AOSIS

UN Women's experience with strengthening evaluation systems in Africa: Enhancing quantity, quality and use of evaluations

CrossMark
click for updates

Author:
Caspar Merkle[1]

Affiliation:
[1]UN Women Regional Office for Eastern and Southern Africa, Nairobi, Kenya

Corresponding author:
Caspar Merkle,
caspar.merkle@unwomen.org

Background: Following the adoption of the Women Evaluation Policy in 2012, a series of systems and mechanisms were introduced in the organisation to strengthen the evaluation function at both central and decentralised levels. They were based on a systemic approach and a Theory of Change for building an enabling environment for evaluation in UN Women.

Objectives: The purpose of this article was to analyse progress made and challenges with respect to establishing evaluation systems and institutionalising an evaluation culture in the UN Women Africa region.

Method: The article draws on UN Women evaluation performance data collected over the past five years, discussions and practical experience by the author of working on evaluation with UN Women since 2009. It also analyses UN Women documents and the broader literature on the topic.

Results: The findings illustrate that the different mechanisms to strengthen the evaluation function in UN Women show progress in the Africa region on four out of the five selected evaluation performance indicators. The Theory of Change to strengthen the UN Women evaluation function is largely validated by the wider literature on evaluation use. External assessments confirm that the UN Women evaluation function is sound overall.

Conclusion: The article concludes that evaluation performance indicators only provide a partial snapshot of the many different factors that help or undermine evaluative thinking and a learning culture within an organisation. Institutional systems and mechanisms are necessary but not sufficient for nurturing an evaluation culture and ensuring utilisation of evaluation for better development effectiveness.

Introduction

The purpose of this article is to analyse UN Women's experience with strengthening evaluation systems in the Africa region since 2010. UN Women was created in 2010 by a merger of four different organisations in the United Nations (UN) system and, following this, has undergone an organisational transition process that included substantial changes in the UN Women evaluation function.

The article examines the Theory of Change for strengthening the UN Women evaluation function and provides an analysis of the changes resulting from the systems and mechanisms introduced in UN Women to institutionalise an evaluation culture over the past 5 years. In the context of this article, 'evaluation culture' denotes an organisational culture that deliberately seeks out information on its performance in order to use that information to learn how to do better and thereby improve its performance (Mayne 2008). The article also compares the findings with another conceptual framework developed by the UN Evaluation Group (UNEG) describing factors that discourage or encourage evaluation use.

Methodology

The research builds on a trend analysis of data on five different evaluation performance indicators in UN Women during the period 2009 to 2015: (1) the number of staff dedicated to monitoring and evaluation (M&E) at country level in Africa, (2) the development of evaluation expenditures, (3) the total number of evaluations completed in the region, (4) the number of evaluation management responses, and (5) the quality of completed evaluations. It should be

mentioned that UN Women has introduced more evaluation performance indicators over the period, e.g. on evaluation coverage, implementation of management response, use of evaluations for programming and number of M&E officers/ focal points who have completed the corporate e-learning course on gender responsive evaluation. These additional performance indicators are omitted in this analysis as they were introduced only recently.

The article draws on the broader literature as well as on a series of UN documents, e.g. UN Women Annual Reports on Evaluation, UN Women evaluation quality assessments as well as recent external reviews of the UN Women evaluation function by the UNEG, the Joint Inspection Unit (JIU) and the Multilateral Organisation Performance Assessment Network (MOPAN). Finally, the author draws on his practical experience of working on evaluation with UN Women in the Africa region since 2009.

Background

In July 2010, the United Nations General Assembly created UN Women, the United Nations Entity for Gender Equality and the Empowerment of Women. The creation of UN Women came about as part of the UN reform agenda, bringing together resources and mandates of four previously distinct parts of the UN system which focused exclusively on Gender Equality and Women's Empowerment (GEWE), namely the Division for the Advancement of Women, the International Research and Training Institute for the Advancement of Women (INSTRAW), the Office of the Special Adviser on Gender Issues and Advancement of Women and the United Nations Development Fund for Women (UNIFEM).

Following the creation of UN Women, the Independent Evaluation Office (IEO) expanded significantly. Since 2010, the UN Women Evaluation Office doubled from eight staff (five based at Headquarters and three Regional Evaluation Specialists in Africa, Asia Pacific and Latin America) to 16 staff in 2014 and 2015 (10 based at Headquarters and six Regional Evaluation Specialists in Western and Eastern Africa, Asia Pacific, Latin America, Eastern and Central Europe and the Arab States). Regional Evaluation Specialists are IEO staff deployed in UN Women Regional Offices who report to the Director of the IEO at Headquarters. They play a key role in providing oversight, quality assurance support and capacity development to the planning, management and follow-up of programme evaluations at country level.

