

UN WOMEN GUIDANCE NOTE

Evaluating Impact in Gender Equality and Women's Empowerment

UN WOMEN GUIDANCE NOTE

Evaluating Impact in Gender Equality and Women's Empowerment

AUTHORS:

Courtenay Cabot Venton and Anne Stephens

PRODUCED BY:

UN Women Independent Evaluation and Audit Services (IEAS)

Lisa Sutton, *Director*, IEAS

Inga Sniukaite, *Chief*, Independent Evaluation Service (IES)

MANAGED BY:

Sabrina Evangelista, *Regional Evaluation Specialist*, IES

Caspar Merkle, *Regional Evaluation Specialist*, IES

ACKNOWLEDGMENTS

This publication was made possible with the contribution from the Evaluation Department at the [Norwegian Agency for Development Cooperation \(NORAD\)](#)

DISCLAIMER: *The views and analysis in this publication are those of the Independent Evaluation Service (IES) and do not necessarily reflect the views of UN Women. This is an independent publication by the UN Women Independent Evaluation and Audit Services (IEAS).*

© March 2021, UN Women. All rights reserved.



Acronyms

CSO Civil Society Organization

EVAWG Elimination of Violence Against Women and Girls

ECA Evaluation Capacity Assessment

GEWE Gender Equality and Women's Empowerment

IE Impact Evaluation

IEAS Independent Evaluation and Audit Services

IES Independent Evaluation Service

J-PAL Abdul Latif Jameel Poverty Action Lab

M&E Monitoring and Evaluation

M&R Monitoring and Reporting

NORAD Norwegian Agency for Development Cooperation

OECD-DAC Organisation for Economic Cooperation and Development -
Development Assistance Committee

QCA Qualitative Comparative Analysis

SDG Sustainable Development Goal

SP Strategic Plan

ToC Theory of Change

TP Thematic Priority

UNDAF United Nations Development Assistance Framework

UNDP United Nations Development Programme

UNEG United Nations Evaluation Group

UNICEF United Nations Children's Fund

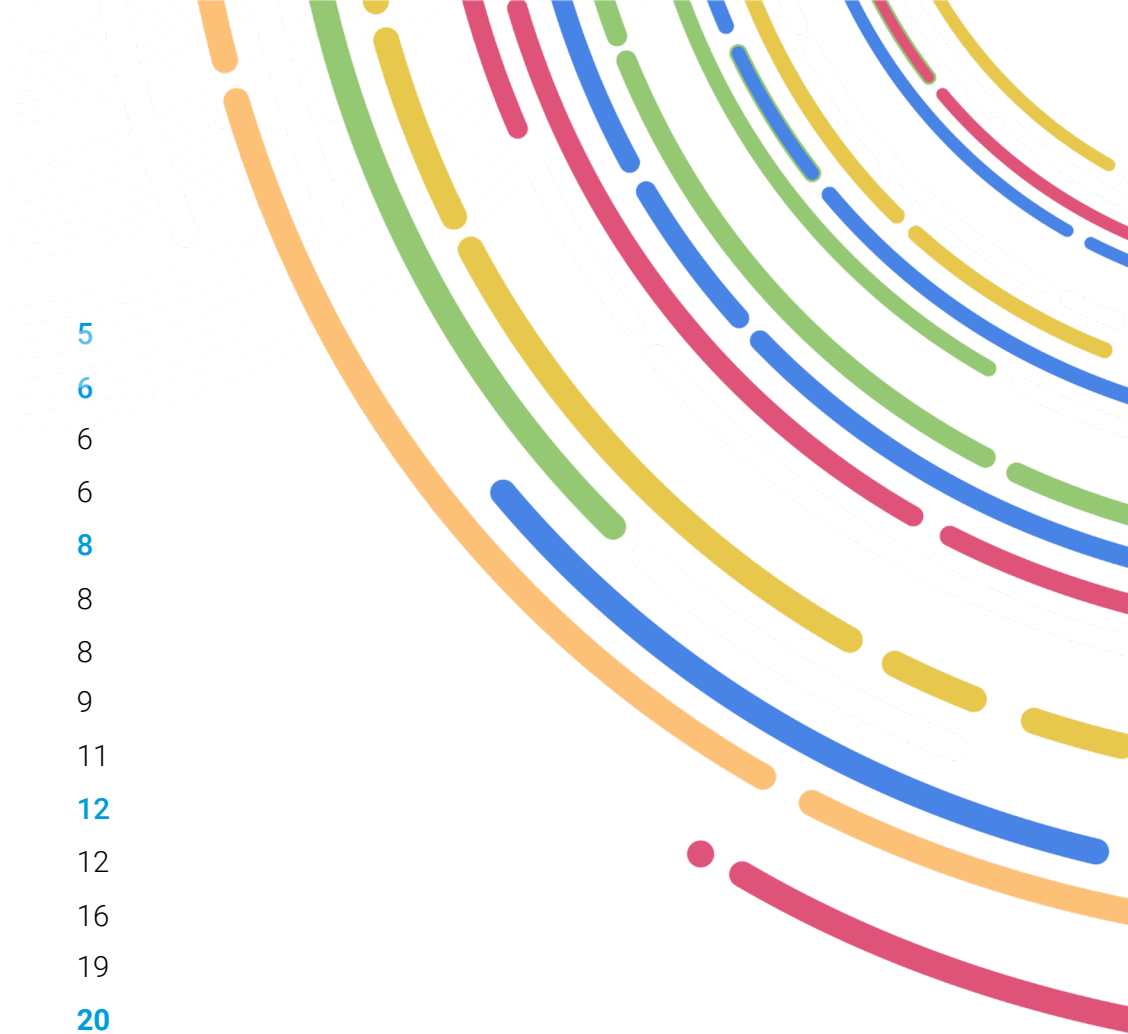
SDCF Sustainable Development Cooperation Framework

