Helpdesk Research Report: Measuring and Evaluating Women’s Leadership

01.06.2012

Query: What types of measures are currently being used or are emerging to measure and evaluate women’s leadership (focusing particularly on studies on Asia – especially Indonesia and the Pacific)? Provide a summary of these measures and, if available, critical appraisal of their application.

Enquirer: AusAID

Author: Oliver Walton (oliver@gsdrc.org)

Contents

1. Overview
   2. Summary of Measures
   3. Critical Appraisal of their Application
   4. References
   5. Additional Information

1. Overview

As donors have become increasingly interested in promoting women’s leadership, a growing range of measures and indicators have emerged to measure and evaluate these efforts. Despite growing interest, however, there has been little sustained analysis of the issues associated with measuring women’s leadership in the literature and ‘surprisingly little existing research or analysis on the impact of women’s leadership’ (Gill et al 2009, p. 25). Because of the lack of analysis in this area, this report is largely based on a small number of available evaluation studies and project documents of women’s development programmes. These examples come mainly from Asia, but also include examples from Jamaica, Northern Ireland, Morocco and Malawi.

Measuring and evaluating women’s leadership is difficult because individual leadership is not easily separated from its context or from its nature as a shared and political process; because efforts to promote leadership are highly varied in their methods and address a wide range of leaders; and because results can often be evaluated only indirectly.

Leadership Development Programmes (LDPs) have often been ambiguous about what is meant by ‘leadership’, with a very small proportion explicitly defining the term (Lyne de Ver & Kennedy 2011). A
key distinction in understandings of leadership is between programmes that see leadership as an *individual attribute* and those that understand it as a *shared process* between leaders and groups (Lyne de Ver & Kennedy 2011). The next section provides examples of measures designed to capture changes in women’s leadership both at the individual attribute level and based on a ‘shared process’ understanding of leadership.

A growing body of literature, led by the recent work of the Developmental Leadership Program (DLP), stresses that leadership is best understood as a *political process* ‘that takes different forms in different contexts’ (DLP n.d., no page number). From this perspective, leadership involves capacity to mobilise people and resources ‘to forge coalitions with other leaders and organizations, within and across the public and private sectors’ (DLP, n.d., no p.n.). LDPs that adopt this political understanding of leadership present more challenges for evaluators since they necessitate a move away from individual skills towards a focus on supporting coalitions of actors (Lyne de Ver & Kennedy 2011). Evaluations of this kind require an understanding of the political context within which LDPs are being conducted, and require donors to adopt a more long-term perspective (Lyne de Ver & Kennedy 2011, Tadros 2011, Waring 2011). As several of the examples highlighted in this report show, few LDPs assess long-term impacts because few donors or implementing agencies adopt a long-term approach to their work, and because measuring impacts over the long-term is expensive and requires detailed planning.

Donor efforts to promote leadership in development are extremely varied in their aims, teaching methods and contents (Lyne de Ver & Kennedy 2011). Gill et al (2009, p.27) note that ‘while leadership is commonly measured by visibility in the political realm or the marketplace, women exert leadership across a range of private and public domains, often outside of the norms and mainstream institutions that dominate public life’. As a result, women’s LDPs tend to address a particularly varied set of audiences, which include existing leaders, potential leaders, grass-roots leaders, or sector-specific leaders.

Women’s leadership programmes use a range of quantitative and qualitative measures to assess immediate outputs and outcomes as well as wider impacts. The examples identified in this report largely confirm Lyne de Ver and Kennedy’s (2011, p.34) finding that the majority of LDPs evaluate ‘only at the individual or organisational level’ and that few programmes examine the wider impact of these programmes on policy or institutions, although some exceptions (UNIFEM Afghanistan, UNDEF Jamaica, AusAID Indonesia) were found. One of the key criticisms of current donor efforts to support women’s leadership has been a tendency to focus on narrow progress at the level of national political leadership at the expense of broader changes in political systems, or women’s leadership at the grass-roots level (Iknow Politics 2008, Tadros 2011, Waring 2011). A number of authors also stress the importance of moving beyond nominal measures of women’s political empowerment (such as the number of women in parliament) towards a more in-depth and contextualised analysis of what women leaders do once they gain office (Iknow Politics 2008, Waring 2011).

The next section provides a brief summary of some key measures currently being used to measure and evaluate women’s leadership. Section three outlines some of the issues associated with measuring women’s leadership.