A revised UN Women Evaluation Policy, which was adopted in 2012, outlines the guiding principles and standards of the UN Women evaluation function and its role in terms of accountability, decision-making and learning. The guiding principles include national ownership and leadership, UN System Coordination and Coherence on GEWE, innovation, fair power relations and empowerment, participation and inclusion, independence and impartiality, transparency,

quality and credibility, intentionality and use of evaluation and ethics.

Following the adoption of the Evaluation Policy, a series of systems and mechanisms were introduced to strengthen the evaluation function at both central and decentralised level. These mechanisms will be described in detail below. They were based on a systemic approach in building an enabling environment for evaluation in UN Women by strengthening (1) the capability of managers to demand and use evaluation and (2) the capability of UN Women offices to produce and supply evaluative evidence. Figure 1 illustrates the Theory of Change for this systemic approach. For the purpose of this article a Theory of Change is understood as the 'beliefs, assumptions and hypotheses about how change happens – about the way humans work, or organisations, or political systems, or eco-systems. The Theory of Change is about articulating these many underlying assumptions about how change will happen in a programme' (Vogel 2012:4). The Theory of Change helps to conceptualise a change process through illustrating, generally in graphical form, the presumed causal relationships between the various elements that are necessary for realising the desired long-term goal. The Theory of Change to strengthen the UN Women evaluation function explicitly takes into account UN Women's role in promoting and advocating for gender responsive evaluation to achieve more effective development for women and girls.

Key institutional elements for strengthening *evaluation demand* include systematic awareness raising mechanisms for UN Women senior management, an oversight dashboard based on evaluation key performance indicators and appropriate financial resources for evaluation. These mechanisms are expected to contribute to changes in managerial behaviour with respect to appreciating evaluation, using evaluation findings and assuming accountability for evaluation in their respective area of work. Key institutional factors for improving the *supply of evaluation* comprise evaluation quality assurance systems, evaluation capacity development opportunities, technical support at different levels and staffing of skilled M&E personnel. These factors are expected to enable relevant evaluation planning and the production of high quality evaluations at country level. Taken together, increased demand for evaluations by senior management and improved supply of quality evaluations will contribute to better use of evaluation evidence in UN Women for decision-making and reporting.