VAWG Violence Against Women and Girls

WHO World Health Organization

Contents

1. Background and objectives	5
2. Conceptual background	6
2.1 Purpose and definition of impact evaluation	6
2.2 Conceptual frameworks and challenges for measuring gender equality	6
3. Illustrative Impact Evaluation Framework for ending violence against women	8
3.1 Introduction	8
3.2 UN Women theory of change for EVAW	8
3.3 The Building Block Framework for evaluating impact	9
3.4 Applications of the Building Block Evaluation Framework	11
4. Conducting an impact evaluation	12
4.1 Preparing for impact evaluation	12
4.2 Potential methods for impact evaluation on EVAW	16
4.5 Determining programme success across the Building Blocks	19
ANNEXES	20
Appendix 1 : Sample impact evaluation framework for each building block	21
Appendix 2 : Organizational capacity for evaluation	23
Appendix 3 : Measuring progress using the building block framework	25
Appendix 4 : Methods for impact evaluation and additional resources	27
Literature list	28



1. Background and objectives

The Independent Evaluation Service (IES) of UN Women's Independent Evaluation and Audit Services has produced this impact evaluation guidance note with support from NORAD with the purpose of outlining approaches to measuring impact in gender equality and the empowerment of women, both within UN Women and other organizations. The guidance note is specific to UN Women's context and programming as it uses the integrated mandate as the main guiding framework for its approach to assessing UN Women's normative, coordination and operational work. The intended audience of this Guide is UN Women staff and Evaluation Consultants. Nevertheless, it may be of interest to the broader community of GEWE advocates, programme managers, and evaluators working on how to evaluate the complex area of GEWE.

Impact evaluations in UN Women should be strategically positioned in the overall Strategic Note/ programme cycle and be seen as one out of several evaluation modalities with the specific purpose of assessing long-term change in the lives of women and girls, boys and men. The decision on undertaking an impact evaluation depends on the specific questions that the evaluation is expected to answer and other factors related to the evaluability of the programme. It is recommended to seek expert advice from UN Women IES prior to planning an impact evaluation.

In the context of the new UN Sustainable Development Cooperation Framework (SDCF) options for joint impact evaluations should be explored to the extent possible.

BOX 1: UN Women's integrated mandate*



NORMATIVE

Strengthening normative frameworks for GEWE by supporting the formulation of global, regional and national gender legislation, policies, norms and standards.



COORDINATION

Leading, coordinating and promoting the accountability of the UN system's work on GEWE, including system-wide gender mainstreaming.



OPERATIONAL

Supporting countries to implement commitments through operational activities, i.e. operational programming on the ground.

*UN Women Global 2018–21 Strategic Plan

2. Conceptual background

2.1. Purpose and definition of impact evaluations

The United Nations Evaluation Group (UNEG) 2013 Guidance on Impact Evaluation* clearly outlines the different approaches to impact evaluation and their relative advantages and disadvantages. Most UN agencies, including UN Women, have adopted the latest OECD/ DAC definition of impact: “the extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects.”

[UN Women’s Evaluation Policy](#) articulates the dual role that evaluations can play both in terms of providing evidence-based information about an intervention, as well as “a means to enhance gender equality and the empowerment of women through the incorporation of gender and women’s rights dimensions into evaluation approaches, methods, processes and use. Accordingly, not only does evaluation act as an important driver of positive change towards gender equality and the empowerment of women, but the way in which the evaluation process itself is undertaken empowers the stakeholders involved.”

Complementing this more general guidance, the UNEG Handbook on evaluating normative work** also provides a wealth of relevant resources given the significant focus within UN Women on normative and coordination activities. The UNEG Handbook includes hands-on methodological guidance; concise practical examples; tools for conducting evaluations of normative work; and is highly relevant to the impact evaluation framework presented in this report.

2.2. Conceptual frameworks and challenges for measuring gender equality

The definition and analytical framework for gender equality and the empowerment of women has important implications for how it can be measured (or not) in the context of an impact evaluation. A quick literature review reveals many different definitions of women’s empowerment and, consequently, the analytical frameworks for measuring empowerment diverge. This paper focuses on three frameworks for measuring gender equality and the empowerment of women:

- the Women’s Empowerment Framework developed by Sara Longwe;
- the Women’s Empowerment Framework developed by Oxfam; and
- the Abdul Latif Jameel Poverty Action Lab (J-PAL) empowerment framework.

Box 2 illustrates some degree of overlap between the different gender equality and empowerment of women frameworks. It also demonstrates how the very definition of gender equality and the empowerment of women has implications for potential indicators relating to measurement, data collection, evaluation questions and overall evaluation design.

* UNEG (2013). “Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management.”

** United Nations Evaluation Group (UNEG) (2013) UNEG Handbook for Conducting Evaluations of Normative Work in the UN System.

BOX 2: Gender analytical frameworks



The Women’s Empowerment Framework developed by Sara Longwe*

This framework includes five dimensions of analysis: **welfare** (improvement in socio-economic status); **access** (increased access to resources); **conscientizations** (recognition of structural forces that disadvantage women); **mobilization** (implementing actions related to the conscientization of women); and **control** (control of resources).



The Women’s Empowerment Framework developed by Oxfam**

This recognizes three levels at which change can take place: **personal** (how a woman sees herself); **relational** (power relations in the household and community; and **environmental** (social norms and attitudes, political and legislative framework).



The Abdul Latif Jameel Poverty Action Lab (J-PAL) empowerment framework***

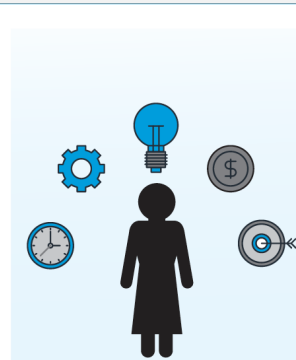
This suggests using the three interrelated dimensions of **resources** (material, human, social); **agency** (participation, voice and influence); and **achievements** (meaningful improvements in well-being and life outcomes) for measuring empowerment.

*As referenced in The UN Women Country Portfolio Evaluation (CPE) Guidance.
** How To Guide to Measuring Women’s Empowerment: Sharing experience from Oxfam’s impact evaluations. ***Glennerster, R, C Walsh, and L Diaz-Martin (n.d.). “A Practical Guide to Measuring Women’s and Girl’s Empowerment in Impact Evaluations.” J-PAL.

The J-PAL Guide provides a useful overview of some of the challenges in measuring women and girl's empowerment, summarized in Box 3.

A wide variety of approaches to overcome these challenges exist, including triangulating data; using panel data; and using a formative research design and strategy, etc. The ethical implications of any work must also be considered and carefully designed, particularly when working with vulnerable women and girls, ensuring that impact evaluation design considers unequal power dynamics.

