recent review culminated in a workshop organised by Sida in 1999, and drew on a database of Member evaluations up to 1998.² The current review will build on this work.

Database for the review

6. The database for the current review includes 2 broad types of evaluations:

- i) Thematic evaluations conducted since 1999 which have been **specifically designed to evaluate gender equality, gender mainstreaming and women's empowerment efforts**. WP-EV's matrix of planned evaluations indicates that between 16 and 21 evaluations may be available for review (see Annex 1), depending on progress with the finalisation of reports.
- ii) Evaluations conducted since 1999 which have **not** been specifically designed to evaluate Members' gender equality policies, but which have included significant gender analysis. Selection of evaluations in this category will give priority to evaluations which: have developed or applied innovative approaches or tools; add to our knowledge of institutional approaches; and/or which demonstrate the impact of including gender perspectives on overall development results. WP-EV's Inventory of evaluations indicates that up to 66 evaluations may be available. It is expected that a number of evaluations will be culled from the review after initial assessment. Assistance from Members with the selection of evaluations which meet these criteria, using the list attached in Annex 2 as a starting point, will expedite the selection process.
- iii) Other material not available on the inventory would also yield relevant and important information on issues, such as reviews, lessons learned and other assessments conducted by agencies on gender analysis and quality.

7. Members who have undertaken evaluations in these categories will be requested to send final or draft reports by email to AusAID's consultant by 15 March 2002. It is expected that this database will capture evaluation experience for activities where gender perspectives have not been mainstreamed, as well as activities where gender equality and

² OECD 1994 "Assessment of DAC Members' WID Policies and Programmes", Paris; Woodford-Berger, Prudence 1999 "Evaluating Gender Equality and Women's Empowerment: A DAC Review of Agency Experiences 1993-1998" Final Draft Report, Prepared for the DAC Working Party on Aid Evaluation; Keller, Bonnie 2000 "Report on the DAC Workshop on Evaluating Gender Equality and Women's Empowerment, 25-26 November 1999, Stockholm, Sweden" Prepared for the DAC Working Party on Aid Evaluation.

women's empowerment were explicit as a primary or secondary objective. In order to undertake the review, full copies of evaluation reports will be required, rather than abstracts. Due to resource constraints, only reports available in English will be included in the Review.

8. Following review of the material and other documentation from agencies, the study will conduct interviews, both electronic and through site visits, to discuss aspects of the tools and methodologies not evident from the published reports. Discussion with those who have used various tools and approaches would allow more in-depth analysis of the variables affecting gender analysis and the issues and gaps in information that may need to be addressed at later stages in the study. A questionnaire may be a useful means of eliciting this information.

Review framework

9. The methodological framework for the review has a positive focus on distilling and sharing of experiences in the following areas, which integrates the 3 options for future work proposed in the Secretariat's Discussion Paper:

- i) Innovative, practical and effective evaluation methodologies, approaches and tools for evaluating gender equality and women's empowerment in Member programs and projects (option 2).

The need for guidance on gender sensitive evaluation approaches has been highlighted in previous work for WP-EV.³ The review will focus on: the effectiveness of different methods for evaluating gender equality and women's empowerment across programs, and for different sectors and types of assistance (depending on the database); quantity and quality of information and analysis, including the use of indicators to assess participation, changes in gender equality, women's empowerment and gender relations; innovative and participatory methods for collecting information (particularly where data is scarce, and where gender perspectives have not been mainstreamed through design, implementation and monitoring); how gender analysis is linked to key evaluation criteria, and to the assessment and attribution of outcomes, benefit and impact.

- ii) Institutional approaches and change mechanisms which have successfully responded to recurrent obstacles to gender mainstreaming (option 1).

³ Woodford-Berger 1999; Keller 2000; and Hunt, Juliet 2001 "Discussion Note for the DAC Working Party on Aid Evaluation: Options for Further Work on Evaluating Gender Equality and Women's Empowerment" Prepared for AusAID for the Working Party on Aid Evaluation Meeting, 22-23 May 2001.

There is an extensive literature on institutional obstacles. However, good practice case studies on the institutionalisation of gender equality approaches are often based on intention, rather than outcome, and few are based on evidence from program and project evaluation. The review will distil experience on addressing institutional blocks and strategies from both donor and partner perspectives, and in relation to different levels of institutionalisation and types of assistance.⁴ This will assist to develop Terms of Reference for further work needed on Option 1.

- iii) Assessing whether gender analysis and gender mainstreaming improves the achievement of overall development results (option 3).

If this aspect of the Review is to yield significant results, it will be important for Members to identify evaluations which make the link between including gender perspectives, and successful or improved development outcomes. The size of the evaluation database for this area is unknown. Of the 3 areas of focus for the review, this presents the most methodological challenges. Depending on information available in the database, the review will attempt to distinguish between activities where activity design has been improved by gender analysis, and activities where a gender sensitive approach to implementation has improved development outcomes, whether or not activity objectives or design were modified to address gender equality concerns.

Output from the review

10. The initial output from the review will be a report (approximately 20 pages, in addition to annexes) which summarises findings in each of the 3 areas above, and which identifies ways forward to complete work on each option. Close attention will be given to identifying steps needed to achieve a relevant and accessible output for each option, which can be disseminated and utilised by evaluation and gender specialists, and other stakeholders in the development process.

11. The WP-GEN will be approached to provide guidance and peer review comments on the study and to facilitate contacts with members in collection and discussion on the material and reports. A progress report will be prepared for members and circulated for comment prior to the May 2002 meeting of the Working Party on Aid Evaluation.

⁴ These issues were raised by Members in comments on the Secretariat's Discussion paper. The extent to which it is possible to focus on these areas (donor and partner perspectives, and different levels of institutionalisation) will depend on attention to these issues in the evaluation database.

Further steps

12. The Discussion Paper recommended that workshops be held to discuss findings from initial work for options 1 and 3. For option 2, discussion of review findings with agency staff and evaluators involved in gender evaluation studies was also seen as an essential step, before finalising a discussion note or toolbox on innovative approaches and tools for gender sensitive evaluation. With an integrated approach to the review of evaluations, an integrated workshop agenda is considered the best option. Collaboration with a Member other than Australia, both in the development of the study and in the planning and organisation of the workshop would be sought.

13. Other donors may also take the lead role in finalising outputs for each option. These were summarised in the Discussion Paper as:

- i) Evaluation methodologies: Development of a discussion note on recommended innovative approaches/tools.
- ii) Institutional approaches: Short policy note or communiqué. The review report will identify further work which needs to be undertaken prior to the workshop on Gender and Evaluation.
- iii) Improving development results: A study which synthesises the implications of gendered approaches. The review report will identify any further work which needs to be undertaken prior to the workshop on Gender and Evaluation.

Scheduling of the review and follow-up work

14. The table below summarises the timetable of work for the review and makes suggestions for scheduling further follow-up work.

Annex B Documents Submitted to the Review ¹

Table B1 **Thematic evaluations on gender equality, gender mainstreaming and women's empowerment**

Agency (year)	Title of evaluation report	Type of evaluation
ADB (2001)	"Special Evaluation Study on Gender and Development", OED, October 2001	Policy
AusAID (2002a)	"Gender and Development: GAD lessons and challenges for the Australian aid program", Published report, September 2002	Policy audit
AusAID (2001)	"Gender and Development Review: Findings and Recommendations", Unpublished report	
Belgium (2001a)	"Report of the Special Evaluator for International Cooperation to the Belgian Parliament", June 2001 "Terms of Reference for the Evaluation of Poverty Reduction Incidences, Part 1: An external evaluation of the gender aspects in a selected number of projects of the Belgium Survival Fund and their incidence on poverty reduction"	Cluster
Belgium (2001b)	"Evaluation of Belgium Survival Fund Support to Kenya Women's Finance Trust (KWFT)", Vol. I & Vol. II (Annexes)	Project
Belgium (2001c)	"An External Evaluation of Gender and Its Influence on Poverty Reduction: Hoima-Kibaale Integrated Community Development Project, Uganda, June–August 2000"	Project
Belgium (2001d)	"Project for Improving Household Food Security and Nutrition in the Luapula Valley of Zambia (IHFSAN)"	Project
Belgium (2001e)	"Evaluation Thematique Genre et Pauvrete du Pim Niger (Programme integre Maradi)"	Project
Belgium (2001f)	"Evaluation Thematique Genre/Reduction de la Pauvrete du projet Fenu au Mali: Aménagement et gestion des terroirs villageois du Sené-gondo"	Project
BMZ (2001b)	"The Integration of Gender Issues in PRSPs: The Example of Ghana", Birte Rodenberg, Deutsche Institut für Entwicklungspolitik, December 2001, German language, Executive summary in English	PRSP review
BMZ (2000)	"German Development Cooperation: Summary of the Series Evaluation 'Gender-specific differentiation of the target group' in selected FC/TC projects - trans-sectoral analysis" (English Summary) "Geschlechterspezifische Differenzierung der Zielgruppe in ausgewählten F2/T2 – Vorhaben"	Cluster

(Continued on next page)

¹ Some agencies did not submit reports to the review. In some cases the consultants relied on web searches of agency evaluation reports and read all reports with relevant titles. Only those reports that included some references to women or gender issues are included in this list, which constitutes the database for the review. Thematic evaluations from multilateral agencies are also included in Table B1.