The Theory of Change makes a number of assumptions about the pre-conditions for achieving the proposed results, such as adequate financial and human resources, a supportive senior management and UN Women Executive Board, continuity in M&E staffing and an results-based management (RBM) organisational culture. The organisational change process outlined in the Theory of Change should not be seen as a linear model but as a complex system where change happens through feedback loops, reversals and sometimes even
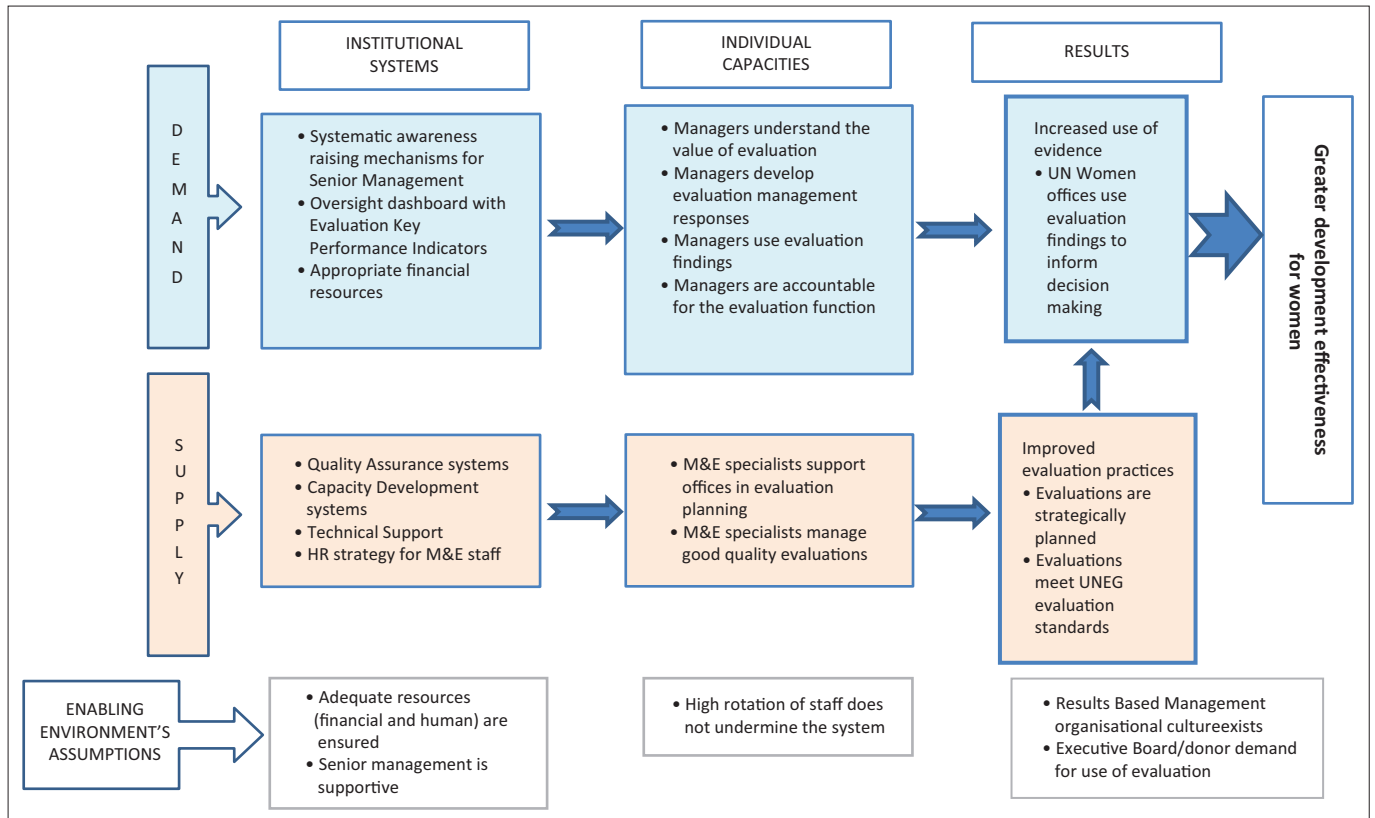
Source: Adapted from UN Women Independent Evaluation Office 2014b

FIGURE 1: Theory of Change to strengthen the UN Women evaluation function.

backlashes (UN Women Independent Evaluation Office 2014b).

A description of some of the key mechanisms that were introduced to strengthen the UN Women evaluation function follows.

Systems and mechanisms to strengthen the UN Women evaluation function

The systems and mechanisms to strengthen the UN Women evaluation function described below were introduced as from 2013 and at different points in time.

UN Women Global Evaluation Advisory Committee: The Global Evaluation Advisory Committee (GEAC) was constituted in December 2013 with the purpose of providing advice to the UN Women Executive Director and the IEO on the overall evaluation function at both corporate and decentralised level. The GEAC is composed of five external independent evaluators representing different geographical areas and institutional backgrounds, two senior evaluation experts from UN agencies and Bretton Woods institutions and two UN Women senior management colleagues from Headquarters and regional level.

UN Women Global Evaluation Oversight System: The Global Evaluation Oversight System (GEOS) was introduced in May 2014 with the aim to build awareness on the evaluation function amongst UN Women senior management. It consists

of a dashboard with seven key evaluation performance indicators covering the following areas: evaluation coverage during a 4-year planning cycle, evaluation delivery, human and financial resources for evaluation, evaluation report quality, evaluation follow-up through management responses and the implementation status of evaluation management responses. The evaluation performance indicators are monitored globally by the IEO on a bi-annual basis and updates with breakdown by geographical region are shared with senior management.

UN Women Global Accountability and Tracking of Evaluation Use System: The Global Accountability and Tracking of Evaluation Use (GATE) system was launched in May 2013. It is an online-based information management system to facilitate UN Women's efforts to effectively plan and use evaluations for accountability, management for results and knowledge management. It represents a repository of all UN Women evaluation reports, along with their quality rating, that is publicly accessible at http://gate.unwomen.org/.

UN Women Global Evaluation Reports Assessment and Analysis System: The Global Evaluation Reports Assessment and Analysis System (GERAAS) was introduced in 2013 with the purpose to provide an objective assessment of the quality of UN Women evaluation reports. It is also aimed at strengthening internal evaluation capacity by providing feedback on strengths and weaknesses of individual evaluation reports and promoting learning through a meta-synthesis of evaluation findings, conclusions and recommendations.

