Promoting Human Rights and Gender Sensitive Evaluations: Key Ideas for Evaluating Gender Equality Results

Prepared By

Julia Espinosa (**)

Complutense University, Madrid

Introduction

Evaluation is regarded as a central tool for learning, improving and accountability in relation to international development policies and programmes. It has been so recognized on an international level where evaluation started to gain importance in the late sixties and continues to do so today. In fact, now evaluation is considered as a key element in the promotion of a greater effectiveness and higher quality of international development and cooperation.

Gender and human rights issues were first incorporated in international development evaluation relatively recently. Indeed, the vast majority of evaluative practices related to gender equality dates back to twenty years ago. The inclusion of gender issues in evaluation has been related to the international focus on gender equality as a human right and a pre-requisite for development since the Beijing Conference (1995). The incorporation of human rights as a central concern of the evaluation dates back to the last decade and it is linked to the acceptance by the United Nations of the Human Rights Based Approach (HRBA) strategy whose goal is to implement human rights in programming (Espinosa, 2013, Ligero et al., 2013 and UNEG, 2011).

Broadly speaking, the incorporation of gender and human rights issues in evaluation is essential for generating learning and knowledge in respect to how to promote these subjects in a more adequate way in international development programmes. Additionally, it is a fundamental exercise for reporting what advances have been made in this direction and what is necessary for promoting human development even more. So, an evaluation sensitive to human rights and gender equality is a central element for effectively promoting HRBA in programming and gender mainstreaming strategy as well as for improving the human development results.

Although there are some key articles on feminist evaluation and human rights evaluation as well as some specific guidelines, the majority of the contributions comes from general evaluation methodologies, gender-sensitive planning frameworks and meta-evaluation studies. In addition, while there is an increasing number of evaluations focused on gender equality and human rights, most of the evaluative practice has not yet included these issues systematically.

This communication presents some key ideas about how to carry out human rights and gender sensitive evaluations and how to measure the results of gender mainstreaming strategy. For that purpose, the communication is based on the guidelines of UNEG (2011) and UN Women (2011). Additionally, it also receives information from the study *How to carry out human rights and gender sensitive evaluation* (Ligero et al., 2013)[1], the doctorate thesis *Gender Equality in the Evaluation of the Development Aid: the cases of British, Swedish and Spanish aid* (Espinosa, 2011) and two papers based on it (Espinosa, 2013 and 2012)[2]. Using these references, this communication tries to answer the following questions:

---

[1] This study, funded by the Minister of Foreign Affairs and Cooperation of the Spanish Government, is based on a thorough analysis of the existing bibliography of material on gender, human rights and development evaluation as well as on opinions of experts –using interviews and workshops-. The complete reference to this study is Ligero, J.A., Espinosa, J., Mormeneo, C. and Bustelo, M. (2013) *Cómo hacer una evaluación sensible al género y al enfoque basado en derechos humanos para el desarrollo*, Madrid, Minister of Foreign Affairs and International Cooperation.

[2] This PhD research is a meta-evaluation based on the analysis of the strategic evaluations of these donors and interviews of the staff of their evaluation units. It is available in Spanish at: http://eprints.ucm.es/13206/

- How can we make evaluations more human rights and gender sensitive**?**

- How can we define human rights and gender sensitive criteria, evaluation questions and indicators?

- What kinds of evaluative methodologies and techniques help to measure the results of gender mainstreaming strategy?

The main goal of this paper is to provide key ideas for carrying out human rights and gender sensitive evaluations in order to improve gender mainstreaming in programming.

How can we make evaluations more human rights and gender sensitive?

A human rights and gender sensitive evaluation is key for promoting a real inclusion of and attention to human rights and gender equality in all the development policies and programmes. Indeed, "those UN interventions that do not consider these principles risk reinforcing patterns of discrimination and exclusion or leaving them unchanged" (UNEG, 2011: 11). But, what does a gender equality and human rights sensitive evaluation mean? The development of evaluations with this sensibility implies considering human rights and gender matters as key elements throughout the evaluative process. In practice this means carrying out two exercises of analysis that complement each other and share the principles of inclusion, participation, fair power relations and the use of mixed evaluation methods. In both cases, it means the identification and analysis of "the inequalities, discriminatory practices and unjust power relations that are central to development problems" (UNEG, 2011: 14).