**2. Summary of Measures**

This section draws largely on evaluation reports and other project documents from seven case study programmes. It also draws on a few examples that are briefly mentioned in the wider literature. The section begins by summarising the key types of indicators used in these examples. More details of the
indicators and some of the key issues associated with them are then provided in a table (the programmes are referred to in the text below using the name of the implementing agency and the country – e.g. UNIFEM Afghanistan). The measures can be distinguished in three main ways:

**Individual vs. broader environmental measures** – One programme (CE Forum Northern Ireland) focused solely on changes to individuals. Most, however, measured both the impact of leadership programmes on individuals or individual organisations and the broader institutional and policy environment.

**Short- vs. long-term measures** – Most programmes predominantly used measures that assessed short-term project outputs (such as the number of women receiving leadership training), while others also addressed wider impacts, such as improvements in the capacities of the Ministry of Women’s Affairs (UNIFEM Afghanistan).

**Quantitative vs. qualitative measures** – Most programmes used a range of quantitative measures (such as the number of women on school boards) and qualitative measures (such as changes in attitudes and perspectives of participants in leadership training programmes).

An overview of the main types of measures (with examples) is provided below:

**Broader Environmental Measures**

**Effectiveness and capacities of key government institutions** - Several programmes used the effectiveness and capacities of key government departments such as the Ministry of Women’s Affairs as key indicators for improvements in women’s leadership (UNIFEM Afghanistan, UNDP Cambodia).

**Policy changes** – A number of programmes used policy changes as key measures of the impact of women’s leadership programmes. Key indicators included the existence of a clear national strategy relating to gender equality (UNIFEM Afghanistan, UNDEF Jamaica, AusAID Indonesia).

**Effectiveness and capacities of key organisations** – Several programmes assessed outputs and outcomes achieved by women’s organisations, using an array of measures which varied according to the goals of the project. The AusAID Indonesia programme used effectiveness at engaging with the media as a key variable. UNIFEM Afghanistan, on the other hand, looked at the effectiveness of women’s entrepreneur organisations in improvement women’s economic security. The UNDP Cambodia project cited the fact that a group of trained women’s leaders had formed a ‘women’s civil servants association’ as a measure of success. A women’s organisation in Malawi (‘Women’s Forum’) was mobilised after a leadership training programme, and as a result three women were allocated tracts of land (Pittman 2011). Sharma & Sudarshan’s (2010) study of a women’s leadership programme in India uses the fact that women’s groups have engaged with local government as a measure of the success. The use of organisational effectiveness as a measure of leadership has been questioned by Jing and Avery (2008, p.74), who state that ‘existing research on the leadership-performance relationship is full of difficulties and has many unsolved problems, including methodological problems. Thus, conclusions cannot be drawn about the extent to which leadership behaviors and styles facilitate the improvement of organizational performance’.

**Initiation of advocacy campaigns** - A Women’s Learning Partnership training programme in Morocco led one group of women to ‘use advocacy tools gained through the training and organized a grassroots campaign to demand equal rights to men in their right to compensation for collective lands’ (Pittman 2011, p.8).
Access to information about women’s rights – UNIFEM Afghanistan measured the number and types of users of the situation analysis, media briefs and database of information on Afghan women, and the number and quality of media strategies/broadcast programmes promoting women’s rights by trained media workers.

Improved outcomes for beneficiaries – For example, AusAID Indonesia used improved access to social protection programmes as a measure of the broader impact of its project.

Gender perspectives and notions of women’s empowerment integrated into UN and donor strategy for rebuilding Afghanistan – UNIFEM Afghanistan measured the number of UN initiatives to support gender equality, number and type of expert opinions provided by UNIFEM.

Individual measures

Changes in attitude – Several programmes measured how women felt after leadership training, with several reporting that they felt more confident (UNDP Cambodia, CE Forum Northern Ireland, UNDEF Jamaica).

Changes in understandings of leadership - Participants in Women’s Learning Partnership training programmes in 2008 from a number of locations around the globe ‘consistently report a sustained shift in their view of leadership from hierarchical and power-oriented models to democratic models based on consensus-building, participation, and cooperation’ (Pittman 2011, p. 8).

A range of individual outcomes and impacts – Mostly quantitative measures measuring short-term project outcomes such as the number of women trained in leadership (UNDEF Jamaica, World Bank Indonesia) and longer-term impacts: number of women on school boards (UNDEF Jamaica), number of women running for office (World Bank Indonesia), career progression of trained women (CE Forum Northern Ireland), impact of trained women on their networks (CE Forum Northern Ireland). Some programmes used qualitative measures for example, Pittman (2011, p.8) describes how after leadership training participants established ‘more equal relations in the family and participate in civil society and political activities at higher levels than non-participants’.