UN Women Online Evaluation Manual: In order to develop internal evaluation capacities, the IEO has developed guidance such as the UN Women online evaluation manual 'How to manage gender-responsive evaluation', available at http://genderevaluation.unwomen.org/en/evaluation-handbook. The evaluation manual explains and provides practical tools for the various steps in the evaluation process starting with evaluation planning, evaluation management to utilising and following up on evaluations recommendations.

UN Women Professionalisation Initiative to strengthen the UN Women evaluation function: In 2015, the IEO launched a Professionalisation Initiative that comprises an e-learning course based on the UN Women Evaluation Manual described above, and a coaching programme whereby the knowledge acquired in the e-learning course is applied through managing a practical evaluation process. The initiative also introduced evaluation process standards for decentralised evaluation (UN Women Independent Evaluation Office 2015b).

UN Women Gender Equality Evaluation Portal: The Gender Equality Evaluation Portal serves as a tool to strengthen and promote the exchange of evaluation findings, conclusions and recommendations in the area of GEWE. It provides a database of currently more than 400 evaluations conducted by the UN and other organisations, available at http://genderevaluation.unwomen.org/.

Findings

Firstly, evaluation resources such as M&E staff at country level and evaluation expenditures by UN Women offices in the Africa region are analysed.

The development of UN Women *staff dedicated to M&E at country level* in Africa needs to be understood in the context of the UN Women regional architecture, which evolved since the creation of UN Women in 2010 with a view to further decentralise and strengthen UN Women's presence at country level. The process of strengthening UN Women country offices contributed to a greater number of offices with dedicated M&E staff. The number of offices with dedicated M&E staff increased from four offices in 2009 to 13 offices in 2015 (Figure 2), which represents more than 50% of the total offices in the Africa region.

Regarding *financial investment in evaluation*, the UN Women Evaluation Policy recommends a minimum level of evaluation investment of 3% of the programme budget. Whilst resources invested in evaluation fluctuate according to annual programme and evaluation planning cycles, this indicator provides an overall indication of the financial commitment of UN Women to the evaluation function. Total global expenditures invested by UN Women in evaluation have increased from 1.3% in 2011 to 2% in 2015. During that period the financial investment in decentralised

evaluations (in all geographical regions) increased from USD 687 000 in 2011 to USD 1 650 000 in 2015 (UN Women Independent Evaluation Office 2014a, 2015a, 2016). The absence of norms to guide evaluation budget allocation in most UN agencies makes it difficult to compare or benchmark across UN entities, as observed by JIU (2014) in the external assessment of evaluation functions in the UN system. However, when looking at UN Women only, the financial increase represents a positive move towards achieving the target of 3% as established in the UN Women Evaluation Policy.

In terms of *number of evaluations completed* by UN Women in Africa, we can observe a steady increase from one evaluation completed by UN Women in the Africa region in 2009 to 10 completed evaluations in 2013 (Figure 3). Following this peak, completed evaluations decreased to seven in 2014 and 2015. Amongst the main reasons for the decline in completed evaluations are insufficient funding for evaluations, weak quality of consultants' work, over-planning and the late initiation of evaluations (UN Women Independent Evaluation Office 2015a). This points towards the need to ensure good quality evaluation planning and implementation.