Specifically, incorporating human rights and gender issues in evaluation implies going through all the evaluation steps and making some important decisions related to these matters. As a starting point, the evaluation team has to:

- "Recognize the gendered nature of development and the importance of gender equality to economic and social development.

- Recognize that poor and marginalized groups of people are gendered, and women and men are differentiated by race, ethnicity, age, disability, class and caste

- Recognize that power relations exist within the home/family, and that these relationships intersect with power relationships in society, policy, and economy

- Consider these dimensions in the evaluation process – approach, methods and use

- Collect and analyze sex, age, ethnic, regional disaggregated data" (Sanz et al., 2012).

Regarding the evaluative process, it means that the Terms of Reference (ToR) requires experts on gender and human rights to be on the evaluation team and the incorporation of gender equality and HR as important issues. It also involves rethinking what should be analyzed, what evaluative methodology should be adopted as well as what type of report should be made and how to disseminate and use it (Espinosa, 2013; UNEG, 2011; UN WOMEN, 2011; DFID, 2008 and CIDA, 2001).

In the international development evaluation, the practice is usually based on criteria and evaluation questions as well as the measuring of results through a set of indicators. Carrying out a human rights and gender sensitive evaluation implies rethinking the criteria, evaluation questions and indicators taking into account the Theory of Change of the programme, the domains of change in gender power and the gender dimensions as well as the specific interest of the stakeholders. Moreover, this type of exercise also means adapting the methodological and technical proposals to the specifics of gender; ensuring the dissemination of the results to both women and men; and promoting the use of reports in favor of gender equality (Espinosa, 2013; UNEG, 2011, UN WOMEN, 2011, DFID, 2008, De Waal, 2006, and FIDA, 1995).

In relation to the evaluation of the results of gender mainstreaming strategy, it is necessary to identify the Theory of Change as well as the areas in which this strategy wants to promote change as a previous step to the formulation of criteria, evaluation questions and indicators. In the same way, the identification of some methodologies and techniques that help to measure these generated changes or results is fundamental. In the following pages, we suggest ideas about how to define criteria, evaluation questions and indicators as well as about relevant methodologies and techniques for evaluating gender mainstreaming strategies and improving programming.

How can we define human rights and gender sensitive criteria, evaluation questions and indicators?

A gender mainstreaming strategy, as defined by the United Nations, seeks to make "women's as well as men's concerns and experiences an integral dimension of the design, implementation, monitoring and evaluation of policies and programmes in all political, economic and societal spheres, so that women and men benefit equally and inequality is not perpetuated" (UN, 1997). Accordingly, gender mainstreaming is not an end in itself, but its objective is to promote gender equality including the interests of women and men throughout the whole policy cycle.

The results of gender mainstreaming strategy can be evaluated considering different levels of action: macro, meso and micro. At the macro level, the analysis centers on how the gender questions are included in the political and economic context, policy, budgeting, strategy, structures and systems. At the meso level, the evaluation is focused on a review, in terms of gender, of the institutional ability, human and financial resources and system management. And, at the micro level, the analysis is of how the gender content is included in the implementation of the development programme and the personal and interpersonal experience caused by it (De Waal, 2006).

However, in order to identify how much gender equality results are generated at these different levels of action, it is important to first ask what the Theory of Change of the programme is to define criteria, evaluation questions and indicators. "Theory of Change is the belief about how change occurs that is embedded in the intervention design and its logical framework" (UNEG, 2011: 29). The reconstruction of the Theory of Change allows for the identification of how the programme works and what results are to be obtained by paying attention to outcomes and outputs as well as processes and structural elements. In this way, not only do we evaluate "if the objectives have been reached but we can also identify the causal mechanisms that have let to them" (Ligero, 2011:4).