Table B1 **Thematic evaluations on gender equality, gender mainstreaming and women's empowerment** (continued)

Agency (year)	Title of evaluation report	Type of evaluation
CIDA (2002a)	"Self-management and sustainability: An Impact Study – Study on PADEL's Contribution to the Reduction of Poverty and Inequities between Men and Women" (Executive Summary)	NGO project
CIDA (2002b)	"Summary of the Systemization Document: Incorporating gender equity into cooperative organizations in Guatemala: a local development initiative", Socodevi, February 2002	
CIDA (2001a)	"Gender Equality Review", Prepared for Indonesia Program, CIDA, Goss Gilroy Inc, Internal document, not available for circulation	Policy audit
CIDA (2000a)	"Caribbean Gender Equity Fund Performance Review: Umbrella Report, Final Report", Universalia	Gender equality fund
CIDA (2000b)	"Review and Monitoring Report: Social and Women's Initiative Project Phase I", Veronica Huddleston	Gender equality fund
CIDA (1999a)	"Evaluative Review of Ghana Girl Child Education Project (GGCEP)", Patrick Cummins and Akosua Anyidoho	Women's project
CIDA (1999b)	"Evaluation: Women's Initiatives Fund (WIF) Project, Phase II, Project No: 344/14537", Mary Lynch, Ali Anwer and Wafik Arif	Women's project
CIDA (1999c)	"South America Regional Gender Equity Funds: Evaluation Report", Juanita Barreto, Marisela Benavides, Celsy Campos, Paola Cappellin, María Cui, Christine Ouellette, and Ana Quiroga	Gender equality fund
CIDA (1999d)	"Final Report: Assessment of the Society of Friends of Nepal (SOFON)", Sheila Robinson and Jamal Devi Shrestha, July 1999	NGO women's project
DFID (2000a)	"Promoting Equality Between Men and Women", SD Scope Paper No. 2, Nazneen Kanji and Sarah Salway	Cluster
DFID (2000b)	"A Participatory Gender Review and Support Consultancy for DFID's Rural Livelihoods Projects", Eva Jordans, Faria Zaman and Amita Dey	Cluster
DFID (2000c)	"Mainstreaming Gender Through Sector Wide Approaches in Education", Andy Norton, Mo Sibbons, David Smawfield, Helen Poulsen, Amanda Gibbard and Amanda Seel	SWAP review
DFID (1999a)	"Mainstreaming Gender Equality in Project Implementation: Botswana, Pakistan and South Africa", Helen Derbyshire	Cluster
ILO (2002)	"ILO Gender Audit 2001-02: Final report", Internal Report of the Bureau for Gender Equality, Geneva, May 2002	Policy audit
JICA (2001)	Chapter 2, Section II Thematic Evaluation: 'Poverty and Gender in Agriculture, Forestry and Fisheries Cooperation (Nepal)'; and 'Poverty and Gender in Agriculture and Forestry Cooperation (Paraguay)' in "Annual Evaluation Report FY 2001", December 2001	Cluster
JICA (2000)	Chapter 3, Section II Thematic Evaluation: 'Women in Development (WID)/Gender (Sri Lanka)' in "Annual Evaluation Report FY 2000", June 2000	Cluster

(Continued on next page)

Table B1 **Thematic evaluations on gender equality, gender mainstreaming and women's empowerment** (continued)

Agency (year)	Title of evaluation report	Type of evaluation
Netherlands (1998) ^a	"Women and Development: Policy and Implementation in Netherlands Development Cooperation 1985-1996", Policy and Operations Evaluation Department "Women in Kenya and the Netherlands Development Cooperation 1985-1995" "Women in Burkina Faso and the Netherlands Development Cooperation"	Policy
NORAD (2001)	"Review of the Regional Diploma Course in Women's Law, Zimbabwe", Patricia Kameri-Mbote, Mette Topnes and Bodil Maal, September 2001	Project
NORAD (2000)	"Joint Review, Women's Economic Programme Phase II (India)", September/October 2000	Project
NORAD (1999a)	"Evaluation of the Gender in Development Division (GIDD) at the Cabinet Office, Zambia"	Project
NORAD (1999b)	"WID/Gender Units and the Experience of Gender Mainstreaming in Multilateral Organisations", Gisela Geisler, Bonnie Keller and Anne-Lene Norman, March 1999	Institution-alisation review
NZAID (2002a)	"... a wind of change: review of NZODA gender and development project in Kiribati", Patti O'Neill and Rose Namoori-Sinclair, February 2002	Project
NZAID (2002b)	"Review of NZODA support for gender mainstreaming in the Mekong River Commission, Report on Review Mission (DRAFT)", Helen McNaught, May 2002	Project
SDC (2000)	"Validacion de la Fase Inicial y Definicion de la Segunda Fase de la Estrategia de Genero Corlap, Cosude Bolivia", José Baldivia, Chrystel Ferret and Sylvia Cardona, September 2000	Project
Sida (2003)	"Reflection on Experiences of Evaluating Gender Equality", Sida Studies in Evaluation 03/01	Policy
Sida (2002a)	"Mainstreaming Gender Equality: Evaluation of Sida's Support for the Promotion of Gender Equality in Partner Countries", Britha Mikkelsen, Ted Freeman, Bonnie Keller et al., Sida Evaluation Report 02/01	
Sida (2002c)	"Sweden's and Holland's Strategies for the Promotion of Gender Equality Policy in Bolivia"	Project
Sida (2001a)	"Swedish-Danish Fund for the Promotion of Gender Equality in Vietnam", Sida Evaluation 01/17, Shashi R. Pandey, Darunee Tantiwiranmanond and Ngo Thi Tuan Dung	Gender equality fund
Sida (2000a)	"Empowerment of Women through Panchayati Raj in Rajasthan and Orissa, India", Sida Evaluation 00/31, D.K. Manavalan	Project
Sida (2000b)	"Vietnam Women's Union: Promoting Gender Equality", Sida Evaluation 00/16, Wanjiku Kaime-Atterhög and Tran Thi Van Anh	Gender equality fund

(Continued on next page)

Table B1 **Thematic evaluations on gender equality, gender mainstreaming and women's empowerment** (continued)

Agency (year)	Title of evaluation report	Type of evaluation
UNIFEM (2002)	"Desk Review: Gender Mainstreaming in the CCA and UNDAF Processes – Presented to the Inter-Agency Taskforce on Gender Mainstreaming in the CCA/UNDAF Process, Inter-Agency Network on Women and Gender Equality (IANWGE)", Alicia Mondesire, June 2002	Institution-alisation review
USAID (2001a)	"Final Evaluation Survey of the Women's Empowerment Program (WEP)", Submitted to PACT by Ava Darshan Shrestha and Janardan Khatri-Chhetri, September 2001	NGO women's project
USAID (2001b)	"Aftermath: Women and Women's Organisations in Postconflict Societies, The Role of International Assistance", Krishna Kumar, USAID Program and Operations Assessment Report No. 28, Center for Development Information and Evaluation, July 2001	Cluster
USAID (1999a)	"Midterm evaluation: developing Mayan-based health care for rural women and children, Project Concern International", Melody Trott and Barbara Schieber, September 1999	Women's project
USAID (1999b)	"Freedom from Hunger Credit with Education for Women Program, Final Evaluation", C. Stark Biddle, March 1999	NGO women's project
USAID (1999c)	"Evaluation of USAID/OTI's (Office of Transition Initiatives) Women in Transition Initiative in Rwanda", Hannah Baldwin and Catherine Newbury, Office of Transition Initiatives and the Center for development Information and Evaluation, USAID	Gender equality fund
USAID (1999d)	"More, But Not Yet Better: An Evaluation of USAID's Programs and Policies to Improve Girls' Education", Chloe O'Gara, Sharon Benoiel, Margaret Sutton and Karen Tietjen, USAID Program and Operations Assessment Report No. 25, June 1999	Cluster
World Bank (2002)	"The Gender Dimension of Bank Assistance: An Evaluation of Results", Report No. 23119, Operations Evaluation Department (OED), 17 January 2002	Policy
World Bank (2001a)	"Integrating Gender in World Bank Assistance", Report No. 23035, OED, 25 October 2001	
WFP (2002)	"Final Evaluation of WFP's Commitments to Women 1996-2001, Full Evaluation Report, Final Report", Camillia Fawzi El-Solh, 26 May 2002	Gender institution-alisation review

^a At its May 2002 meeting, the Working Party on Aid Evaluation agreed to include reports produced prior to 1999 in special cases. The innovative methodology used qualified these reports for special consideration.