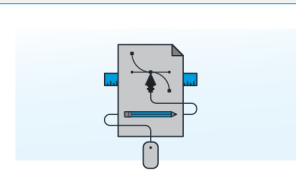
BOX 3: J-PAL Guide - Challenges in measuring women and girl's empowerment



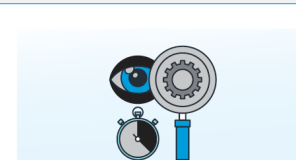
Measuring people's ability to make important life choices

is challenging because we rarely observe decision-making directly.

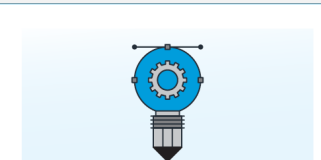
Empowerment is not just about changes in people's well-being, but also people's agency in achieving these changes, and this can be hard to track and measure.



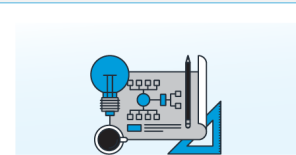
Empowerment is a process, and it takes effort and creativity to measure it well.



Many aspects of empowerment are susceptible to **reporting bias**, and it can take a good deal of time and effort to design tools that can capture empowerment while mitigating for bias.



Empowerment means different things, in different contexts. This means that **indicators need to be locally tailored**, which can hamper comparability.

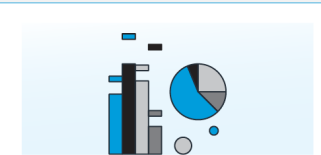


Prioritizing outcome measures

is difficult. Empowerment spans across every aspect of a woman's life, and it can be difficult to choose the most relevant outcomes to measure.



Measuring women's preferences is challenging in contexts where women have **internalized society's views**.



Disempowerment can heighten **data collection challenges** – when women lack power and voice, it can be difficult to collect data about their aspirations and opinions.

3. Illustrative Impact Evaluation Framework for ending violence against women

3.1. Introduction

The design of any impact evaluation begins with an articulation of the theory of change. A theory of change helps to establish causal pathways and explain how activities are understood to produce a set of inter-related result assumptions and context that contribute to achieving outputs, outcomes and the final intended impact. UN Women has developed theories of change for all 12 of its Thematic Priorities.* The following section introduces the impact evaluation framework and applies it to the ending violence against women thematic area of work for illustrative purposes. The impact evaluation framework is expected to be adapted as appropriate based on the thematic area and programme.

3.2. UN Women theory of change for ending violence against women (EVAW)

As identified in the theory of change developed by UN Women**, the goal for Thematic Priority 8 is “Violence against women and girls is prevented and quality essential services are available and accessible and to victims and survivors.” Thematic Priority 8 is supported by three outcomes and ten outputs as summarized in Table 1.

TABLE 1: ORIGINAL THEORY OF CHANGE FOR THEMATIC PRIORITY

GOAL: Violence against women and girls is prevented and quality essential services are available and accessible to victims and survivors		
OUTCOME 1: An enabling legislation and policy environment in line with international standards on ending violence against women and girls (EVAWG) and other forms of discrimination is in place and translated into action.	OUTCOME 2: Favourable social norms, attitudes and behaviours are promoted at community and individual levels to prevent violence against women and girls.	OUTCOME 3: Women and girls who experience violence can use available, accessible and quality essential services so the impacts of violence are addressed, and perpetrators are held accountable.
Output 1.1 Women’s voice and agency strengthened to advocate for the development and implementation of laws and policies on EVAWG.	Output 2.1 Prevention strategies and action plans in line with international human rights standards developed and implemented by the UN and national actors in a coordinated manner.	Output 3.1 Capacity of service providers to adapt and integrate global standards and guidelines for essential services into service delivery is enhanced.
Output 1.2 Laws and policies are regularly reviewed, reformed and resourced to conform with international human rights standards and evidence.	Output 2.2 Women, girls, men and boys at community and individual levels are mobilized in favour of respectful relationships and gender equality.	Output 3.2 Capacity of service providers to deliver quality, coordinated essential services; to hold perpetrators to account; and the collection and use of data in an ethical manner is strengthened.
Output 1.3 Capacity of institutions to implement legislation on EVAWG and other forms of discrimination strengthened.	Output 2.3 Educational curricula and programmes addressing gender equality and violence against women and girls are developed and integrated into formal and non-formal education.	Output 3.3 Availability and accessibility of quality essential services for women and girls who have been subject to violence has improved and, in the case of gender-related killings, support to victims’ families is strengthened.
		Output 3.4 Women understand and can exercise their rights to quality essential services.

*UN Women (2017) Theories of Change for UN Women’s Thematic Priorities: Achieving Transformative Results for Gender Equality and Women’s Empowerment.