To reconstruct the Theory of Change, the starting point used is the written documents of the programme. However, at the time of the evaluation of the programme, there might exist quite a distance between what was programmed and what was really implemented (Weiss, 1998, in Ligero et al., 2013). So, the incorporation of the points of view of the different stakeholders might favour the development of a logical model that adapts better to the reality of the intervention. The final result can be represented in a table like the one that follows. It helps us to learn about the logic of the programme. Specifically, we can analyze if it works, why it works or not and direct the conclusions and recommendations of the evaluation to improving future programming.

Box. Theory of Change
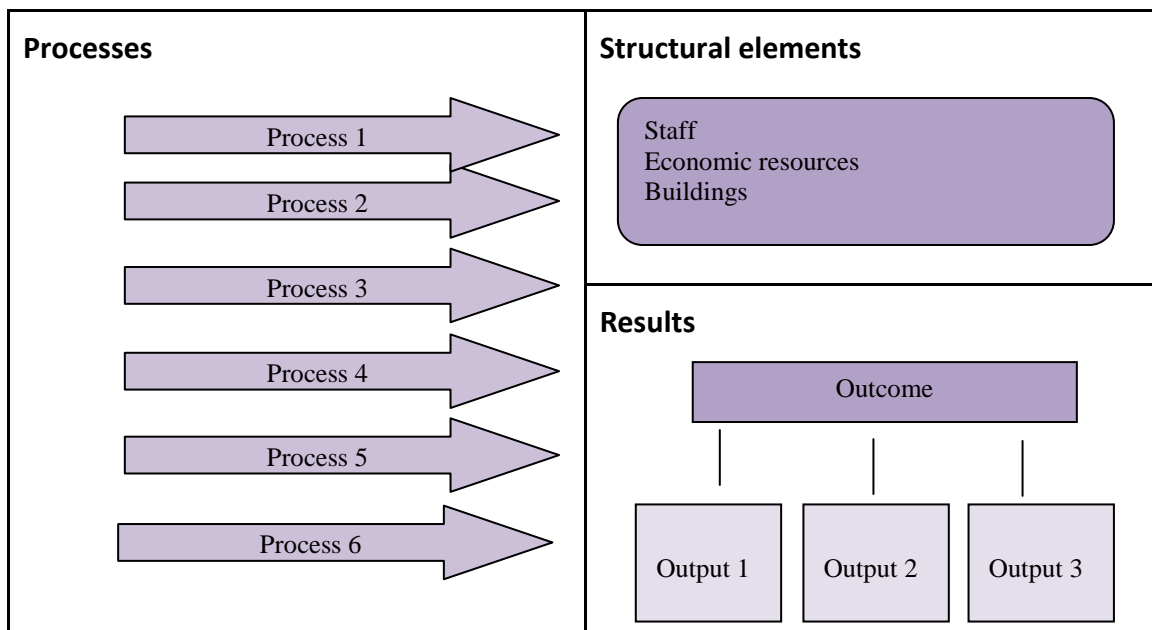
Source: Ligero et al. (2013).

In addition, to evaluate gender equality results, it is central to analyze how the domains of change in gender power and the gender dimensions are taken into account and how they are impacted by the programme evaluated. On the one hand, the domains of change in gender power describe where gender power structures operate. On the other, the gender dimensions help us to visualize some gender inequalities that often go unnoticed (Espinosa, 2011). The analysis of the domains of change and gender dimensions constitutes a fundamental starting point for defining the criteria, evaluation questions and indicators. The attention given to certain domains of change and gender dimensions will depend on the sector and specific context being studied. In any case, this analysis helps us to know what domains and dimensions are taken into account and impacted by the programme and if it is necessary to include new lines of action regarding them.