Note: Two thematic evaluations ongoing at the time of the review were not included in the review's database: Danida, "Terms of Reference: Evaluation /Impact Study of Four Training Projects for Farm Women in India" (report available mid-2003); and EC, "Thematic Evaluation of the Integration of Gender in EC Co-operation with Third Countries: Terms of Reference (Final)" (report available early in 2003). One evaluation was received too late to incorporate findings: DFID, "DFID China Gender Review", Nazneen Kanji and Du Jie, Department for International Development, UK, January 2003.

Total number of thematic evaluations: 42

Total number of DAC member agencies with thematic evaluations: 12

Table B2 **General evaluations that included references to gender issues or gender analysis**

Agency (year)	Title of evaluation report	Type of evaluation
AusAID (2002b)	"Water and Primary Health Care for Tibetan Villagers", Quality Assurance Series No. 29, February 2002	Project
AusAID (2000a)	"Evaluation of Australian Government Funded NGO Projects in Africa", Quality Assurance Series No. 25, December 2000	NGO program
AusAID (2000b)	"Developing Integrated Rural Health Care Systems: An Evaluation of the Kadavu Rural Health Project in Fiji", Quality Assurance Series No. 24, December 2000	Project
AusAID (2000c)	"Developing Rural Communities in Marginal and Rainfed Areas: Contributions of Agricultural Projects in the Philippines", Quality Assurance Series No. 23, September 2000	Cluster of 3 projects
AusAID (2000d)	"Qinghai Community Development Project: Evaluation Report", Quality Assurance Series No. 21, June 2000	Project
AusAID (2000e)	"Improving Access to Land and Enhancing the Security of Land Rights: A Review of Land Titling and Land Administration Projects", Quality Assurance Series No. 20, September 2000	Sector
AusAID (2000f)	"Increasing Rural Incomes: An Evaluation of Three Rural Sector Projects in Samoa", Quality Assurance Series No. 19, July 2000	Cluster of 3 projects
AusAID (2000g)	"Assisting Local Communities: Evaluation of Government Funded NGO Projects in Vietnam", Quality Assurance Series No. 18, March 2000	NGO program
AusAID (1999a)	"Wells, Taps and Toilets: Safe Water and Sanitation for Eastern Indonesia", Quality Assurance Series No. 17, November 1999	Project
AusAID (1999b)	"Targeting Poor Farmers: Contributions to Rural Development in Thailand", Quality Assurance Series No. 16, June 1999	Cluster of 3 projects
AusAID (1999c)	"Growing Rice and Protecting Forests: An Evaluation of Three Food Production Projects in South East Asia", Quality Assurance Series No. 15, June 1999	Cluster of 3 projects
AusAID (1999d)	"Building for Development: An Evaluation of Four Infrastructure Projects in Tonga and Samoa", Quality Assurance Series No. 14, May 1999	Cluster of 4 projects
BMZ (2001a)	"Main Report on the Series Evaluation: Prospects of Success of Basic Education Projects" (English version) "German Development Cooperation with Senegal: Summary of the Sector Evaluation 'Prospects of Success of Primary Education Projects', Partial Evaluation Senegal (Summary of the Sector Evaluation)" (English Summary)	Sector

(Continued on next page)

Table B2 **General evaluations that included references to gender issues or gender analysis** (continued)

Agency (year)	Title of evaluation report	Type of evaluation
CIDA (2001b)	"Infrastructure Services Performance Review: Executive Report", Performance Review Branch, October 2001	Sector
CIDA (2001c)	"Infrastructure Services Performance Review: Background Report", Valerie Young, Philip Chan and Chris Stanley, Performance Review Branch, October 2001	
CIDA (2001d)	"Infrastructure Services: Policy Dialogue Study, International Telecommunication Union (ITU), Evaluation Report 2001", John Gilbert, Performance Review Branch (May 1999) "Infrastructure Services: Global Knowledge, A Policy Dialogue Study, Canada's Role at the Global Knowledge 97 Conference and in the Global Knowledge Partnership, Consultation Draft, 2001", Charles Morrow and Associates, Performance Review Branch (July 1999)	
CIDA (2000c)	"Canada-Thailand Trilateral Environment (CTTE) Project: End of Project Review, May-June, 2000", Fred Meth, September 2000	Project
CIDA (2000d)	"Regional Aids Training Network, Mid-Term Evaluation, Final Report", Ron Titus, Joe Decosas and Bruce Waring, August 2000	Project
CIDA (2000e)	"Evaluation of Inter Pares' Program in Peru: Final Report", Catherine Gander, NGO Division of CIDA, May 2000	Institutional NGO review
CIDA (2000f)	"Final Report on Evaluation of Inter-Church Action for Development, Relief and Justice for CIDA", John Doran, March 2000	Institutional NGO review
CIDA (2000g)	"Final Program Evaluation Report, Presbyterian Church in Canada: Presbyterian World Service and Development (April 1st 1997 – March 31st 2000 Program)", Ursula Mount, Canadian Partnership/NGO Division, March 2000	Institutional NGO review
CIDA (2000h)	"India Oilseeds Development: End of Project Evaluation, Final Report", Ron Titus, Bob Jiggins, Nancy Higginson, Elise Rafuse and Sanjay Phansalkar, January 2000	Project
CIDA (1999e)	"Final Report, Operation Eyesight Universal (OEU) Institutional Assessment", Sheila Robinson and Philip Cox, NGO Division, CIDA, December 1999	Institutional NGO review
CIDA (1999f)	"Evaluation: Strengthening of Civil Society and Rural Development in Nampula Province, 1997/99, Cooperation Canada Mozambique (COCAMO), Final Report", Brigitte Bagnol, September 1999	Institutional NGO review
CIDA (1999g)	"Private Enterprise Accelerated Resource Linkage (PEARL): Mid-term Review Report", Raynald Brassard and M.M. Lynch Consultants International Inc, April 1999	Project
Danida (2002a)	"In the Wake of a Flagship, The Noakhali Project in Bangladesh"	Program
Danida (2002b)	"Danish Assistance to Vocational Education and Training"	Sector
Danida (2001a)	"Private Sector Development Programme", October 2001	Program
Danida (2001b)	"Evaluation: Danish-Burkinabè Development Cooperation 1996 – 2000" (2001/4)	Country program
Danida (2000a)	"Evaluation of Danish Bilateral Assistance to Health 1988 – 1997" (2000/4)	Sector
Danida (2000c)	"Rakai District Development Programme in Uganda"	Program
Danida (1999)	"Danish Support to Promotion of Human Rights and Democratisation"	Sector program

(Continued on next page)

Table B2 **General evaluations that included references to gender issues or gender analysis** (continued)

Agency (year)	Title of evaluation report	Type of evaluation
DFID (2002a)	"From Projects to SWAps: an Evaluation of British Aid to Primary Schooling, 1988-2001", Samer Al-Samarrai, Paul Bennell, Christopher Colclough, September 2002	Program
DFID (2001b)	"ODA/DFID Support to Health Sector Reform and Health Management: Synthesis Study", Andrew Cassels and Julia Watson, January 2001	Sector
EC (2001a)	"Evaluation of Voter Education in the Context of EU Electoral Support, Final Report", Karen Kenny, Sophie Lagueny and Florence Burban, September 2001	Sector
EC (2001b)	"Synthesis Report on EC Activities in the Field of Human Rights, Democracy and Good Governance, Synthesis Note, 10 August 2001", Emery Brusset, Emma Achilli, Christine Tiberghien, Evaluation Unit of the EuropeAid Co-operation Office	Sector
EC (2001c)	"Rapport du Synthese des Activites de la Communaute Europeenne dans le Champ des Droits de L'Homme, de la Bonne Gouvernance et de al Democratie: Rapport de référence, 10 août 2001", Emery Brusset, Emma Achilli, Christine Tiberghien	
Netherlands (2002)	"Health, Nutrition and Population, Burkina Faso, Mozambique, Yemen – Evaluation 1995-1999", IOB Evaluation No. 288, Policy and Operations Department, May 2002	Program
NORAD (1999c)	"Performance Assessment of IPPF: Policy and Effectiveness At Country and Regional Levels – Synthesis Report", Julie Skjaeraasen, Bo Stenson and Ian Thomas, March 1999	Program
SDC (1999a)	"Strengthening of Vocational Training Centers (SVTC) in Vietnam. Joint Evaluation. Final Report"	Project
SDC (1999b)	"Country Programme South Africa 1994-1999 Evaluation of the Activities of the Swiss Department of Foreign Affairs"	Country program
SDC (1999c)	Social Forestry Support Project, Phase 2. Report on Mid-Term Review, 15 October 1999	Project
SDC (1999d)	External Evaluation of Southern Highlands Dairy Development Project (SHDDP)	Project
SDC (1999e)	Penipe Project Phase III Evaluation Report	Project
Sida (2001b)	"Democracy and Human Rights: An evaluation of Sida's support to five projects in Georgia", Sida Evaluation 01/11, Birgitta Berggren and Patrik Jotun	Cluster of 5 projects
Sida (2000c)	"Strengthening Democracy on the Atlantic Cost in Nicaragua", Sida Evaluation 00/19, Hans Peter Buvollen, Mario Rosales Ortega and Leticia Velásquez Zapeta	Project