The UN Women Evaluation Policy mandates a *formal management response to evaluation recommendations* within 6 weeks upon evaluation completion. The number of completed management responses can be interpreted as
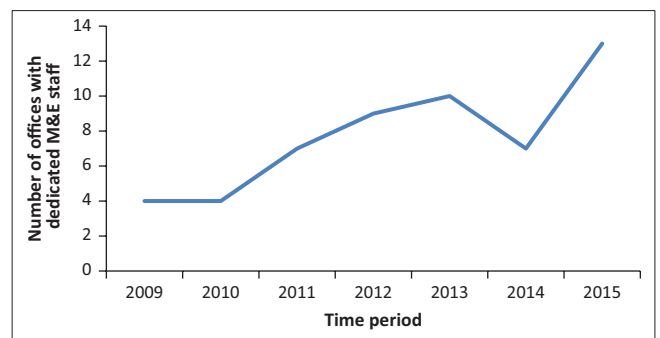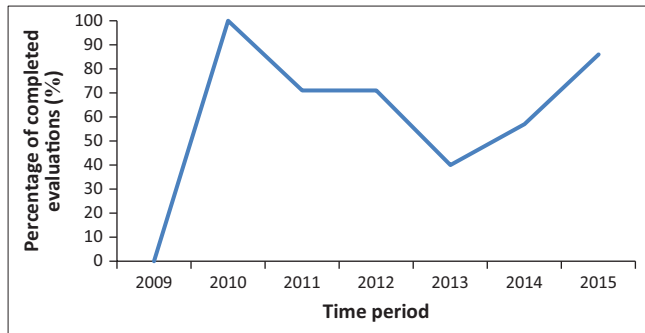
Source: UNIFEM Evaluation Office 2010, 2011; UN Women Independent Evaluation Office 2012, 2013, 2014a, 2015a, 2016
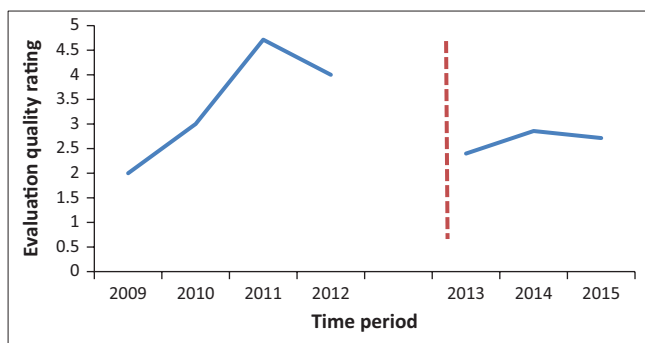
FIGURE 2: UN Women offices in Africa with M&E staff.
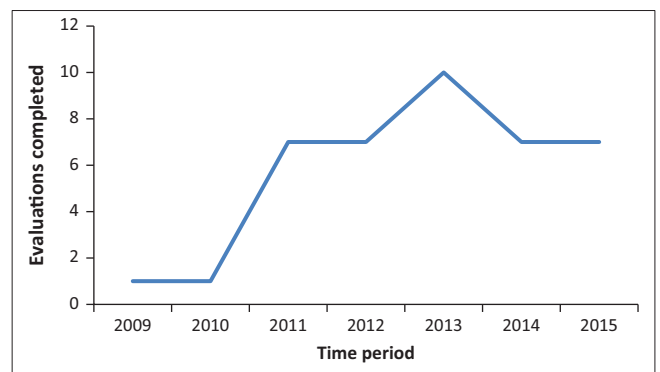
Source: UNIFEM Evaluation Office 2010, 2011; UN Women Independent Evaluation Office 2012, 2013, 2014a, 2015a, 2016

FIGURE 3: Evaluations completed by UN Women offices in Africa.

Source: UNIFEM Evaluation Office 2010, 2011; UN Women Independent Evaluation Office 2012, 2013, 2014a, 2015a, 2016

FIGURE 4: Percentage of completed evaluations in Africa with management response.

Source: UNIFEM Evaluation Office 2010, 2011; UN Women Independent Evaluation Office 2012, 2013, 2014a, 2015a, 2016

FIGURE 5: Average evaluation quality in Africa.

proxy indicator for evaluation utilisation. The trend in terms of evaluation management responses fluctuated (Figure 4). In 2010, there was 100% management response completion; however, this was based on one evaluation only. This decreased to 40% completion in 2013, when four out of 10 evaluations had a formal management response. Since then the trend has been positive: there was a 83% management response completion rate in 2015, when six out of seven evaluations had a management response. External assessments have confirmed that UN Women systems for facilitating use of evaluation results, such as the GATE system and the GERAAS system described above, are positive contributions to ensuring utility. However, more could be done to facilitate lesson learning and to enhance utility of evaluations (UN Women Global Evaluation Advisory Committee 2015). Potential factors associated with the use of and demand for evaluations will be discussed in the section on analysis and conclusions.