As to the domains of change, as shown in the following box, there are four different ones and they are interconnected (Rao and Kelleher, 2005). The evaluation of the results of gender mainstreaming strategy must formulate specific questions relative to the changes promoted: in women's and men's individual consciousness (knowledge, skills, political consciousness, commitment); in women's objective condition (rights and resources, access to health services and safety, opportunities for a voice); in informal norms, such as inequitable ideologies, and cultural and religious practices; and in formal institutions, such

as laws and policies. The more domains of change promoted by the programme, the more transformative it is.

Box. Domains of change in gender power
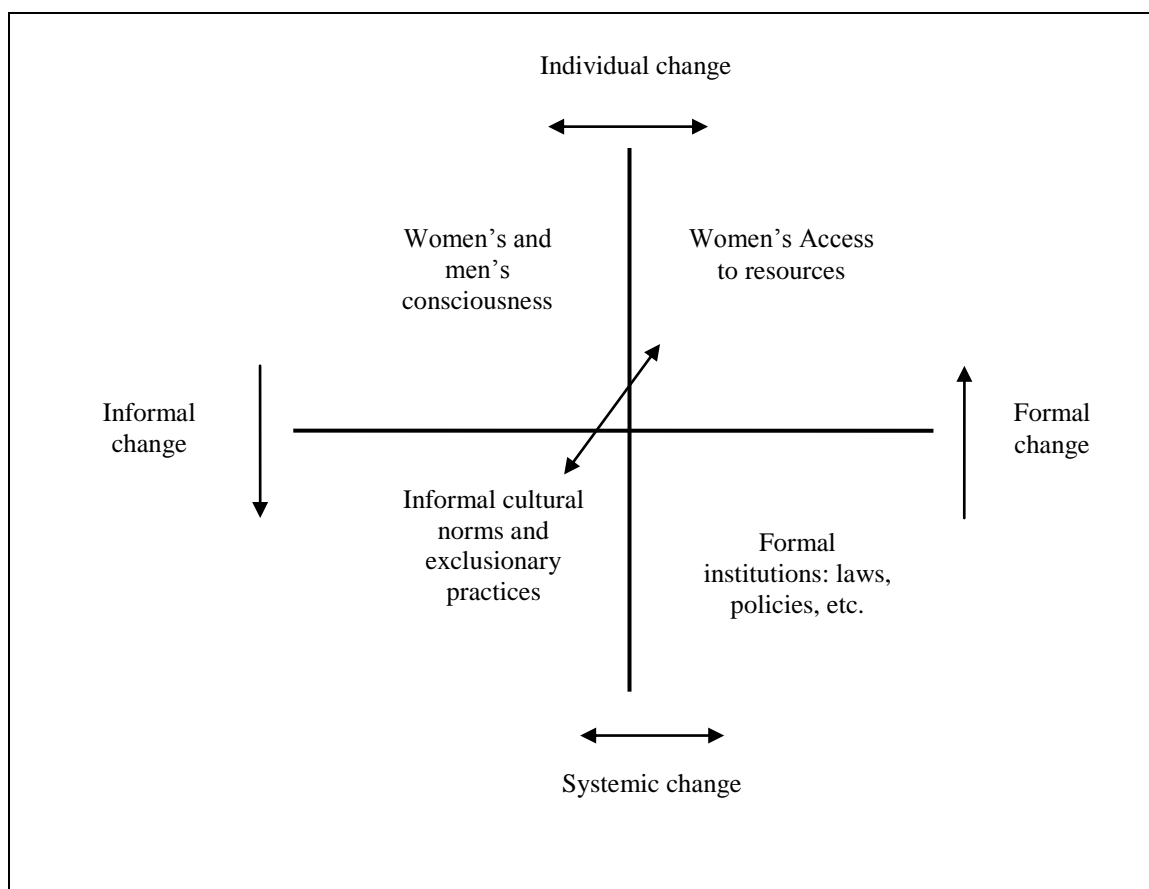
In relation to gender dimensions, six different ones have been identified: sexual division of labor and different gender roles, participation of women and men in private and public spheres, the control of the use of women's bodies, practical and strategic gender needs, different use of time by women and men, and the unequal access to and control over resources, benefits and services (Hunt and Brouwers, 2003; UNDP, 2001; DFID, 1999 and Miller and Razavi, 1998, in Espinosa, 2011). Regarding the evaluation of the gender equality results, the following table includes some key evaluation questions relative to each dimension.