**Ibid

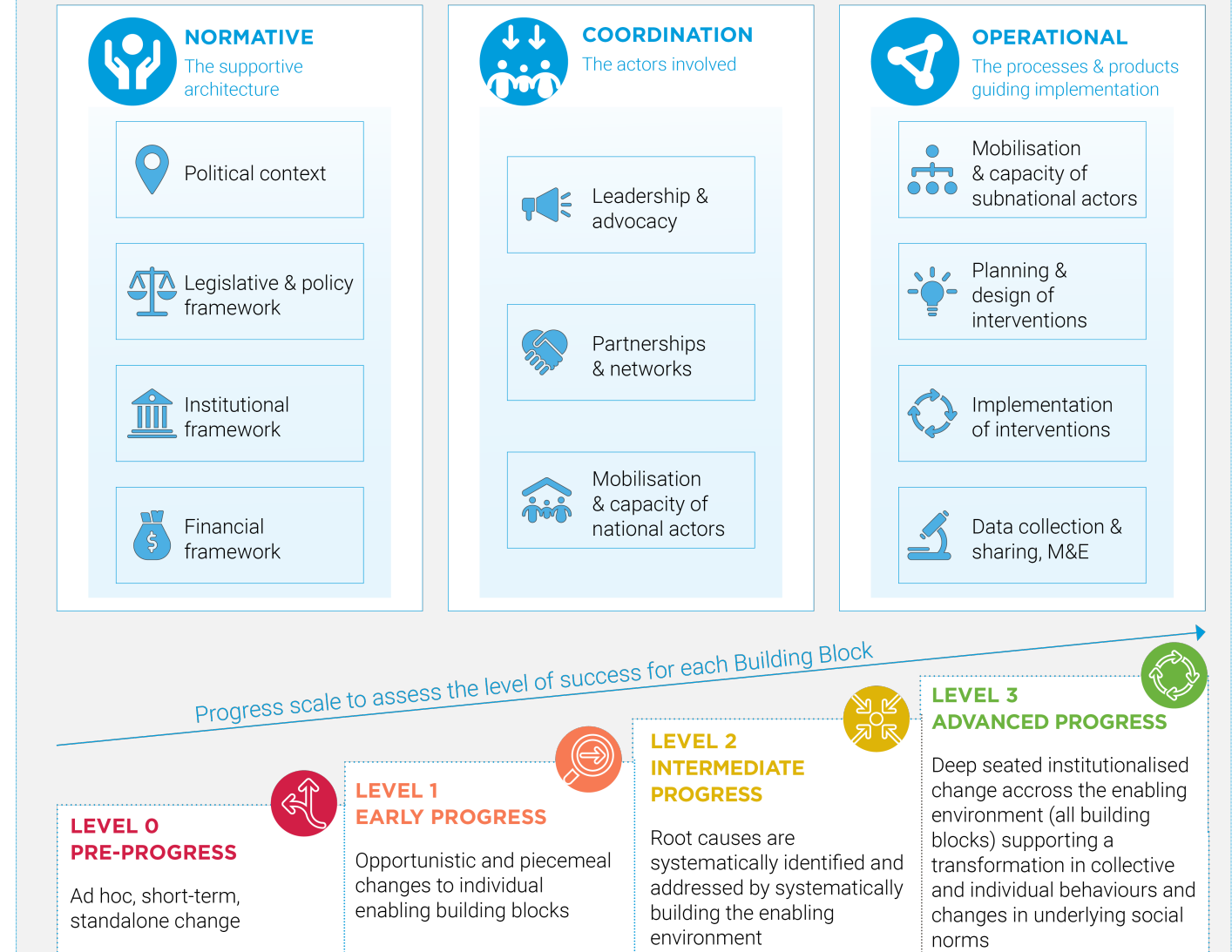
The analysis of the existing theory of change for ending violence against women girls reveals several limitations: it insufficiently reflects UN Women's triple integrated mandate and other complexities, including the Entity's contributions to multiple stakeholders' results. Therefore, this guide proposes to reconfigure the existing theory of change for Thematic Priority 8 to better capture UN Women's triple mandate with three pillars of inter-related work. The proposed framework also identifies potential enabling factors, as well as barriers to change, and helps unpack the "how" and "why" a change may or may not have happened.

3.3. The Building Block Framework for evaluating impact

The starting point for the Building Block Framework* is to (a) re-align the existing theory of change for ending violence against women and girls (Thematic Priority 8) against the three pillars of UN Women's work; and (b) identify the enabling factors or building blocks (and by default the barriers) for change.

Figure 2 on following page provides an overview of the Building Block Framework applied to Thematic Priority 8 with a high-level summary of respective impact pathways for UN Women's normative, coordination and operational work. It is not intended to describe a linear process, but rather summarizes the main categories of impact along the Entity's integrated mandate.

FIGURE 1: The building blocks enabling change across the three pillars of UN Women's work



* This analysis was informed by a UNDP Risk Governance Building Block Framework used in the Pacific for enabling transformation of the development agenda in support of resilient development. A review of the risk governance building blocks identified 10 that are enshrined in UN Women's Strategic Plan. One additional building block – an enabling political framework (e.g. political commitment, political will, champions) was highlighted as a key enabling factor in UN Women's Strategic Plan.

FIGURE 2: Summary of the Building Block Framework for TP 8













BUILDING BLOCKS	ACTIVITIES Work done with inputs to produce specific outputs	OUTPUTS Changes or new products and services resulting from the completion of activities	OUTCOMES Medium-term changes resulting from collective efforts	EVIDENCE OF CHANGE	IMPACT Long-term effects
Political, legislative, policy, institutional & financial frameworks	 Technical assistance to strengthen the political, legal/policy, institutional & financial frameworks	 Laws & policies reformed (1.2); prevention strategies developed (2.1), institutional capacity to implement (1.3), service provider capacity to integrate standards (3.1)	 Political commitment; integration into decision making & codes of conduct; institutional arrangements & budget allocations established	 Political, legal, policy, institutional & financial frameworks mainstream EVAWG	Enabling environment supports VAWG prevention & responsiveness
Leadership, advocacy, partnerships networks, mobilisation & capacity	 Leadership, advocacy, brokering partnerships & networks, and mobilising national actors	 Coordinated implementation of action plans (2.1); women's agency & voice strengthened (1.1)	 Increase in collaboration, coordination collective action, dialogue, engagement, awareness and capacities of national actors	 Multi-sector stakeholder partnerships, networks, decision making & action plans for EVAWG	Leadership, collaboration & capacities of national actors ensures coordinated prevention & response
Mobilisation, capacity, planning, design, implementation, data, information, M&E	 Mobilising & developing capacity of women's organisations, communities, girls and boys; training to service providers; development of curricula; awareness raising & outreach on service access	 Women's agency & voice strengthened (1.1) mainstreaming into education (2.3), community mobilised (2.2), women understand/exercise rights (3.4), delivery & availability of essential services (3.2 & 3.3)	 Increase in awareness, understanding, knowledge, participation demand for services; and change in provision & use of services	 Use of knowledge & skills, action to prevent VAWG, changes in use / uptake of services, quality of service provision, & implementation of EVAWG interventions	Changes in behaviours & practices address root causes & impacts of violence

Table 1 describes each of the outputs aligned with this figure

Figure 2 provides an overview of the Building Block Framework applied to EVAW with a high level summary of impact pathways. Appendix 1 contains the full Building Block Framework for ending violence against women and girls. It shows how the three areas of UN Women's work are mutually supportive, and together build the enabling environment for change by putting in place individual building blocks. It is important to consider other relevant frameworks and how they align, for example in the case of this framework elaborated in Appendix 1, alignment with the RESPECT framework* has been considered.