Finally, the progress in *evaluation quality* in the Africa region since 2009 is examined. UN Women used a six-point scale (excellent, very good, good, average, weak, very weak) for rating evaluation quality during the period 2009 to 2012. This rating system was replaced by a four-point scale (very good, good, satisfactory, unsatisfactory) in 2013. This modification in the evaluation quality assessment system poses a challenge to comparing evaluation quality before and after 2013. Examining data before 2013, the quality of

evaluations in Africa improved from an average rating of 'weak' in 2009 to an average of 'very good' in 2011 (Figure 5). In 2012, evaluation quality dropped to an average of 'good'. After 2013 and based on the new 4-point scale rating system, we can note an overall improvement in average evaluation quality in the Africa region from 'satisfactory' to 'good' (Figure 5). Whilst external assessments of the UN Women evaluation function have confirmed high levels of institutionalisation regarding implementation of evaluation policies and systems in UN Women, they also note that action is necessary for ensuring the quality and credibility of evaluation reports (UN Women Global Evaluation Advisory Committee 2015).

Analysis and conclusions

The different systems and mechanisms to strengthen the evaluation function in UN Women show progress in the Africa region on four out of the five selected evaluation performance indicators discussed above. Since 2009, there has been a progression in the areas of M&E staffing at country office level, evaluation financial investment, evaluation management response completion and evaluation quality. Evaluation delivery in the region remains a concern. The author has observed that the IEO's reporting on evaluation performance indicators to UN Women senior management in itself has helped to reinforce awareness on the importance of evaluation in general and has created positive competition amongst UN Women offices.

It is useful to analyse the findings on selected evaluation key performance indicators in the Africa region in the context of the Theory of Change to strengthen the UN Women evaluation function presented above. The UN Women Theory of Change will also be compared with another conceptual framework developed by the UNEG for enhancing evaluation use. When interpreting the findings, it is important to note that the data from 2009 to 2015 on the five selected evaluation performance indicators only partly reveals changes in the evaluation culture in the organisation. Evaluation performance indicators provide a partial snapshot of the many different factors that help or undermine attempts to foster an evaluative thinking and learning culture within an organisation. The Theory of Change for the UN Women evaluation function includes many more elements necessary for and assumptions regarding nurturing an evaluation culture and greater evaluation utilisation. As Mayne (2013:6) points out 'a lot can go wrong. Many diverse factors have to come together to make evaluation work well, that is, get used in organisations. This might explain why it is so difficult to have good utilization of evaluation in an organisation; many factors have to line up'.

Looking at the UN Women example, whilst the follow-up to evaluation recommendations through the formal management response mechanism constitutes a necessary step for institutionalising a system for utilising evaluations, the focus on the management response potentially misses out

on other factors and mechanisms that encourage or discourage use of evaluation findings. Data on other critical elements of the Theory of Change, in particular 'soft' aspects such as awareness, appreciation and commitment towards evaluation amongst senior management, is not tracked systematically. It would not be realistic from a practical and methodological perspective to capture such data.

Another example is the role and level of engagement of stakeholders and potential users during the evaluation process.

TABLE 1: Framework for factors (potentially) associated with use and demand.

Factors	Use and demand
Context in which the evaluation takes place	
(1) Evaluation culture	Support and commitment by senior management / Strong evaluation champions
	Presence of evaluation advisory groups
	Good organisational understanding of evaluation
	Presence of evaluation policy
	Respect for/visibility of the evaluation function
	Quality of the M&E infrastructure in place
	Evaluative thinking/use of evaluation information in decision-making
	Previous positive experience with evaluation/demonstrated use
	Availability and awareness of evaluation reports and products
(2) Organisational structure and incentives	Strength and position of the Evaluation Unit
	External pressures for accountability/information
	Root of the demand for evaluations
	Systems in place to feed evaluation findings into decision-making processes
	Incentives to learn
	Incentives to act on evaluation recommendations
	Ability of the organisation to implement recommendations
	Organisational stability / Human resources
	Resource availability
(3) Characteristics specific to the evaluation activity	Presence of evaluation champion for the specific evaluation
	Nature of the subject being evaluated
	Resources dedicated to this specific evaluation
	Presence of individuals /networks/intermediaries/brokers to facilitate the evaluation
	Timing of the evaluation in the policy cycle
	Political climate
	Main users' capacity/ability to receive and implement findings
	Nature of the evaluation findings
Evaluation activities	
(4) Evaluation design and process	Involvement of potential users at an early stage
	Planning for timeliness in the policy cycle
	Planning communication and dissemination early in the evaluation process
	Appropriateness and relevance of the evaluation approach
	Rigour and quality of the design, data sources and triangulation
	Encouragement of stakeholders to reflect critically on the project
(5) Stakeholder/user involvement	User involvement from the beginning
	Mechanism(s) for user involvement at all stages of evaluation
	User involvement in defining evaluation questions
	Participatory relationship between evaluation staff and programme staff
	Sharing preliminary findings with users
	Face-to-face meetings with producers and users of knowledge take place
	Direct participation of users in evaluation teams
(6) Evaluator qualities	Personal qualities: independence, credibility, openness, flexibility, adaptability
	Procedural qualities: communication and facilitation skills
	Technical competences
(7) Evaluation product	Timeliness
	Relevance of findings
	Quality/validity of findings
	Report is user-friendly / final deliverables are tailored to different audiences
(8) Post-evaluation process	Dissemination/communication strategy is sound and implemented
	Formal system for managers to respond to findings
	Formal follow-up system to verify whether recommendations have been implemented
	Evaluators maintain significant involvement in follow-up activities
	Findings are actively disseminated to users
	Meta-evaluation and/or interagency lesson learning activities take place