Table. Gender dimensions and evaluation questions

Gender dimensions	Evaluation questions
Sexual division of labor and different gender roles	How much does the programme take into account the three roles (productive, reproductive and community)? Does it promote a well-balanced distribution of tasks between the sexes?

Participation of women and men in private and public spheres.	Is there a balanced participation of women and men in the intervention? Are there activities designed and developed to cause a change in rules of participation of women and men?
The control of the use of women's bodies.	Does the programme promote the idea that women be the ones who make the decisions about their own bodies?
Practical and strategic gender needs, understanding the first one as the needs linked to their daily life and the traditional and unequal customs; and the second ones as the needs that challenge gender division of labor, power and control.	How much does the programme deal with both kinds of needs since they are considered to be important complementary parts of action?
Different use of time by women and men.	Are schedules established to allow for the equivalent participation of women and men? Are actions that promote a balanced use of time between sexes established?
The unequal access to and control over resources, benefits and services.	Are measures taken to increase the access to and control over resources, benefits and services by women? Is a greater balance between sexes in this way promoted?

Source: Espinosa (2011).

According to the sector and the context of the programme, the evaluation of gender equality results must include, in the definition of its criteria, part of all of these questions and translate them into specific indicators. To this effect, it is central to reinterpret the classical criteria to include this content in their definition and extend new criteria if necessary. In this regards, the DAC criteria are not considered the more relevant in this kind of evaluations. These criteria are useful for determining the evaluation scope and unifying their format. However, the use of just the DAC criteria can reduce the utility of the evaluative results because they do not take into consideration the different features of the programmes and their context. In addition, they are not sensitive to the different interest of the stakeholders (Espinosa, 2013 and Ligero, 2011).

In regard to the indicators, the measurement of results in terms of human rights and gender requires a combination of quantitative and qualitative indicators (Espinosa, 2013; Ligero et al., 2013; UNEG, 2011; and UN Women, 2011). Indeed, the advances in human rights and gender relations are not always quantifiable. Son, qualitative indicators, that track changes in behavior and attitudes as well as the perceptions of women and men of their own process of change need to be developed. Moreover, "many authors suggest formulating

indicators in a participatory manner in order to ensure the active involvement of women and men and their ownership of the evaluative process" (CIDA, 1997; Murguialday et al., 2008; World Bank, 2005 in Espinosa, 2013: 178).

However, in practice of international cooperation, the evaluative indicators are not usually linked, in a direct way, to the assessment of design, implementation and results of the programme under evaluation. On the contrary, it is about long-standing indicators that "take on a life of their own (…) becoming an end in themselves rather than a means" (Batliwala and Pittman, 2010: 9). In this respect, the first step to be taken is to define indicators according the key evaluation questions of the specific programme. This is the only way to assess to what extent the programme is generating substantive changes in people's daily lives to get useful knowledge to inform future programming.

In addition, emphasis is given to SMART (Specific, Measurable, Achievable, Relevant and Time-bound) indicators but they cannot always measure complex social changes (Batliwala and Pittman, 2010). These kinds of indicators play an important role in result-based management and in the discussion on accountability. However, Roche (1999) proposes combining these indicators with SPICED (Subjective, Participatory, Interpreted and communicable, Cross-checked and compared, Empowering, Diverse and disaggregated) indicators to explain how change occurs and to generate more knowledge to improve programming. Their features are described below.

- Subjective: Informants have a special position or experience that gives them unique insights which may yield a very high return on the investigators time. In this sense, what may be seen by others as 'anecdotal' becomes critical data because of the source's value.

- Participatory: Indicators should be developed together with those best placed to assess them. This means involving a project's ultimate beneficiaries, but it can also mean involving local staff and other stakeholders.

- Interpreted and communicable: Locally defined indicators may not mean much to other stakeholders, so they often need to be explained.