The table below provides more detailed information about the seven projects and programmes reviewed for this report. Where information has been available it highlights the main goals of each programme, key indicators used, the methods used to evaluate the programme and key measurement issues.
<table>
<thead>
<tr>
<th>Project name</th>
<th>Goal/ Objectives</th>
<th>Key Indicators</th>
<th>Methods</th>
<th>Key Measurement Issues</th>
</tr>
</thead>
</table>
| UNIFEM Afghanistan Women’s Leadership programme (2002-2006) (Saltmarsh e & Kuovo 2008) | Increase options and opportunities of Afghan women to transform the overall development of Afghanistan into a more equitable and sustainable process. This was to be achieved through strengthening women’s capacity and opportunities to participate in politics and governance, promoting gender justice and women's human rights and increasing women’s economic opportunities. | - Effectiveness and capacities of the Ministry of Women’s Affairs (MOWA), including the existence of a clear strategy, a national gender policy, and policies for engaging with a variety of other partners, activities of outreach centres (women’s development and community centres).  
- Effectiveness of Afghan legislative and policy frameworks (existence of laws, number of women participating in governance processes, levels of women’s access to legal redress for rights violations).  
- Access to information about women’s rights (number and types of users of the situation analysis, media briefs and database of information on Afghan women, number and quality of media strategies/broadcast programmes promoting women’s rights by trained media workers).  
- Effectiveness of organizations and groups in achieving women’s economic security (number and type of organisations established, number of women benefitting from the Afghan Business Women Council’s (AWBC) outreach, number of important issues raised by AWBC, number of producers and entrepreneurs supported).  
- Gender perspectives and notions of women’s empowerment integrated into UN and donor strategy for rebuilding Afghanistan (number of UN initiatives to support gender equality, number and type of expert opinions provided by UNIFEM) | Data collection in 2008 (six years after project inception). Survey undertaken with 3 units to measure perceptions of success. Interviews with 100 key informants. | High pressures of work and deadlines are ‘leading to oversights in planning processes and reporting’ (p.64).  
‘The failure to systematically use a logframe as a reporting tool that links strategically determined activity makes it hard to judge the extent to which the programme is achieving its purpose. It would be valuable to formally clarify the programme outcomes and the goal’ (p.65).  
‘Reporting is largely narrative in nature creating large quantities of data that are not easy to assimilate. Reporting takes place for purposes of accountability and recording progress rather than to understand process …Given this, it is not unsurprising that there is a shortage of systematically structured qualitative indicators to judge programme performance. Similarly, quantitative indicators are not systematically recorded in a way that can produce information about programmes. Finally, there is an absence of baseline data against which to measure change’ (p.65). |
| **UNDP Cambodia Women’s Leadership Programme (2003-2006) (Platt & Vutheary 2006)** | Builds the skills of Ministry of Women’s Affairs (MoWA) and key line Ministry staff to carry out their work more effectively in the promotion of gender equity in Cambodia. The programme also aims to ensure all national policies and programmes are gender responsive. | No detailed information about indicators used is provided in the report. Nevertheless, some of these can be inferred from the evaluation’s findings:  
‘[t]he trained women clearly value the WLP course and feel significantly more confident and better equipped to fulfil their roles’ (p.5).  
The evaluation also highlighted another key impact in the decision by trained women’s leaders to form a ‘women’s civil servants association’. | The evaluation adopted a participatory approach. Interviews were conducted with a wide range of stakeholders. Participants were asked to reflect on the programme and to assess its value. | ‘Proxy indicators such as promotion of women after training and the increased networking of women resulting from alumnae meetings and study tours provide some measure of the impact of the programme. However, the lack of ongoing monitoring and reporting against established baselines and using indicators over the period of the programme makes it difficult to provide precise evaluation of impact. Furthermore, the programme covers 24 Ministries with wide variations in their cultures and commitment to supporting women as leaders’ (p.8). |