Source: UNEG 2016

UN Women strongly encourages the establishment of participatory feedback mechanisms throughout the evaluation process, e.g. through Evaluation Reference Groups. As a result of the decentralised nature of evaluations managed at country level, however, there is limited systematic information on the extent to which staff and partners are actually engaged during the evaluation process. As mentioned above, UN Women introduced evaluation process standards for decentralised evaluations in 2015 only as part of its Professionalisation Initiative.

A recent study by UNEG, which builds on a large body of literature on evaluation use and links it to the UN context, confirms that 'a large share of evaluations are not used, or are not used enough' (UNEG 2016). This applies to evaluations both within and outside the UN system. The UNEG study, based on an extensive literature review, identifies eight individual factors affecting evaluation use, grouped into two categories as outlined in Table 1.

The conceptual framework developed by UNEG in 2016 largely confirms the Theory of Change to strengthen the UN Women evaluation function presented above. It includes important additional elements related to the timing of the evaluation in the policy cycle, stakeholder involvement during the evaluation process and evaluator's qualities, for example. These need to be factored in when discussing potential entry points for enhancing evaluation use.

External assessments by UNEG (2014), MOPAN (2014) and JIU (2014) confirm that whilst systems such as GATE and the Gender Equality Evaluation Portal introduced in UN Women can be seen as positive contributions to ensuring the utility of evaluations, there is room for improving the use of evaluation and dissemination of lessons learned. This includes approaches to further incentivise the use of evaluative information by UN Women senior management at country, regional and global level. An example is the recently (2015) introduced key performance indicator on evaluation use in UN Women: in 2015, 75% of the UN Women field offices reported instances where they had used evidence from recent evaluations to inform their interventions and programming (UN Women Independent Evaluation Office 2016). The UNEG framework described above also underlines the importance of the post-evaluation process, which includes deliberate activities around facilitating the sharing of knowledge from evaluations, e.g. through generating knowledge products and actively disseminating evaluation findings to users.

Both the UN Women Theory of Change and the UNEG framework identify the quality of the evaluation process and product as a key factor for enhancing evaluation utilisation and organisational learning. UN Women, like many international development organisations, has developed evaluation guidance and quality assurance systems with the purpose of enhancing the quality of the evaluation process and product. It is worth recalling that the final evaluation product is the result of an iterative process between evaluation commissioner, evaluator and evaluand. The interplay and often complex relationship between these actors – what de Laat (2013) refers to as 'tricky triangle' – has implications for the independence of the evaluation but also the potential utilisation of evaluation results. Similar to most evaluation functions of UN agencies, UN Women relies heavily on consultants to conduct evaluations. Whilst the GERAAS system provides a benchmark and analytical framework for evaluation quality assurance, the performance of the evaluator remains an important – in most cases external – determinant of evaluation quality (Merkle, Carbon & Addai 2011).

In summary, building evaluation systems and changing the evaluation culture in an organisation is a complex task that takes time. UN Women has put in place structures to strengthen evaluation that, taken together, rank the UN Women evaluation function within the highest cluster along with highest performing evaluation functions of UN entities such as UNICEF, ILO, UNFPA and WFP (Joint Inspection Unit 2014). However, many mechanisms for strengthening the UN Women evaluation function were introduced fairly recently (after 2013) and it may be too early to expect radical improvements across all evaluation performance indicators.

The paper has also illustrated that institutional systems and mechanisms are necessary but not sufficient for nurturing an evaluation culture and ensuring utilisation of evaluation for better development effectiveness. The assumptions identified in the UN Women Theory of Change in particular regarding the commitment to and support of evaluation amongst senior management and existing M&E capacity at office level illustrate the importance of the 'human factor' that will continue to play a key role in institutionalising the UN Women evaluation culture in the Africa region.