The table in Appendix 1 identifies individual change pathways from building blocks through outcomes to impact:

- **Building blocks** – aligned with the relevant integrated mandate category.
- **EVAW activities** – taken from UN Women's theory of change for EVAW.
- **EVAW outputs** – associated with each activity. This is useful as a visual tool to see how many times different outputs may include activities that cut across a range of building blocks and normative/coordination/operational work.
- **Success criteria/ indicators** – for each building block
- **Outcomes** – institutional and behavioural changes that occur between the completion of outputs and the achievement of goals, usually requiring the collective effort of partners.
- **Evidence of change** – for the building block pathways.
- **Impacts** – positive and negative long-term effects on population produced by a development intervention, directly or indirectly, intended or unintended.

*UN Women, RESPECT Women: Preventing violence against women implementation package (2020): <https://www.unwomen.org/en/digital-library/publications/2020/07/respect-women-implementation-package>

The Building Block Framework can improve programme planning, effectiveness and impact through identifying entry points and potential gaps in a particular intervention. It can highlight enabling factors that are not being targeted through current work on ending violence against women and girls.

3.4. Applications of the Building Block Evaluation Framework

The reframing of the initial theory of change for ending violence against women into the Building Block Evaluation Framework could potentially be applied as follows:

It facilitates impact evaluation across all three pillars of UN Women's work. The Framework accommodates inter-relationships between normative, operational and coordination work. For example, the Framework allows an assessment of how strengthening the institutional framework building block (normative work) results in changes in institutional practice and service delivery (operational work).

The Framework helps standardize impact evaluation across thematic priorities, countries and interventions. The Framework allows comparison across countries and application of the building blocks to different thematic priorities and different contexts. For example, using a standardized framework for evaluating impact in ending violence against women in different countries, potentially highlights whether certain enabling factors feature more prominently across countries; whether contextual differences are relevant; or whether implementation failure or theory failure play a role. It also provides a framework for collecting baseline data in different contexts and for enabling evidence synthesis such as meta-analysis that can draw conclusions across individual impact evaluations.

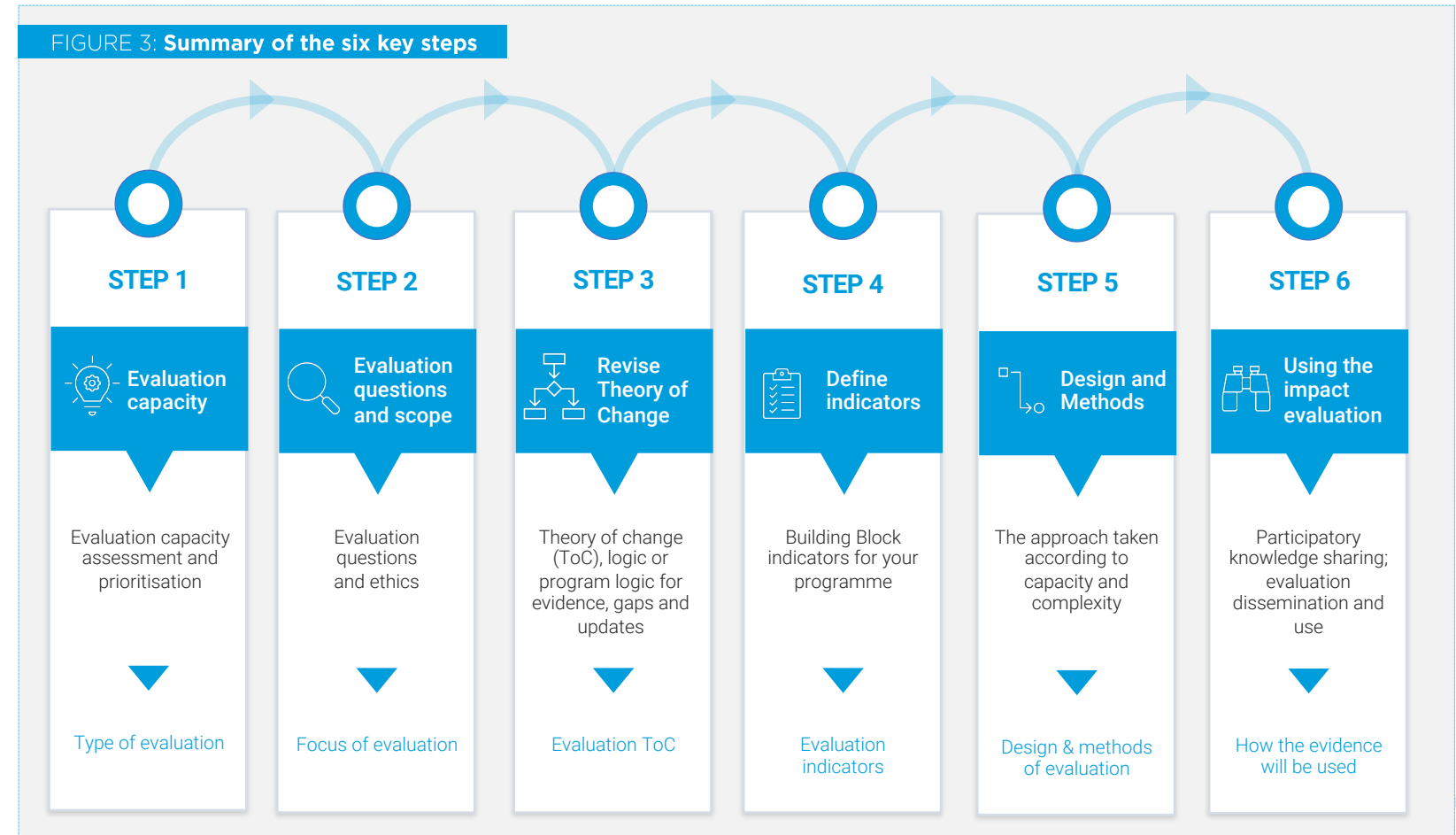
The Framework helps identify UN Women's contribution to he specific thematic area. The analysis of building blocks and the identification of outputs and outcomes linked to each of the enabling factors helps identify the specific contribution that UN Women has made to the specific thematic area. This is critical for any articulation of UN Women's impact, as progress might be hindered due to a lack of action across stakeholders, and/or because UN Women initiatives were not effective. This is particularly important in the context of measuring UN Women's contribution to the Sustainable Development Goals (SDGs). The SDGs are a multi-stakeholder framework that can only be achieved via a combination of normative, coordination and operational work across stakeholders. The Building Block Framework provides a broader, system-wide analytical tool that can help to identify which enabling factors are inhibiting progress towards achieving SDG 5. For example, UN Women might provide normative support to develop standards and guidelines to ensure women have access to a range of key services; however, without operational work by other agencies, e.g. community mobilization around ending violence against women, uptake of these services will be limited.