| **UNDEF Jamaica Strengthening Women’s Leadership Project – implemented by Women’s Resource and Outreach Centre) (2009-2011) (UNDEF 2011)** | To address the under-representation of women in decision making positions within Jamaica, and in particular on private boards and public commissions. This was to be done by:  
- Increasing the participation of women on boards and commissions at the national level through training and awareness building; and  
- Increasing the participation of women in leadership in community-based organizations (CBOs), including school boards, also through training and awareness building.  
- Creating a national conversation on the need to open spaces for the greater participation of women in decision making positions in Jamaica. | Policy changes (development of new criteria for participation on school boards by the National Council of Education)  
- Increased number of women trained (and their demeanour and sense of empowerment),  
- Increased national awareness and increased participation of women in leadership  
- Increased number of women on boards (future indicator) | The evaluation field work took place in July 2011, well after the project had been completed. The evaluators relied on 1) project and other documents; 2) interviews; and 3) project-related products. | The approach to impact evaluation was not very rigorous. The project did not measure changes in public awareness and therefore the evaluation stated that it was ‘impossible to assess the degree of change and action generated by project activities’ (p.17). However, the evaluators stated that based on ‘anecdotal information’, it is evident that the project reached two to three times the number of women trained and in all likelihood, well beyond that through the word of mouth by the participants and community activities’ (p.17). |
<table>
<thead>
<tr>
<th>Programme</th>
<th>Goal</th>
<th>Theory of Change</th>
<th>Monitoring and Evaluation</th>
<th>Impact</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search for Common Ground Indonesia Transformational Leadership Project (SFCG 2011)</td>
<td>Empower the elite and emerging women politicians in Indonesia as agents of social change.</td>
<td>No info</td>
<td>No info</td>
<td></td>
<td>The programme had a clearly articulated theory of change, but this was ineffective and 'led to partial approach to addressing the problem' (p.6). An elaborate M&amp;E plan was devised at the outset, but this was not followed 'due to limitations of budget' (p.6). Baseline data collection was not conducted, and some monitoring data was not collected 'due to tensions with one partner organisation' (p.6). 'Too many indicators of dubious value were included in the log frame' and that no approach was established 'to capture evidenced changes as they occurred (other than through quarterly reporting to donors)' (p.6)</td>
</tr>
<tr>
<td>AusAID Indonesia MAMPU Programme (AusAID 2012, 2012a)</td>
<td>To build networks and coalitions led by strengthened women's and gender interested organisations who will influence government policies, regulations and services, and in selected private sector arenas, in order to improve the access of poor women to critical services.</td>
<td></td>
<td></td>
<td></td>
<td>No info</td>
</tr>
<tr>
<td>World Bank Indonesia Sustaining Women's Leadership (World Bank 2010)</td>
<td>Reduce poverty and vulnerability among female-headed households in the poorest parts of Indonesia by facilitating the training of women leaders and develop the capacity of local women's organizations to set up special programmes that benefit female heads of households in poor provinces.</td>
<td>A range of measures including, for example:</td>
<td>No info</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- To provide opportunities for senior women to develop themselves as leaders, to enhance their visibility and to enable them to network and influence the wider external environment.
- To develop a programme for senior women to increase the number of women progressing up the career ladder to higher level management positions within the next five years.
- To increase the representation of women at chief executive level and above from its 2003 level of approximately 12% of the public service chief executives in Northern Ireland to 20% in five years and 30% in ten years.

Participants’ current grade, promotion and career history in order to track senior women’s career progression.
- Extent to which the programme has enabled participants to develop themselves as leaders, to enhance their visibility and to enable them to network and influence the wider external environment, system-level effects of the Programme (organisational ‘ripple effects’)

An online survey of all those who attended the programme between 2003 and 2005 was conducted, followed by telephone interviews with a sample of participants, and with programme sponsors and facilitators.

No info
Critical Appraisal of their Application and Best Practice

All of the seven LDPs examined for this report established some measures to assess the programmes’ impacts. In several cases, baseline indicators were established as part of a monitoring and evaluation plan drawn up at the start of the project, but this was not implemented and indicators were not monitored consistently (UNIFEM Afghanistan, UNDP Cambodia, SFCG Indonesia). In the UNIFEM Afghanistan and UNDP Cambodia programmes, work pressures meant that impact was not measured consistently and baseline data were not properly collected. In the case of SFCG Indonesia, a theory of change was developed but this was flawed. The elaborate monitoring and evaluation plan was not followed due to budget restrictions. In two other cases (Chief Executives Forum Northern Ireland and UNDEF Jamaica) there were no rigorous attempts to monitor indicators of impact against baselines. Instead, impact measures were simply examined after the project during the evaluation phase.

Lyne de Ver and Kennedy (2011, p.34) find that the majority of LDPs evaluate ‘only at the individual or organisational level’ and that ‘most provide only anecdotal evidence of participant satisfaction, used as evidence of effectiveness of programmes’. So, for example, a ‘population leadership’ programme targeting health professionals evaluated the impact on health professionals, but not the wider intended impacts in the areas of family planning and reproductive policy, nor the wider social impacts the programme may have had. This finding is largely supported by the examples provided in the last section, which focused mostly on the impact of programmes on individuals or organisations. Only a few examples appeared to successfully measure the wider impact on institutions or policy. A common indicator of programme success, highlighted both in some of the examples examined in this report and in Lyne de Ver and Kennedy’s (2011) wider survey, is an impact assessment that only addresses participant satisfaction, measured through participant feedback surveys.