Acknowledgements

References

De Laat, B., 2013, 'Evaluator, Evaluand, Evaluation Commissioner: A tricky triangle', in M. Läubli Loud, J. Mayne (eds.), *Enhancing evaluation use, insights from internal evaluation units*, SAGE, London.

Joint Inspection Unit (JIU), 2014, *Analysis of the evaluation function in the United Nations system*, Geneva, viewed 1 April 2015, from https://www.unjiu.org/en/reports-notes/JIU%20Products/Forms/AllItems.aspx

Mayne, J., 2008, *Building an evaluative culture for effective evaluation and results management*, Institutional Learning and Change (ILAC) Brief, November, Rome, 1–4.

Mayne, J., 2013, '*Issues in Enhancing evaluation use*', in M. Läubli Loud, J. Mayne (eds.), *Enhancing evaluation use, Insights from internal evaluation units*, Sage Publications, London.

Merkle, C., Carbon, M. & Addai, E., 2011, 'Enhancing evaluation quality through better consultant management – Experiences from different UN agencies', paper presented at 3rd South African Monitoring and Evaluation Association Conference, Johannesburg, South Africa, 5–9 September.

Multilateral Organisation Performance Assessment Network (MOPAN), 2014, *Synthesis report: United Nations Entity for Gender Equality and the Empowerment of Women,* viewed 1 April 2015, from http://www.mopanonline.org/publications/

UNIFEM Evaluation Office, 2010, *UNIFEM Annual Evaluation report 2009*, New York.

UNIFEM Evaluation Office, 2011, *UNIFEM Annual Evaluation report 2010*, New York.

UN Evaluation Group (UNEG), 2014, *Professional Peer Review of the Evaluation Function of UN Women*, New York, viewed 1 April 2015, from http://www.uneval.org/document/peer-review-reports

UN Evaluation Group (UNEG), 2016, *Evaluation Use in the UN System: Conclusions from the Data*, New York, viewed 4 Sep 2016, from http://uneval.org/document/detail/1911

UN Women Global Evaluation Advisory Committee, 2015, *Report of the Global Evaluation Advisory Committee on the external assessments of the evaluation function of the United Nations Entity for Gender Equality and the Empowerment of Women*, New York, viewed 1 April 2015, from http://papersmart.unmeetings.org/executive-boards/un-women/first-regular-session-2015/documents/

UN Women Independent Evaluation Office, 2012, *Report on the evaluation function of the United Nations Entity for Gender Equality and the Empowerment of Women 2011*, New York, viewed 1 April 2015, from http://www.unwomen.org/en/executive-board/documents

UN Women Independent Evaluation Office, 2013, *Report on the evaluation function of the United Nations Entity for Gender Equality and the Empowerment of Women 2012*, New York, viewed 1 April 2015, from http://www.unwomen.org/en/executive-board/documents

UN Women Independent Evaluation Office, 2014a, *Report on the evaluation function of the United Nations Entity for Gender Equality and the Empowerment of Women 2013*, New York, viewed 1 April 2015, from http://www.unwomen.org/en/executive-board/documents

UN Women Independent Evaluation Office, 2014b, *2014–2017 evaluation strategic plan*, New York, viewed 1 April 2015, from http://www.unwomen.org/~/media/headquarters/attachments/sections/about%20us/evaluation/evaluation-strategicplan-2014-2017-en.ashx

UN Women Independent Evaluation Office, 2015a, *Report on the evaluation function of the United Nations Entity for Gender Equality and the Empowerment of Women 2014*, New York, viewed 1 April 2015, from http://www.unwomen.org/en/executive-board/documents

UN Women Independent Evaluation Office, 2015b, 'Professionalization initiative to strengthen the UN Women Evaluation Function', unpublished guideline developed by UN Women Independent Evaluation Office, New York.

UN Women Independent Evaluation Office, 2016, *Report on the evaluation function of the United Nations Entity for Gender Equality and the Empowerment of Women 2015*, New York, viewed 1 September 2016, from http://www.unwomen.org/en/executive-board/documents

Vogel, I., 2012, 'Review of the use of 'Theory of Change' in international development', report for the UK Department of International Development, viewed 1 September 2016, from https://www.gov.uk/government/news/dfid-research-review-of-the-use-of-theory-of-change-in-international-development