Total number of general evaluations: 43

Total number of DAC member agencies with general evaluations: 10

Table B3 **Tools and other documents on gender-sensitive evaluation methodology**

Agency (year)	Title of document
APC/WNSP (2002?)	"GEM – A Guide to Integrating a Gender Analysis into Evaluations of Initiatives that Use Information and Communication Technologies"
CIDA (2002c)	"Draft Framework for Gender Equality Integration into Programs for CIDA-CPB Program Partners", May 2002
CIDA (2002d)	"Draft Checklist for Gender Equality Integration into Projects for CIDA-CPB Project Partners"
CIDA (2001e)	"How to Perform Evaluations: Gender Equality", Performance Review Branch, No. 4, May 2001
CIDA (2000i)	"CIDA Evaluation Guide – Work in Progress", Performance Review Branch, January 2000, and other guides in the "How to Perform Evaluations" series
CIDA (2000j)	"A Results Approach to Developing the Implementation Plan: A Guide for CIDA Partners and Executing Agencies", March 2001
CIDA (2000k)	"RBM Handbook on Developing Results Chains: The Basics of RBM as Applied to 100 Project Examples", Results-Based Management Division, December 2000
CIDA (1997)	"Guide to Gender-Sensitive Indicators", Tony Beck and Morton Stelcner, August 1997
DFID (2002b)	"Gender Manual: A Practical Guide for Development Policy Makers and Practitioners", April 2002
ILO (1995)	"Guidelines for the Integration of Gender Issues into the Design, Monitoring and Evaluation of ILO Programmes and Projects", Evaluation Unit, Bureau of Programming and Management, January 1995 <ul style="list-style-type: none"> • Checklist 2: Gender considerations in progress review reports • Checklist 3: Gender considerations in self-evaluation reports • Checklist 4: Gender considerations in terms of reference (TORs) for independent evaluation missions
World Bank (2001b)	"Gender in Monitoring and Evaluation in Rural Development: A Tool Kit"

Table B4 **Evaluation reports from other agencies, and other reports**

Agency (year)	Title of document received
AfDB (2002)	"Monitoring and evaluating gender issues and women's empowerment in Kenya", Paul A. Ogula in "Monitoring and Evaluation Capacity Development in Africa", Selected proceedings from a seminar and workshop organised by the Development Bank of Southern Africa, the African Development Bank and the World Bank, Johannesburg, 25-29 September 2002
Bridge (2001)	"Gender and Monitoring: A Review of Practical Experiences", Bridge Report No. 63, Paola Brambilla, 2001
DFID (2002c)	"Sector Wide Approaches: Opportunities and Challenges for Gender Equity in Health", Sally Theobald, Rachel Tolhurst and Helen Eelsey
DFID (2002d)	"What's Behind the Budget? Politics, Rights and Accountability in the Budget Process", Andy Norton and Diane Elson
DFID (2001a)	"First Annual Review of the Gender Equality Strategy, Progress on Country Strategy Objective on Gender Equality", Lina Payne
DFID (2000d)	"Gender Mainstreaming: Emerging Lessons from Ghana", Elizabeth A. Akpalu, Esther Ofei-Aboagye and Helen Derbyshire
DFID (2000e)	"Gender Equality Strategy for DFID Bangladesh", Julie Lawson-McDowall
DFID (1999b)	"Report on the First Phase of DFIDB Gender Review", Anne Coles
Heinrich Böll Foundation (1999)	"Empowerment, A study of Women's Projects Abroad", Birte Rodenberg & Christa Wichterich, Heinrich Böll Stiftung, November 1999
Netherlands (2001b)	WP-GEN Task Force on Gender Equality in SWAps – Reference Guide & Case Studies on Gender Mainstreaming in Bangladesh: Annex 1 – Experiences in health sector programs Annex 2 – Experiences in education sector programs Annex 3 – Experiences in agricultural sector programs
Novib (2001)	"En Route: Evaluation of the Gender Route Project", Ria Brouwers and Donny Meertens, October 2001
OCAA (2001)	"The Final Workshop Report on Philippines Field Gender Audit, 21–30 August 2001"
SPC (2001)	"Review of the Pacific Women's Resource Bureau, Secretariat of the Pacific Community", August 2001
Netherlands	"Good Practice in the Development of PRSP Indicators and Monitoring Systems", David Booth and Henry Lucas, Overseas Development Institute, July 2002 "Recommendations for Integrating Gender into the Poverty Reduction Strategy Paper (PRSP) and Related Processes", Tanzania Ministry of Community Development, Women's Affairs and Children, no date "Information Pack on Key Gender Issues in Different Sectors", Mary Rusimbi, Gemma Akilimali, Tanzania Gender Networking Programme (TGNP), Debbie Budlender, Community Agency for Social Enquiry (CASE)

Annex C Summary of Indicators Used in Evaluations¹

A brief summary of the types of indicator used in evaluation reports is presented in the following tables. These are: gender mainstreaming indicators useful at the agency level, indicators of women's and men's participation, and results indicators that describe benefits for females and males.

Table C1 **Indicators used in thematic reports**

Agency & report	Gender mainstreaming indicators	Participation indicators	Results indicators
AusAID (2001) Gender mainstreaming policy audit	Focus on project design, including project objectives, partner capacity, assessment of socio-economic & political context, identification of constraints & strategies, resources, monitoring, & capacity of contractor.	Women and men involved in data collection: women's involvement in decision making groups.	Women's perception of benefits: practical gender needs & strategic gender interests addressed.
CIDA (2002a) Padel (NGO credit & training project)		Women members of community-based organisations/groups; active women members (attendance at meetings); women in executive positions; women participants in training; % of active women members participating in training; women with access to credit; organisations with women's committees.	
CIDA (2001a) Gender mainstreaming audit of a country program	A 3-page questionnaire – can be reformulated as indicators. Focus is on attention to gender issues in project documents. Includes quality of analysis, baseline study, gender strategy, resources, MOU. Also a 4-page questionnaire for executing agencies.		Focuses on whether project documents identify expected gender equality results at output, outcome & impact levels.
CIDA (1999a) Ghana Girl Child Education Project	Focus is on attention to gender issues in education plans, policies, existence of a permanent unit within the Ministry of Education, an information system on girls' education & transmission of information to Ministry of Education.	Number of female teachers; participation in classroom activities	Retention rates, positive role models, attention to female students, greater participation of women in community decision making, reduction in barriers to girls' education, gender-sensitive curricula & teaching materials, gender-sensitive teachers & education officers, gender-sensitive infrastructure.

(Continued on next page)

¹ Many evaluations included questions that could be reformulated as indicators. These are not included here except where noted. Other evaluations (eg ADB 2001 in Table B1) developed specific gender-sensitive indicators for each activity, but these were not included in the ADB evaluation report.

Table C1 **Indicators used in thematic reports** (continued)

Agency & report	Gender mainstreaming indicators	Participation indicators	Results indicators
CIDA (1999b) Women's Initiative Fund Egypt		Number of enterprises created, loan funds disbursed, jobs created.	
ILO (2002) Gender Audit, Annex VIII Document Analysis Sheet	Overall attention to gender issues in documents, including analysis of context, gender mainstreaming in objectives, indicators & budget, technical co-operation, capacity building in the ILO work unit, knowledge & information management in the organisation, M&E systems, choice of partner agency, products & public image, human resources, & organisational culture.	Participation and decision making by women in meetings.	
USAID (2001a) Women's Empowerment Program, Nepal			Empowerment indicators include decision making for self-needs, children's needs family needs & for community activities. Decision-making indicators distinguish between sole decision making, joint decision making, participation in decision making but with less than decisive influence, and participation with no influence. Indicators for expenditure include direct wellbeing, indirect wellbeing, and not related to wellbeing.
USAID (1999d) Girls' Education	Covers regulations regarding pregnancy & age, policy/practice regarding repetition, policy frameworks for quality improvement.		Access indicators focus on enrolment, completion & retention, availability & appropriateness of facilities. Indicators for quality need to be extrapolated from the findings of the report, which focus on quality of instruction, curricula, family & teacher support, & engagement with communities.
World Bank (2001a, 2002) Evaluation of gender policy results	Annex IIA of the 2001 report includes indicators & rating systems for attention to gender issues in a range of documents including country assistance strategy, poverty assessments, public expenditure reviews, sector reports & project appraisal documents. Overall indicators for relevance included: <ul style="list-style-type: none"> • soundness of underlying gender analysis work • adequacy of integration of gender issues into country assistance strategy – understanding of critical gender issues; indicators & benchmarks; realism of the strategy proposed in the country context. 		Results of bank assistance were assessed according to 3 broad criteria/questions: <ul style="list-style-type: none"> • increased human capital – outcomes in education, sustainability in education, results in health, sustainability in health, institutional strengthening in health & education sectors • increased participation of women in economic development – increased opportunities for income generation, increased access to credit & other services, increased participation in training or skills upgrading • improved/strengthened institutional framework – strengthened development institutions to deliver gender aware programs, strengthened NGOs or community groups.