Lyne de Ver and Kennedy (2011) stress the importance of developing a clear theory of change, to guide the programme’s methods and contents, and to enable the programme to evaluate its effectiveness. In a general survey of LDPs, Lyne de Ver and Kennedy (2011, p.v) found that the vast majority (57 out of 67) did not clearly explain ‘the processes through which leadership was developed, and how this leadership then creates change’. They argue that a ‘well-conceived and well-run LDP should, at least, provide:

Rigorous evaluations or impact assessments that validate a programme’s theory of change; allow for the identification of short-term and long-term impact; and enables the programmes to make necessary adjustments when output, outcomes or impact fall short of expectations’ (p.42).

Measuring women’s political leadership

Tadros (2011, p.10) argues that a ‘policy shift is required from the current focus almost exclusively on getting women into legislatures to providing women opportunities for political apprenticeship, and ultimately women’s leadership’. Tadros (2011) recommends a broader focus on leadership within NGOs, clubs, community centres, universities, schools and in the workforce, although also stresses that the suitability of any approach will be heavily dependent upon context.

These issues have also been highlighted by Waring (2011), who has written critically about donor efforts to enhance women’s leadership. She argues that improvements in women’s leadership often fail to promote wider improvements in women’s societal position, because they are not driven by disempowered people. Instead, women leaders are more likely to ‘get elected’ and promote themselves. ‘The same old figures’ show up at regional women’s leadership programmes in the
Pacific region, emphasising that who is included in these programmes is of critical importance to their transformative potential. She advocates the use of more ‘textured data’, where measures of impacts are more closely calibrated to specific political contexts; using a mixture of qualitative and quantitative methods; and the use of participatory approaches (Waring 2011).

In an online discussion on women’s leadership from 2008, one practitioner stated that it is important not simply to measure numbers, but also to measure the impact that women have once they are involved in politics (Iknow Politics 2008). This discussion also included responses from some practitioners who stated that they used indicators such as the number of women in various political institutions, the number of women who ran as candidates and won, and the number of women who received training – “[t]his helps to see where women are falling through the political cracks” (Iknow Politics 2008, no page number). Another practitioner suggested that it was important to get beyond nominal participation, by assessing the extent to which women are actually involved in political decision-making. This could be done, for example, by measuring ‘whether women parliamentarians sponsor any pieces of legislation, how they vote on legislation, and how active they are in committee or plenary debates’ (Iknow Politics 2008, no page number). This, however, may be difficult to do since it would require defining which positions are involved in decision-making, which may be politically sensitive. Ascertaining this would also be time-consuming (Iknow Politics 2008).

**Reform coalitions**

Recent research on women’s leadership has emphasised the importance of coalition-building in driving social and political change. The Developmental Leadership Programme has adopted the following definition of leadership for developmental purposes:

*The process of organizing or mobilizing people and resources in pursuit of particular ends or goals, in given institutional contexts of authority, legitimacy and power (often of a hybrid kind). Achieving these ends, and overcoming the collective action problems which commonly obstruct such achievement, normally requires the building of formal or informal coalitions of interests, elites and organizations, both vertical and horizontal* (Lyne De Ver 2009, p.9).

Tadros (2011a) argues that recognising the need for coalitions implies a more complex understanding of change that requires better political analytical skills from donors to read the interactions between leaders and enabling factors that facilitate the formation of coalitions. Reform coalitions may also imply the need for donors to accept that success cannot be attributable to a single variable such as the presence of seasoned women’s activists, or their expertise in gender issues, but rather that success or failure will depend on an interaction of their capacities with the shifting political and institutional context (Tadros 2011).

### 6. References


http://www.norad.no/no/resultater/publikasjoner/norads-samlede-rapporter/publikasjon?key=109804


7. Additional Information

Experts Consulted:

Anne Marie Goetz, UNIFEM
Marilyn Waring , AUT University
Ana Lukatela, UNIFEM

Key websites:

Association for Women’s Rights in Development (AWID) http://www.awid.org/ 
Developmental Leadership Program (DLP) http://www.dlprog.org/ 
International Knowledge Network of Women in Politics (Iknow Politics) http://www.iknowpolitics.org/ 
NORAD Evaluation http://www.norad.no/en/evaluation 
Pathways of Empowerment http://www.pathwaysofempowerment.org/

About Helpdesk research reports: Helpdesk reports are based on 2 days of desk-based research. They are designed to provide a brief overview of the key issues, and a summary of some of the best literature available. Experts are contacted during the course of the research, and those able to provide input within the short time-frame are acknowledged