In addition to impact evaluation, the Building Block Framework can improve programme planning, effectiveness and impact through identifying entry points and potential gaps in a particular intervention. It can highlight enabling factors that are not being targeted through current work on ending violence against women and girls.

4. Conducting an impact evaluation

4.1. Preparing for impact evaluation

Impact evaluation requires planning, resourcing, and the establishment of systems at the programme formulation stage. It involves several preparatory steps (outlined in this section) to facilitate the testing of the Framework presented above in a potential impact evaluation at country level. This section is not intended to provide detailed methodological impact evaluation guidance as the UNEG guidance* and other sources can be referred to. The key steps are described below and summarized in Figure 3. Detailed guidance on the evaluation process is available in the UN Women [Handbook on Managing Gender Responsive Evaluation](#).



*UNEG (2013). "Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management."

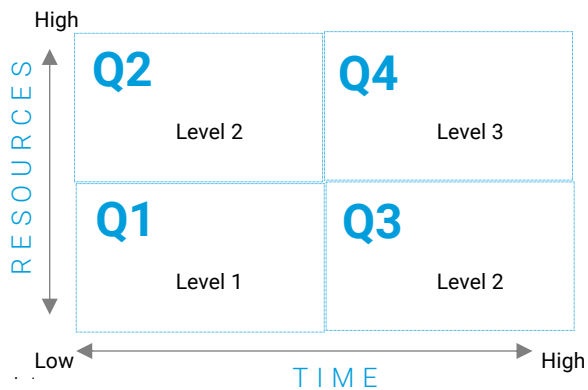


STEP 1. Evaluation capacity, assessment and prioritisation

The eight parameters for prioritizing the interventions to be considered for impact evaluation, as outlined in [UN Women's Evaluation Handbook](#), can be used here: i) relevance of the subject; ii) risk associated with the intervention; iii) significant investment; iv) demands for accountability from stakeholders; v) potential for replication and scale-up; vi) potential for joint or United Nations Development Assistance Framework (UNDAF) Evaluation; vii) feasibility for implementing evaluation; and vi) knowledge gaps.

Step 1 should also include an evaluation capacity assessment to determine whether an impact evaluation is appropriate given the time required, internal knowledge and expertise and funding. Appendix 2 "Evaluation Capacity Assessment (ECA)" Tool provides detailed guidance on the steps involved. The tool is designed to support the self-assessment of organizational capacity to manage an impact evaluation. Below is a quick overview of the tool.

The graph below shows four quadrants. Each quadrant has a high/low scale on the vertical and horizontal axes. The vertical axis is RESOURCES. The horizontal axis is TIME. Answer each question and rate your response against both axes. Considering both the resources and time, and place an X anywhere in the quadrant you think most accurately reflects your position.



QUESTIONS

1

LENGTH OF PROGRAMME
How long is the programme of work funded for?
1 year = Low
2 to 3 years= Medium
4 to 5 years= High

2

FUNDING FOR M&E ACTIVITIES
What % is allocated to evaluation?
UN Women evaluation policy recommends 2-3% of programme expenditure for evaluation; and 8-10% for monitoring)

1% ← → 10%
Low High

3

RESOURCES FOR EXTERNAL CONSULTANT
Are resources sufficiently high to fund an external evaluator for baseline, mid-term and a summative Impact Evaluation?

No evaluation work can be undertaken = low resources and time
Programme evaluation at end-point only = high resource / low time
Yes, but selective about baseline or midterm = low resource / high time
Yes, resources and time is adequate = high resources and time

4

M&R SPECIALIST
Does the CO have a M&R specialist for the duration of the programme of work?
0 Full-time equivalent staff = Low
2 Full-time equivalent staff = High
Do you have the IE expertise to manage the programme of work at the country office? (If no, the resource axis is low). Does the programme of work require more than one person in a full-time capacity to manage stakeholders and participants? Are the timeframes of the programme of work able to accommodate recruitment and procurement processes to secure country-level expertise?

5

EVALUABLE PROGRESS
(relationships with stakeholders, donors, political sensitivities, budget, execution of programme on track, evaluation plans, monitoring systems)
Not on track, budget is limited = low / low
Commenced (on time), but key deliverables cannot be met due to unforeseen circumstances (preventing field work, stakeholder engagement low, participant recruitment etc.) = high / low
On track, on budget and meeting all project milestones = high / high

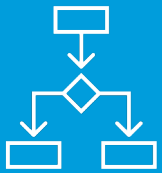


STEP 2. Evaluation questions and scope

The next step is to establish the key evaluation question(s) that will be answered during the impact assessment. A set of focused, key questions are central to any evaluation, which help to provide focus and structure. Importantly, the choice of design and methods will flow out of these overarching evaluation questions, which might include:

- Did the intervention make a difference?
- What was the intervention's contribution?
- Could the intervention be replicated?

This step also includes the consideration of ethical issues to ensure the rights of individuals involved in an evaluation are respected. For details please refer to the [UNEG Ethical Guidelines for Evaluation](#) in the UN system. When working with women and children in EAWG evaluation, the [WHO \(2016\) Ethical and safety recommendations for intervention research on violence against women](#) should be consulted.



STEP 3. Revise Theory of Change

Most interventions will only contain some of the activities outlined in the Building Block Framework. Therefore, the theory of change should be revised and validated to reflect the specific activities and context of the programme being evaluated.

At this stage it is important to recall that change often occurs in a non-linear, emergent or iterative manner. The ToC needs to unpack the underlying assumptions accordingly.



STEP 4. Define indicators

This step involves developing indicators in the revised theory of change for programme-specific activities. These indicators should include short, medium and long-term indicators of success that can be used to test impact along a pathway.

Sources for collecting qualitative and quantitative data* on indicators may include official national statistics, programme monitoring data and records, (repeat) baseline surveys etc.