Table C2 Indicators used in general reports

Agency & report	Gender mainstreaming indicators	Participation indicators	Results indicators
AusAID (2000g) NGO program in Vietnam	Strategies incorporated to address "WID/GAD issues" in the planning & implementation of the project.	Efforts were made to involve women in project activities.	The project had a positive impact on women.
AusAID (1999a) Water & sanitation in Indonesia		Organised community structures (for management of water facilities at village level) have adequate representation of rich & poor, men & women.	A sufficient majority have access to water & sanitation facilities (rich/poor, men/women).
AusAID (1999b, p. 39) Rural development in Thailand		Indicators for achievement of "WID/GAD objectives" were: women have a say in development decision making; increased women's participation in implementation.	Impact was measured by: enhanced women's role and status in public.
CIDA (2000e, pp. 34-35) Inter Pares Peru (NGO program)	Indicators for institutional results include: adoption of gender policy, strengthening capacity of women's organisations, & strengthening gender focus in other counterpart organisations.	Number of women trained (in various areas, including training of women leaders), forming micro-enterprises, using seed banks, educated about health issues.	List of indicators focuses on developmental results (economic, social, political) & individual results (increased self-esteem & confidence). Economic: increased yields & income. Social: improved nutrition, decreased rates of domestic violence, improved awareness of domestic violence in the community. Political: improved capacity of women's leaders, improved gender perspectives in decision making, increased ability of women's organisations to negotiate with government, increased respect for women leaders (from family, community).
CIDA (1999e, pp. 113-121) NGO eyesight program	Gender policy adopted, supported & understood within donor & partner NGOs, sex-disaggregated data collected on management & administration activities, gender balance in volunteer profile, board is representative of both men & women.	Degree to which women participate as full & equal partners throughout & after the project.	

Annex D Gender Audits

Gender auditing, a new area of evaluation methodology, was pioneered by NGOs, drawing on social auditing principles rather than those of financial auditing (Hunt 2000). Distinguishing features of a gender audit are accountability for gender equality and mainstreaming policy commitments, and a broad agency or program scope of enquiry.

Four gender audits were submitted to the review – two from DAC members (AusAID and CIDA)¹ and two from other agencies (ILO and OCAA).² The AusAID, ILO and OCAA audits involved groups of staff in assessing mainstreaming processes, whereas the CIDA audit used questionnaires to gather data on mainstreaming.

Two of the gender audits (those from the ILO and OCAA) are examples of participatory self-assessment methods using facilitated workshops. Both were designed to promote learning by participants and ownership of findings in addition to policy accountability, in the expectation that this would lead to direct action to improve gender mainstreaming approaches. This is a major strength of these approaches, although it is too soon to assess whether they will fulfil their promise as effective methods for promoting action.

One weakness of the audit methodologies is that their use of standard evaluation criteria (relevance, effectiveness and sustainability) is not explicit. Links to development results are also weak in most gender auditing approaches, particularly the assessment of the impact on women and on gender relations. Beneficiaries are not generally involved in the process, although OCAA's method is an exception to this.

The **ILO** audit included a desk review of documents and publications, a series of participatory 2-day gender audit workshops in 15 ILO work units at headquarters and in the field, interviews with key staff and meetings with constituents, partners and women's organisations. The participatory self-assessment approach contributed to team building, and identified a number of more general organisational learning and management issues that senior management are committed to address.

The Gender Audit has set in motion for the first time a collective process of gender analysis ... enhanced capacity for gender analysis ... and identified good practices and ways of strengthening performance. Some participants made important discoveries that will change the ways they work: for instance that incorporating the gender dimension in HIV/AIDS work means putting men back into the picture; or that doing gender or women-specific projects without a gender analysis ... falls short of genuine transformative outcomes.³

¹ AusAID (2001) and CIDA (2001a) in Table B1. Neither AusAID nor CIDA identified their evaluations as audits. This terminology was applied by the consultants for the current review.

² ILO (2002) and OCAA (2001) in Table B1 and Table B4.

³ ILO (2002, p. 23) in Table B1.

According to ILO staff, the audit resulted in rich learning on strategies for institutionalising gender equality commitments, and some work units have already begun to implement recommendations. At the agency level, the ILO has made a commitment to continue the process in other work units.

CIDA undertook a desk review of 26 activities in the Indonesia program, which aimed to provide baseline data of the integration of gender equality considerations against which future progress could be compared. Tools developed for the review included a project file questionnaire for reviewing documentation, and questionnaires for CIDA field representatives and executing agencies. The questionnaires included a comprehensive set of gender mainstreaming questions for the activity cycle, focusing on: gender analysis requirements, gender equality integration into activity design, the development of a gender equality strategy, budget allocations, and integration of gender equality commitments into contractual documents. Most of the questions could readily be adapted for use as gender mainstreaming indicators.

AusAID's gender and development review used a "rapid panel assessment" tool to assess 20 randomly selected activities that were being implemented. Each activity was assessed using 14 gender mainstreaming indicators focused on design, professional management of the activity by the contractor and partner agency, the likelihood of the activity achieving its objectives and gender-related benefits, and the likely sustainability of gender-related results. Panels of AusAID staff and external consultants assessed the attributes of each activity based on summaries of activity documents, in addition to interviews with team leaders and AusAID staff. The indicators allowed staff to systematically assess gender mainstreaming across activities and sectors. The method also involved staff in a participatory assessment of gender mainstreaming processes, which promoted peer learning and ownership of the findings. However, the methodology provided little insight into benefits and impact.

A facilitated participatory self-assessment approach was also used in the **OCAA**'s field gender audit in the Philippines in 2001. Partner organisations, head office staff and field staff used a gender audit tool to guide their assessment of selected activities in a 6-day audit workshop. The audit process also included a further three days of field visits to communities, where qualitative methods were used to assess benefits and impacts, particularly community perceptions of changes in gender relations and the major causes of those changes (which sometimes had little to do with OCAA activities). Partners and staff collectively analysed the implications of their assessments, focusing on: gender analysis requirements through the project cycle, strategies for working with communities to address gender issues and promote equality and empowerment, and implications for strengthening partner capacity in these areas. The gender audit tool included 25 questions focusing on project design features, the application of key gender analysis concepts, monitoring processes, the approach to addressing gender equality and empowerment, project

resources, and how to strengthen the focus on women's empowerment and gender equality in NGO partner work.

Annex E Features of Evaluation Methodology that Facilitate Gender Analysis

Some of the critical features required to ensure that evaluations pay adequate attention to gender issues are presented in the following tables. The AusAID evaluations are used to illustrate overall findings. Table E1 presents the features of those evaluations that were rated as having given satisfactory attention to gender issues, and Table E2 presents features of those that have been rated as unsatisfactory or highly unsatisfactory.¹

Addressing gender issues in TORs is undoubtedly necessary as a first step towards a gender-sensitive evaluation, and the employment of local consultants is often cited as being desirable for many different reasons. However, the AusAID examples appear to suggest that TORs have little impact on the level of attention given to gender issues in an evaluation. Similarly, the use of local consultants appears to have little impact on whether gender issues are adequately addressed.

Table E1 **AusAID evaluations that gave satisfactory attention to gender issues**^a

Activity (location, sector)	Gave attention to gender issues in TORs ^b	Included local consultant on the team	Collected quantitative information	Collected qualitative information	Analysed information	Used gender-sensitive indicators
China community development	Brief	No	Yes	Yes	Yes	No
Vietnam NGO	Brief	Yes	Yes	Yes	Yes	Yes
Indonesia WSS	Substantial	Yes	Yes	Yes	Limited	Yes
Thailand rural development	Brief	Yes	Yes	No	Yes	Yes
SE Asia agriculture	Brief	Yes	Yes	Limited	Limited	No
Samoa rural	No TORs	Yes	Limited	Yes	Yes	No

^a Evaluations rated as having given satisfactory attention to gender issues assessed the participation of women in activities, and provided at least one piece of information on how women benefited from activities, or they made at least one statement about the impact of activities on women and/or on gender relations. ^b No TORs means that no terms of reference were included in the evaluation report.