- Cross-checked and compared: The validity of assessment needs to be cross-checked, by comparing different indicators and progress, and by using different informants, methods, and researchers.

- Empowering: The process of setting and assessing indicators should be empowering in itself and allow groups and individuals to reflect critically on their changing situation.

- Diverse and disaggregated: There should be a deliberate effort to seek out different indicators from a range of groups, especially men and women. This information needs to be recorded in such a way that these differences can be assessed over time.

What kinds of evaluative methodologies and techniques help to measure the results of gender mainstreaming strategy?

The development of human rights and gender sensitive evaluations requires a specific consideration of the evaluative methodologies and techniques to be used to measure

changes in this area. The quantitative techniques, such as surveys or the census, seek to achieve an objective measurement, a demonstration of causality and a generalization of the results through the study of the whole group or a statistically significant sample. The qualitative techniques, on the other hand, are used to study cultural and symbolic codification which constitutes our way of looking at and living social reality. These techniques facilitate an in-depth analysis of depth cases, facts and subjects. Some of the qualitative techniques are: interviews, direct observation and focus groups. The combined use of quantitative and qualitative techniques therefore allows for a more complex analysis of the social change analyzed. In this respect, mixed methods have become more and more accepted and it is argued that the combination of quantitative and qualitative methods and techniques contributes to a better understanding of human rights and gender change (Ligero, 2013; Espinosa, 2013; UNEG, 2011; UN Women, 2011; Moser, 2007; and Murphy, 1997).

Furthermore, along with the recognition of mixed methods, there is also recognition of the reflexivity of the evaluative process and of the need to pay attention to the different voices. So, the evaluative methodologies have to promote stakeholder participation and the inclusion of the most vulnerable people and they have to pay attention to social systems and specific contexts. Focus is on people and how they live the change promoted by the programme. In this way, the use of participatory techniques, such as the Most Significant Change (MSC) technique and the Change Assessment and Scoring Tool (CAST), is promoted for measuring gender equality results. The MSC, on the one hand, identifies the principle changes according to the information provided by the different stakeholders and the importance of each of these changes is debated in small groups. On the other hand, the CAST provides a general view of the changes as perceived by the groups involved in the intervention under evaluation (Espinosa, 2011 and 2012).

This methodological decision first responds to the complexity of measuring transformations in gender relations and the limited data disaggregated by sex. It is also linked to a desire to promote greater involvement of people in the decision making and a greater ownership of evaluation results.

Regarding the application of information-gathering techniques, this should include women and men in the sample or study population; address the potential obstacles to women's participation; and take into account the different time availability of women and men and their diverse needs and interests. In addition, the evaluation team should have skills in working with men, women and diverse groups as well as the skills to enhance participation and ownership of the evaluative practice and not reinforce the traditional leadership and the concentration of power (González and Murguialday, 2004; Rodriguez et al., 1999; World Bank, 2005 in Espinosa, 2013).

In any case, there is not recipe for designing the evaluation methodology. The evaluation team has to be innovative and creative and take into consideration specific sector and context features. The methodological design has to be based on feminist and gender theory and be defined by experts but in a collaborative way with stakeholders. The design must generate useful information in order to assess the programme and to contribute to future programming.

Bibliographical references

BATLIWALA, S. (2011) *Strengthening Monitoring and Evaluation for Women's Rights: Thirteen Insights for Women's Organizations*, Toronto, AWID.

BATLIWALA, S. and PITTMAN, A. (2010) *Capturing Change in Women's Realities. A Critical Overview of Current Monitoring & Evaluation Frameworks and Approaches,* Toronto, AWID.

CIDA (2001) *How to Perform Evaluations–Gender Equality*, Quebec, CIDA.

DAC (2010) *Development Evaluation Resources and Systems. A Study of Network Members*, OCDE.

DAC (1998) *Directrices y guía de conceptos del CAD sobre la igualdad entre mujeres y hombres,* Madrid, Ministerio de Asuntos Exteriores.