By reviewing the available information, gaps that need to be addressed by collecting new data can be identified. Then tools can be designed to gather the necessary information for measuring the selected indicators.

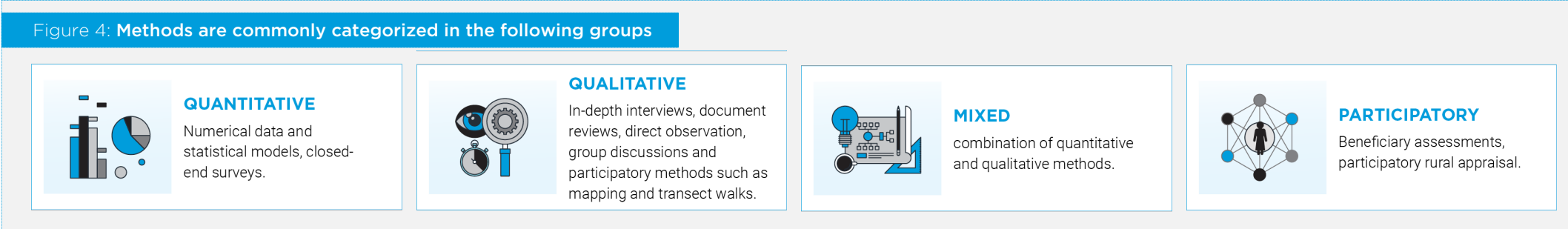
*https://www.unicef-irc.org/publications/pdf/brief_10_data_collection_analysis_eng.pdf

STEP 5. Evaluation design and methods

This step involves identifying the most appropriate evaluation design and methods suitable for tracking impact, see the following section and Table 2 below for more details. As outlined in UNEG guidance, evaluation designs are usually categorized as experimental, quasi-experimental and non-experimental. There are advantages and disadvantages to each evaluation design as explained in the UNEG guidance and other sources (for example, UNICEF’s 2014 methodological brief on quasi-experimental

design,* or the BOND guide to impact evaluation**). Selection of methods need to be both appropriate for the context, fit for purpose and ethical (See Figure 4 below and the next section for more details).

*White, H and S Agarwal (2014). "Quasi-Experimental Design and Methods". UNICEF Office of Research.
**BOND (2015). "Impact Evaluation: A Guide for Commissioners and Managers".



STEP 6. Using the Impact Evaluation

The impact evaluation findings should be validated in a participatory approach with stakeholders and programme partners. Engaging participants in interpreting evaluation results is an important step towards stakeholders understanding the outcome and impact, and committing to the interventions ongoing successes, working to make improvements, or discontinuing practices that are ultimately found to not work or may even cause harm. *

It is essential to widely disseminate the impact evaluation so that it contributes to the global body of knowledge on what GEWE interventions work and in what context. Knowledge transfer by hosting events is important to enable ongoing capacity development and knowledge sharing. The use of short videos may be used to present information in easily digestible formats that can be disseminated broadly. For all dissemination efforts, it is important to consider the audience and the purpose of sharing the information so that it can be tailored appropriately.

*Participatory Evaluation': https://www.betterevaluation.org/en/plan/approach/participatory_evaluation#:~:text=Participatory%20evaluation%20is%20an%20approach,the%20reporting%20of%20the%20study

4.2. Potential methods for impact evaluation on ending violence against women

Specific methodological approaches will be relevant to answer different evaluation questions. The overarching recommendation – also aligned with the UNEG guidance – is that a mixture of methods be used to ascertain impact. This is particularly relevant for UN Women’s work, where the interplay between normative, coordination and operational work is key. The use of multiple lines of enquiry and multiple data sources is advised to allow for different parts of the causal chain to be understood.

It is important to recall that the UN Women Evaluation Policy highlights the role of evaluation as a tool to empower the stakeholders involved, in line with feminist or gender-responsive approaches to evaluation. Experimental designs based on random assignments in intervention versus control groups may treat rights holders as objects of research rather than empowered individuals who claim their rights to engage in defining how change in their own lives is defined and measured. Experimental designs do not necessarily represent the rights of those that are meant to benefit from development programming.

The use of quasi-experimental or non-experimental designs and a mixed methods approach allows more space to incorporate the voices of women within the evaluation and to ensure that the process itself is empowering.

Multi-site locations

Multi-site studies can provide a good opportunity for making comparisons between implementation sites. Testing the intervention in diverse settings can address important questions about the initial situation and how it has changed in relation to differing contexts. However, multi-site studies require data collection systems and baselines that allow for a robust comparative analysis.

Frequency of data collection

A baseline study is a critical element in assessing change over time. Baseline data should be collected prior to the start of the programme or during the inception phase. It is important to advocate for resources to dedicate to a robust baseline study. Baseline studies should include information to allow tracing of respondents for mid-term and end-term studies to ensure that data is accurately linked across the chain of outcomes. If a baseline study has not been conducted or is not available, baseline conditions can be constructed retroactively at the risk of a much weaker evaluation design. Mid-term process evaluation may be advantageous to assess project conduct and to make corrections and adjustments prior to a final impact evaluation. Baseline data can be reviewed and monitoring systems refined. However, it is important to consider the time between baseline – mid-term and impact evaluation because if the time is too short it will not allow for demonstrable change to be picked up and it may not be a worthwhile investment.

TABLE 2: EVALUATION DESIGN AND METHODS FOR UN WOMEN PILLARS OF WORK

	DESIGN (Non-experimental, quasi-experimental or experimental)	METHODS
NORMATIVE	All three types can be used for normative work, though most impact evaluations use non-experimental design due to the level of rigour and attribution required for experimental/quasi-experimental approaches.	Mixed methods (qualitative and quantitative) are strongly recommended.
COORDINATION	All three types can be used for coordination work, though the majority of impact evaluations use non-experimental design due to the level of rigour and attribution required for experimental/quasi-experimental approaches.	Qualitative data complements quantitative data and is necessary to understand the narratives of women who are affected by violence and tease out the ‘how’ and ‘why’ change is happening.
OPERATIONAL	Quasi-experimental can be useful to determine attribution and causality where affordable and practical. However, for violence against women and girls programming specifically, such an approach may not be ethical, and non-experimental approaches are more appropriate.	A variety of approaches can help reveal causality and provide opportunities for triangulation.