¹ The notes to the tables define the ratings “satisfactory”, “unsatisfactory” and “highly unsatisfactory”. Inevitably, as with any rating system based on qualitative assessments, there were cases that were difficult to rate. Where difficulties arose in whether to rate an evaluation as having given satisfactory or unsatisfactory attention to gender issues, the evaluation was rated as having satisfactory gender analysis.

Table E2 **AusAID evaluations that gave unsatisfactory or highly unsatisfactory attention to gender issues**^a

Activity (location, sector)	Gave attention to gender issues in TORs	Included local consultant on the team	Collected quantitative information	Collected qualitative information	Analysed information	Used gender- sensitive indicators
Fiji health	Substantial	Yes	Limited	Limited	Limited	No
Philippines agriculture	Brief	Yes	Limited	Limited	Limited	No
Africa NGO	Brief	Yes	Limited	Limited	Limited	No
Tibet WSS	Brief	No	Yes	Limited	No	No
Pacific infrastructure	Brief	Yes	Limited	Limited	No	No
Global land titling	No TORs	No	No	No	No	No

^a Evaluations rated as having given unsatisfactory attention to gender issues assessed the participation of women in activities in a limited way. They did not provide information on how women benefited, nor did they include any statements about the impact on women. Evaluations rated as having given highly unsatisfactory attention to gender issues did not assess women's participation, their benefits or the impact on women, although they may have included references to women or gender issues in their reports. ^b No TORs means that no terms of reference were included in the evaluation report.

Annex F Examples of Benefits for Women, Gender Analysis and Links to Development Outcomes

This annex provides examples from evaluation reports that provide detail of the benefits for women or the changes in gender relations, some of the qualifications provided by evaluators about those benefits and, sometimes, assessments of how those benefits arose. Links to overall development outcomes are also made where possible. This annex is not a comprehensive list of benefits identified across the evaluations reviewed.

Examples of gender analysis and mainstreaming

Systematic approaches to gender analysis and mainstreaming in the activities evaluated were rare. To illustrate the extent of this problem, in a review of 21 basic education projects in seven countries, only one had undertaken gender analysis and the recommendations were only partially implemented.¹ Moreover, in cases where analysis had been undertaken and documented in activity design, that analysis was often incomplete. For example, in the ADB's evaluation, activities were selected for evaluation precisely because they had given some attention to gender issues in design or implementation:

All the projects reviewed provided practical benefits to women and men ... However, in general, the project designs were based on inadequate baseline data on gender issues, and as such, there was an overall tendency for the project provisions for gender equity to be vague. Accordingly, the designs lacked recognition of the most relevant gender issues, and [lacked] effectiveness in targeting provisions.²

Sida also selected activities on the basis that some gender mainstreaming had been undertaken in either their design or implementation. However, the evaluators found that:

... rather than a coherent and integrated mainstreaming process, most interventions exhibited the presence of only a few mainstreaming elements, and in general showed only embryonic evidence of working with gender mainstreaming processes.³

AusAID's policy evaluation presents a positive assessment of the attention given to gender mainstreaming, which is not really reflected in the database as a whole.⁴ However, it also identified the lack of gender-sensitive indicators in logframes as a serious constraint

¹ BMZ (2001a, p. 40) in Table B2.

² ADB (2001, p. iii) in Table B1.

³ Sida (2002a, pp. xxii, 3) in Table B1.

⁴ AusAID (2001) in Table B1.

during implementation and monitoring, and noted that most activities collect sex-disaggregated data on only male and female participation in training.

Examples of benefits for women⁵

Usually, if evaluation reports identified benefits that addressed practical needs, evaluators also noted that these benefits were limited to a few areas of activity. Some also mentioned that there were no strategic changes in gender relations, or identified negative impacts alongside the benefits.

For example, BMZ's thematic evaluation of financial and technical co-operation concluded that women were benefiting from activities, but less than men were. The most positive results identified for women included better food, higher income, lower workload and improved health. However, little improvements were identified in social, economic and judicial equality, women's access to resources remained less than men's, and no changes in traditional roles were identified.⁶ Danida's evaluation of the Rakai district development program in Uganda identified positive results for girls in education and for women from a targeted credit scheme, but little priority was given to women's interests in other program areas including health and infrastructure activities.⁷

There are some examples in the evaluation database of women benefiting from agricultural activities, but the findings were often mixed. For example, AusAID's evaluation of three rural projects in Samoa found that practical needs had been met through increased family income, and that some women had benefited directly from training and employment as extension agents. However, women had less access to needed agricultural information than men did, and the introduction of village fish reserves restricted women's and older men's access to shellfish collection sources.⁸ The World Bank found that scant attention had been paid to gender issues in agricultural activities, such as disparity in access to land, and different roles and responsibilities for agricultural production:

Although 42 of the 180 interventions were in the agricultural sector, only 14 projects included components or activities to benefit women. The components were ad hoc and appended, but resulted in small positive results for women.⁹

⁵ More examples are provided in the following sections.

⁶ BMZ (2000) in Table B1.

⁷ Danida, (2000c, p. 22) in Table B2.

⁸ AusAID (2000f, pp. xiii, 47) in Table B2.

⁹ World Bank (2002, pp. 14-15) in Table B1.

Belgium's evaluation of an integrated community development project in Uganda found that poverty reduction initiatives addressed practical needs such as health care and water. The project did not invest in women's strategic needs, but sometimes women's ability to make decisions about external resources was increased, even though this had not changed social and cultural factors that prevent women from exercising equal decision making.¹⁰

Participatory design was a key factor documented in the success of a number of activities in achieving benefits for women, which then had other spin-off effects for development in general. This occurred across sectors and was a key factor in the success of a community development and watershed conservation project in Nepal.¹¹ While income increases were marginal, evaluators identified "self-development skills" (confidence and skills gained) and "group development capability" (including decision making and organisational capacity) as key qualitative impacts.

Examples of a causal relationship between the attention given to gender issues and the benefits for women

Understanding the constraints that had prevented women from accessing benefits in the past, and using that analysis to target women more effectively, emerged as two important design features in the WFP's review of its commitments to women. For example, increasing the number of female food monitors increased the effectiveness of reaching female target groups and ensured that they were aware of their entitlements. In some programs, women from female-headed and polygamous households received food aid directly through a ration card system that recognised their status as individuals.¹²

Specifically targeting women and understanding the constraints to their participation also emerged as key design features in JICA's policy reviews. Of five activities included in the JICA 2000 evaluation, two clearly benefited women and the others either did not benefit women, or men benefited more than women did. The two that clearly benefited women either undertook gender analysis or incorporated targeting strategies. Access to training and other resources were also key factors in whether women benefited. Of eight activities included in the JICA 2001 evaluation, one included gender analysis in its design. This activity also involved a female social analyst and included a series of seminars for forestry extension workers that integrated social and gender issues. This activity and another had

¹⁰ Belgium (2001c) in Table B1.

¹¹ JICA (2001, p. 54) in Table B1.

¹² WFP (2002a, para. 11-12) in Table B1.

specific strategies for targeting women. Evaluators found that only these two of the eight activities reviewed showed benefits for women.¹³

Of the three activities included in AusAID's review of rural development in Thailand, only the Ubon-Ratchathani Land Reform Area Development Project (ULRAP) had any significant gender analysis or a gender strategy. The small gains for women in the other two activities were found to be largely unsustainable because of the lack of gender strategies and poor monitoring. In contrast, evaluators found that ULRAP had a substantial impact on women's participation, status and public roles. This impact was achieved by a strategy of starting with women's income-generation and savings groups to build confidence, and moving promptly to encourage women's participation in mainstream development or local government activities. Although the increases in family income were small, the leaders of these groups moved in a short time to elected positions in local government.

The impact of effective gender strategies was that the level of local development capacity was realised more quickly and the balance of local decision making was improved, compared to the likely rate of progress with male-centred strategies ... While women's participation in public life remains low relative to men, it has increased markedly in recent years and local people attribute this to opportunity, training and encouragement provided through this project ... Village leadership skills have increased for both women and men ... Projects that corralled women into traditional craft groups seemed to make less progress and have less active and effective local government.¹⁴

The evaluation concluded that giving serious and effective attention to gender issues in activity design and implementation yielded faster progress in community development generally and more balanced local government decisions. Other key features of the strategy to address gender issues were:

- an emphasis on genuine participation by local men and women, using NGO input
- close alignment with the policies and priorities of the partner government agency, and
- effective monitoring of the achievement of gender and development objectives and impacts.

In 2002 CIDA requested an evaluation of the impact of a gender strategy introduced into a co-operative development project in Guatemala in 1998. Elements of the strategy included equipping women with skills, organising women, sensitising both women and men to gender issues, and forming a co-operative with women. Extensive data were provided on the impact of the strategy, including significant increases in the number of women that were active co-operative members, in executive positions, participating in training, and

¹³ JICA (2000, pp. 277-286) and JICA (2001, pp. 48-71) in Table B1.