DE WAAL, Maretha (2006) "Evaluating Gender Mainstreaming in Development Practices" en *Development in Practice*, vol. 16, nº 2, pp. 209-214.

DFID (2008) *How to Note on Integrating Gender within Evaluation*, United Kingdom, DFID.

DFID (1999) *Gender Planning Frameworks*, United Kingdom, DFID.

ESPINOSA, Julia (2013) "Moving towards a gender-sensitive evaluation? Practices and challenges in international development evaluation" in *Evaluation*, Volume 19 Issue 2, pp. 171-182.

ESPINOSA, Julia (2012) "Gender issues in the evaluation of international aid. The experience of British, Spanish and Swedish development agencies", 10th European Evaluation Society International Conference.

ESPINOSA, Julia (2011) *La igualdad de género en la evaluación de la ayuda: los casos de la cooperación oficial británica, sueca y española*, Madrid, Complutense University of Madrid, available at Spanish at http://eprints.ucm.es/13206/

FIDA (1995) *Directrices Básicas para la Evaluación de las Cuestiones de Género*, Rome, FIDA.

GONZÁLEZ, L., and MURGUIALDAY, C. (2004). "Evaluar con enfoque de género", *Cuadernos Bakeaz*, 66, 1-16.

HUNT, J., and BROUWERS, R. (2003) *Review of Gender and Evaluation*, Final Report to DAC Network on Development Evaluation, DAC Evaluation Series, OCDE.

LIGERO, J.A.; ESPINOSA, J.; MORMENEO, C.; and BUSTELO, M. (2013) *Cómo hacer una evaluación sensible al género y al enfoque basado en derechos humanos para el desarrollo*, Madrid, Ministry of Foreing Affairs and International Cooperation.

LIGERO, J.A. (2011) *Dos métodos de evaluación*, Madrid, CECOD.

MILLER, C. and RAZAVI, S. (1998) "Gender Analysis: Alternative Paradigms*"* in UNDP (Ed.), *Gender in Development Monograph Series,* nº6**,** Geneve, UNDP.

MOSER, A. (2007) *Gender and Indicators,* Brighton, Institute of Development Studies.

MURGUIALDAY, C.; VÁZQUEZ, N.; and GONZÁLEZ, L.(2008) *Un paso más: Evaluación del impacto de género,* Barcelona, Cooperacció and AECID.

MURPHY, J. (1997) "Tracing Gender Issues Through Institutional Change and Program Implementation at the World Bank." in S. Chelimsky y W.R. Shadish (Eds.), *Evaluation for 21st Century. A Handbook*, London, Sage Publications, pp. 251-259.

RAO, Aruna y KELLEHER, David (2005) "Is there life after gender mainstreaming?" in *Gender and Development*, vol 13, nº 2, pp. 57-69.

RODRÍGUEZ, G., MELÉNDEZ, N., VELÁZQUEZ, E., and FUENTES, M. (1999) *Tomándole el pulso al género. Sistemas de monitoreo y evaluación sensibles a género,* San José de Costa Rica, Unión Mundial para la Naturaleza y Fundación Arias para la Paz y el Progreso Humano.

ROCHE, Chris (1999) *Impact Assessment for Development Agencies*, Oxford, Oxfam/NOVIB.

SANZ, B. and Evaluation Office, UN Women (2012) "Steps to commission and carry out Gender Equality Evaluations", in Workshop *Gender-sensitive evaluation: key ideas to commission and carry out it*, 10th EES Biennial Conference.

UNDP (2001) *Gender Analysis. Learning and Information Pack*, Gender in Development Programme, UNDP.

UNEG (2011) *Integrating Human Rights and Gender Equality* http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/IOS/temp/HRGE%20Handbook.pdf

UN Women (2011) *UN WOMEN'S Guide to Gender Equality/ Human Rights Responsive Evaluation*, available at http://unifem.org/evaluation_manual/

UN (1997) Report of the UN Economic and Social Council: "*Mainstreaming the gender perspective into all policies and programmes in the UN system*".

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