¹⁴ AusAID (1999b, pp. xiii, 30-34, 65) in Table B2.

accessing credit. The report also asserted that there were qualitative changes, including an improvement in the quality of women's participation in training activities, women's opinions being taken into account in co-operatives, and increased recognition of the role of women in the family. However, the evaluation noted that, while passive resistance to women's involvement in decision making had been overcome, their benefits were limited to income generation, and they had not set up their own organisations. The evaluation noted the importance of sensitising men (husbands and male leaders) to ensure that women can participate equally. One difficulty faced by the implementing NGO was the great personal effort (in terms of time and unpaid labour) required of women participating, which in some cases led to women becoming fatigued and disillusioned, causing them to drop out of the project. Unfortunately, the evaluation did not investigate whether women retained control over any income or assets earned from the use of credit.¹⁵

Example of a gender strategy detrimental to addressing gender issues

One example was found in the review database of a gender strategy that had a detrimental impact on an activity's capacity to address gender issues properly. CIDA's evaluation of the Regional AIDS Training Network in southern Africa found that gender analysis undertaken early in the activity's implementation resulted in an unrealistic gender strategy and recommendations that were beyond the scope of the activity. Failure to have an appropriate gender strategy meant that some areas where gender issues should have been addressed were overlooked.¹⁶ The review recommends a mainstreaming approach (rather than a separate planning process to address gender issues) to ensure that gender strategies are relevant and specific to the scope of the activity.

Findings on girls' education and development outcomes

The World Bank's thematic evaluation found that overall reductions in gender disparities in the education sector were relatively small. Girls' enrolments increased by 2% and 4% respectively in Bangladesh and Gambia, but did not increase in other countries where efforts were made to address the issue. In Vietnam and Sri Lanka, girls' enrolments increased despite largely gender-blind activity design. However, a study on girls' education by USAID showed significant increases in girls' enrolments (between 50% and 89%) in Guinea, Malawi and Pakistan, and a decrease in enrolments in Egypt. Boys' enrolments also increased over the same period, but not by the same amounts. A study by CIDA in Ghana found that boys' and girls' enrolments increased at approximately the

¹⁵ CIDA (2002a, 2002b) in Table B1.

¹⁶ CIDA (2000d) in Table B2.

same rate due to institutional problems at the national level and project management problems. However, the evaluators concluded that activities targeted at increasing enrolments were responsible for the small gains that have been made, particularly in increased community and parental commitment to girls' education.¹⁷

The USAID evaluation concluded that targeting girls' schooling improves overall development outcomes in education:

When systems are geared up to solve the problems that keep girls out of school or prevent them from learning in school, the solutions have broad applicability and relevance to both sexes. Boys, especially those belonging to vulnerable groups or who live in remote rural areas, face many of the same problems as girls meeting their basic learning needs: lack of nearby schools, poor school quality, and lack of parental resources, support or participation in a child's education.¹⁸

Girls' education initiatives benefited all children if they:

- strengthened investment in and the capacity of primary education institutions by, for example, training teachers and supplying instructional materials
- increased the supply of primary schools, including the supply of school places and different options for schooling
- staffed schools with female teachers, who were preferred by parents for teaching young boys in some places, and
- reduced schooling costs.

Four evaluation reports also presented evidence that the education of women through non-formal literacy, empowerment or credit programs had a positive impact on education enrolments of both boys and girls – three cases in Nepal¹⁹ and one in India.²⁰ An evaluation of a credit initiative in Kenya reported that women were able to pay school fees for children, which they were not able to do before.²¹

Three evaluations considered the quality of education in relation to initiatives to increase girls' enrolment.²² Each indicated that efforts to improve quality had limited impact and identified this as a serious sustainability issue for the future, with the potential to undermine the gains that had been made in increasing both girls' and boys' enrolments.

¹⁷ World Bank (2002, pp. 8-10), USAID (1999d, p. 39) and CIDA (1999a) in Table B1.

¹⁸ USAID (1999d, pp. 6-7) in Table B1.

¹⁹ USAID (1999d), CIDA (1999d) and USAID (2001a) in Table B1.

²⁰ CIDA (2000g) in Table B2.

²¹ Belgium (2001b) in Table B1.

²² USAID (1999d), CIDA (1999a) and World Bank (2002) in Table B1.

Examples of the benefits of gender-sensitive programming in post-conflict situations

Four evaluations included in this review concluded that gender-sensitive programming led to improved outcomes at the community level in post-conflict situations. These were a review by the WFP of its commitments to women, an extensive USAID study of the impact of international assistance to women's organisations in post-conflict societies, an evaluation of USAID small grant assistance to women's associations in Rwanda, and an evaluation of a CIDA-funded NGO program in Peru.²³

The USAID study noted that the international community channelled assistance through women's organisations because they were more effective than mixed or male-dominated organisations at reaching out to women faced with extreme poverty and deprivation, and because they were successful in empowering women by raising gender awareness, facilitating political participation and putting gender issues on the national agenda of governments.²⁴

In Rwanda, the USAID-funded Women in Transition Initiative provided funds to local rural women's associations for agricultural, livestock and micro-enterprise activities. The evaluation found that this contributed to the participation of women in new political structures because of the opportunity these activities provided for women to take part in local decision making, including implementation of local development activities. In addition to directly addressing poverty (the women's associations were successful at targeting the most vulnerable, including female-headed households), the evaluation concluded that these activities had a direct effect on reducing social tensions and promoting unity:

Grants to women's associations in Rwanda are the foundation for rebuilding local communities and social trust, and are making a contribution to peaceful reconstruction.²⁵

The Inter Pares program in Peru focused on helping women to identify how they were affected by political violence, and to communicate this to government. This had a positive influence on government policy and programming. The evaluation concluded that the provision of assistance through women's organisations and federations resulted in both a reactivation of agricultural production and a "reweaving of the social fabric that was torn during the violence".²⁶ Inter Pares worked with 12 NGOs with the aim of promoting the economic and democratic reconstruction of communities affected by violence. Institutions

²³ WFP (2002), USAID (2001b), USAID (1999c) and CIDA (2000e) in Table B1 and Table B2.

²⁴ USAID (2001b) in Table B1.

²⁵ USAID (1999c, pp. 3, 5) in Table B1.

²⁶ CIDA (2000e, p. 43) in Table B2.

were strengthened by providing technical assistance in agriculture, and training in gender, human rights, leadership and political participation, in addition to training that specifically targeted women's organisations and federations to empower women. Results included improved gender focus and programming with partner NGOs, women's involvement in decision making and leadership at the community and local political levels, the election of a number of women as municipal councillors, women's organisations better able to negotiate with local government and other institutions, which became more responsive to women's demands, and evidence of men supporting these changes in women's roles.

The WFP's findings highlighted the importance of male involvement and support, and the danger of marginalisation:

A positive impact has been noted where women of different ethnic groups have been actively supported to participate jointly in food management and distribution committees. In contrast to the elders' committees dominated by men and which operate along ethnic divides, women in this case generally tend to work and cooperate with one another, thus contributing to overcoming the divides that are the cause of much civil strife and conflict. This also points to the possibly positive impact of supporting women-only groups when care is taken to ensure that they are not socially and economically marginalized. Though available documentation also reveals that such groups will tend to be more sustainable and therefore more likely to attain the anticipated impact where male support is actively sought and secured.²⁷

Examples of how women's participation helped to achieve activity objectives and more effective governance processes

In addition to the evidence from post-conflict situations (above), other evaluations made some link between the participation of women in activities, which enhances their leadership capacity, and either the achievement of activity objectives or more effective governance processes. Having participatory planning and monitoring processes, and ensuring that women have access to training appear to have been key factors. However, many of these activities were either implemented by NGOs or of a relatively small scale.

An AusAID evaluation of NGO work in Vietnam highlighted the following example:

The Dengue Surveillance and Control Project was outstanding in its use of gender sensitive participatory process in implementation, monitoring & evaluation despite women's lack of involvement in project design. Utilization of women collaborators from the Vietnam Women's Union has proved a very effective mechanism to convey information to communities about the dengue mosquito and to promote household stagnant water cleanup campaigns around housing areas ... Women have achieved considerable status as a result of

²⁷ WFP (2002, para. 151) in Table B1.

training and community activities and appear to be very well respected as leaders by both men and women within their communities.²⁸

The evaluation of NZAID's gender and development project in Kiribati reported that women had gained confidence to speak up about domestic violence and to report incidents. This was achieved through training of the Kiribati Police Force and a program of community awareness on this issue.²⁹

Sida's support for the empowerment of women through Panchayati Raj in India focused on training and preparing women to become effective partners in decision making in local government structures. The evaluation reported that women learned about panchayat administration, gained respect in the community, took initiatives to solve community problems, acquired confidence to share their experiences and learn from their peers, were able to meet with higher authorities to put their problems before them, acted as agents of change, identified problems in panchayat and state government administration, became sensitised to social problems, demonstrated an eagerness to work for social justice for the marginalised and underprivileged, and started organising women's groups for income generation. Women's participation (through meetings, suggestions and involvement in decision making) resulted in increased efficiency in the running of local government programs, the completion of programs, and the extension of social justice and entitlements of these programs to the needy.³⁰

CIDA's Social and Women's Initiative Fund in Vietnam funded 31 small activities over five years, with the aim of empowering Vietnamese women through a series of training programs and other institutional strengthening activities. Results included an increased number of Vietnamese women in economic and political decision making at local and national levels, and an increase in the capacity of some public and private institutions to plan and deliver programs that benefit women.³¹

In India, local NGOs organised women into groups around income-generation activities. As a result, women participated in project planning and implementation and began to take a greater place in civil society, collectively petitioning local government for the provision of essential services such as water, sewerage, electricity and garbage removal in urban slums.³²

²⁸ AusAID (2000g, p. 35) in Table B2.

²⁹ NZAID (2002a) in Table B1.

³⁰ Sida (2000a, p. 18) in Table B1.

³¹ CIDA (2000b) in Table B1.

³² CIDA (2000g) in Table B2.

In the Caribbean, evaluators found that small activities focused on the political participation of women (supported through CIDA's Gender Equity Fund) increased popular participation in politics in general, and were expected to increase accountability to electors.³³

Other evaluations of activities in the governance sector either gave inadequate attention to gender issues to draw conclusions about the overall impact of activities on development (the lack of time on field visits was noted by evaluators as a constraint in some cases) or the evaluators indicated that poor gender analysis in activity design and implementation, the lack of sex-disaggregated information and weak monitoring processes made it impossible for them to evaluate gender impacts.³⁴

AusAID's evaluation of the Qinghai Community Development Project found that women were poorly served by cash for work, credit and training inputs, with men dominating benefits received in all areas. Despite this overall finding, some useful lessons were learned about the effectiveness of targeting women and the contribution they can make to achieving development objectives. The evaluation found that employing female credit extension agents improved the targeting of both the poor in general and women, and resulted in better loan recovery performance. Social and group cohesiveness, where the poor helped one another to repay loans and get through times of hardship, also occurred where women were members of the group that controlled their own loans. The report acknowledged that loan repayments placed exceptional demands on women in terms of increased labour time (especially if they had children at school), but that this did not have a serious impact on loan repayments where groups were cohesive and recognised joint responsibility. The conclusion was that working with women, particularly in minority national areas, may reduce the vulnerability of the poor by reinforcing patterns of cohesiveness and mutual support.³⁵

Examples of women's control over income and links to poverty reduction

There were numerous evaluations of activities in the review database that included a credit or small business component, either targeted at women or at communities in general. Most of these did not provide sex-disaggregated data on critical issues such as who benefited from and controlled loans, assets or income from the use of loans.³⁶ Without such

³³ CIDA (2000a, p. 13) in Table B1.

³⁴ EC (2001a), EC (2001b), Sida (2001b) and Sida (2000c).

³⁵ AusAID (2000d, pp. xiii, 23-49) in Table B1.

³⁶ See CIDA (1999b, 2002a), USAID (1999b, 1999c) and AusAID (1999a, 1999b, 2000c) in Table B1 and Table B2.

information, it is impossible to draw links between the targeting of women for credit provision, the empowerment of women, and poverty reduction. To be fair, in some of these cases the evaluators did note that no detailed sex-disaggregated information was available on the uses of credit. The following evidence is from evaluations that provided more detailed information.

The USAID-funded Women’s Empowerment Program in Nepal aimed to empower women in three strategic areas – literacy, legal rights and economic participation. Activities that targeted women (literacy training, awareness-raising and credit) resulted in increased decision making by women on a range of household matters, which had both direct and indirect impacts on poverty. The direct impacts were more household resources being spent on family wellbeing – on food, clothing, education and health care. Indirect impacts related to women taking individual or collective action with local authorities to improve various aspects of their lives (for example, domestic violence, alcohol abuse, property settlement after divorce, polygamy, and community perceptions of women’s work and behaviour).³⁷

JICA’s evaluation of five activities in Sri Lanka found that women involved in producing handicrafts spent their money on domestic necessities.³⁸

AusAID’s evaluation of the Qinghai Community Development Project in China found that men dominated credit recipients and that, even where women did take out loans, this was generally in name only. When women did manage to control income (cash for work), there was evidence that they spent it on basic family and household needs such as school fees, medicines and food, whereas men tended to regard additional cash as a “windfall” and spent it on gambling and alcohol.³⁹

Most credit activities included in the World Bank evaluation were assessed as gender blind in design. In one case, a gender-blind design benefited both women and men, but this was principally due to the involvement of NGOs, which ensured women’s access. However, both the World Bank and the ADB evaluations found mixed results on female control over assets or income from credit. At one end of the spectrum, women could borrow only through their husbands and had little or no control over spending. At the other end, there was a mix of male, female and joint male–female control over expenditure of income earned from credit that targeted women, with female control in a minority.⁴⁰

³⁷ USAID (2001a) in Table B1.

³⁸ JICA (2000, p. 283) in Table B1.

³⁹ AusAID (2000d, pp. ix, 43) in Table B2.

⁴⁰ World Bank (2002, pp. 18-19) and ADB (2001, pp. 6-9, 46-100) in Table B1.

Examples of water supply activities

Despite the considerable amount of work that has gone into raising awareness and developing tools for ensuring that gender considerations are taken into account in water supply and sanitation projects, few evaluations documented that women participated in planning, implementing or managing activities in this sector. The World Bank evaluation found that most water supply activities were gender blind and that sustainability was a major concern. However, where women had been involved (in the Philippines), the World Bank concluded that their involvement improved the use and sustainability of water facilities.⁴¹

No other such lessons emerged from evaluations of other water supply activities. In one case, the evaluation itself had very poor quality gender analysis, which assessed participation in a limited way and did not assess benefits or impact.⁴² In another case, evaluators concluded that there was no gender strategy and that the results would have been better if strategies to involve women had been put in place, although they did find that the whole community benefited from the water supply construction.⁴³ In another case, the evaluation reported positively about the attention given to gender issues in water supply activities, but no analysis was made of links to poverty reduction or other development objectives.⁴⁴ BMZ's evaluation noted that the drinking water activities evaluated seldom saw the need for gender-specific information.⁴⁵ One water supply and sanitation activity was included in the ADB evaluation. Although this activity did have gender provisions in the design, these were poorly implemented, with water user committees being male-dominated and female sanitation promotion volunteers inactive.⁴⁶

⁴¹ World Bank (2002, p. 16) in Table B1.

⁴² AusAID (2002b) in Table B2.

⁴³ AusAID (1999a) in Table B2.

⁴⁴ CIDA (2001c, p. 45) in Table B2.

⁴⁵ BMZ (2000) in Table B1.

⁴⁶ ADB (2001, pp. 8, 74-79) in Table B1.

References

- Development Co-operation Directorate, Development Assistance Committee, Working Party on Aid Evaluation (2001), “Options for Future Work on Gender Equality by the DAC Working Party on Aid Evaluation: A Discussion Paper (Note by the Secretariat)”, Submitted for discussion to the 35th meeting of the Working Party on Aid Evaluation on 14-15 November 2001.
- Hunt, Juliet (2000), *Institutionalising Gender Equality Commitments in Development Organisations and Programs*, Winston Churchill Memorial Trust of Australia, Canberra.
- Keller, Bonnie (2000), “Report on the DAC Workshop on Evaluating Gender Equality and Women’s Empowerment, 25-26 November 1999, Stockholm”, Prepared for the DAC Working Party on Aid Evaluation.
- OECD (Organisation for Economic Co-operation and Development) (1994), *Assessment of DAC Members’ WID Policies and Programs*, OECD, Paris.
- OECD (1998), *DAC Sourcebook on Concepts and Approaches Linked to Gender Equality*, OECD, Paris.
- OECD (1999), *DAC Guidelines for Gender Equality and Women’s Empowerment in Development Co-operation*, Development Co-operation Guideline Series, OECD, Paris.
- OECD (2001), *The DAC Guidelines: Poverty Reduction*, OECD, Paris.
- OECD (2002), *Glossary of Key Terms in Evaluation and Results Based Management*, OECD, Paris.
- Woodford-Berger, Prudence (1999), “Evaluating Gender Equality and Women’s Empowerment: A DAC Review of Agency Experiences 1993-1998, Final Draft Report”, Prepared for the DAC Working Party on Aid Evaluation, Stockholm.
- World Bank (2001), *Integrating Gender in World Bank Assistance*, Report No. 23035, Operations Evaluation Department, World Bank, Washington, 25 October.
- World Bank (2002), *The Gender Dimension of Bank Assistance: An Evaluation of Results*, Report No. 23119, Operations Evaluation Department, World Bank, Washington, 17 January.