TABLE 3: UNICEF METHODOLOGICAL BRIEF - THREE MAIN SAMPLING TECHNIQUES

CLUSTER OF SAMPLING OPTIONS	WHAT ARE SOME SPECIFIC METHODS?	RISK OF INTRODUCING BIAS
Probability: Use random or quasi-random methods to select the sample, and then use statistical generalization to draw inferences about that population	<ul style="list-style-type: none"> • Simple random • Stratified random • Multi-stage • Sequential 	This cluster has specific rules about selection of the sampling frame, the size of the sample and managing variation within the sample
Purposive: Study information- rich cases from a given population to make analytical inferences about the population. Units are selected based on one or more predetermined characteristics and the sample size can be as small as one	<ul style="list-style-type: none"> • Confirming and disconfirming • Critical case • Maximum variation • Outlier • Snowball • Theory-based • Typical case 	This cluster encourages transparency in case selection, triangulation, and seeking out of disconfirming evidence
Convenience: These sampling options use individuals who are available or cases as they occur	<ul style="list-style-type: none"> • Readily available • Volunteers 	This type has the lowest credibility but requires less time, money and effort

Source: Peersman, G., (2014). Methodological Brief No.10: Overview: Data Collection and Analysis Methods in Impact Evaluation. United Nations Children's Fund (UNICEF). https://www.unicef-irc.org/publications/pdf/brief_10_data_collection_analysis_eng.pdf

Sampling

For any impact evaluation it is important to remember strengths and weaknesses of potential sampling strategies used. Table 3 summarizes three main sampling techniques.

Unforeseen or emergent impacts

Emergent outcomes (planned or unplanned) are an important consideration to be mindful of when planning for and during an impact evaluation. Methods that enable responsive adaptation are preferable to enable evaluators to track and trace new emerging pathways, or tangents, that may not have been considered previously. Emergent outcomes need to be considered against the programme ToC or logic model, as they provide new insights into knowledge gaps, reinforcing and positive impacts of the theory, or negative impacts that need to be addressed.

Evaluation methods

The UNEG Guidance Notes on Impact Evaluation and on Evaluating Normative Work describe several evaluation methods, of which three approaches to measuring complex causal chains are described in Boxes 4–7 below.

Box 4: Value for money analysis

Increasingly, questions concerning the cost and return of funding social investment are being asked. Funders would like to know is the cost justifiable for the level of impact?

A value for money analysis can be conducted prior to a programme being delivered or after it has been executed to assess likely or actual impact against cost. Expert guidance is recommended for the design and conduct of value for money analysis.

Box 5: Outcome Harvesting

“Unlike some evaluation methods, Outcome Harvesting does not measure progress towards predetermined outcomes or objectives, but rather collects evidence of what has been achieved, and works backward to determine whether and how the project or intervention contributed to the change.”

Outcome Harvesting works best when the focus is on outcomes, rather than activities, particularly in complex programming environments where a range of activities could be intersecting to create a specific outcome.

For example, in the case of ending violence against women and girls, the outcome would be an observed change in violence against women and girls. Outcome Harvesting would then be used to reverse the logic of a traditional impact evaluation, and instead of evaluating the changes brought about by a specific activity, it would look at observed changes in the lives of women and girls and trace these back, using the Building Block Framework presented in this document, to activities and therefore enabling factors or building blocks that have contributed, indirectly or directly, partially or entirely, to the outcome in question. Importantly, Outcome Harvesting seeks to establish the causal pathway to understand what was achieved and how it was achieved. When aligned with the Impact Evaluation Framework for ending violence against women, this process can highlight where certain building blocks have been instrumental in creating change, and also where a lack of attention to some building blocks may be hindering progress.

Wilson-Grau, R. and H Britt (2012) Outcome Harvesting, Ford Foundation.

Box 6: Process Tracing

Process Tracing adopts a case-based approach and uses causal inference to establish linkages in a theory of change. It has been used within political science, comparative politics, organizational studies and international relations. Process Tracing begins with the evaluator identifying the intended outcomes of a programme of work, and the activities intended to support the achievement of the outcome.

Process Tracing is designed to understand a range of explanations, or pathways, by which a change came about.

Process Tracing is one of the key approaches of Oxfam’s Women’s Empowerment Measurement Framework. It is used to assess interventions which often work to achieve specific intermediary and final outcomes where there are too few units of assignment to permit tests of statistical significance between intervention and comparison groups. These might, for example, include increased ability to participate meaningfully in governance processes; changes in the practices of target state institutions; or policy change.

Oxfam’s guidance document highlights the following key steps:

- a. **Construct a theory of change** for the intervention to determine what it is trying to change (outcomes); how it is working to effect these changes (activities); and what assumptions it is making about how it will contribute.
- b. **Identify up to three intermediate or final outcomes** considered to be the most critical.
- c. **Systematically assess the activities** undertaken to achieve the outcomes.
- d. **Identify from the evidence**, the extent to which the selected outcomes have actually materialized.
- e. **Identify plausible causal explanations** for the evidenced outcomes.
- f. **Gather data** to assess the degree to which each of the explanations identified in the previous step are (or are not) supported by the available evidence.

Box 7: Qualitative Comparative Analysis (QCA)

“Developed in the 1970s, QCA enables the analysis of multiple cases in complex situations, using a case-based approach. It uses both quantitative and qualitative analysis, and is based on two assumptions: i) that change can be the result of different combinations of factors (rather than a more linear model where change is linked to one individual factor); and ii) that different combinations of factors can produce similar changes.

QCA includes the following steps:

- Develop a Theory of Change.
- Identify a set of cases to be analysed – ideally some where the outcome was realized and others where it was not.
- Develop a set of factors/conditions whose presence or absence may contribute to the outcome in question.
- Score each of the factors depending on whether or not it has been achieved.
- Analyse the dataset of scored factors to look for patterns (computer software packages have been developed that allow for more rigorous evaluation of patterns in the data).
- Interpret the findings to determine the factors that are most frequently present when the outcome is achieved

4.3. Determining programme success across the Building Blocks

Success criteria can be provided for each building block at all levels of change (e.g. output, outcome, impact). Criteria can be shaped by participants and stakeholders. It can be used to identify progress, from a basic to advanced level.

The Building Block Progress Tool (see below) has been developed to highlight where there may be blockages in the system that may be impeding – or accelerating – progress on specific activities. It provides a practical qualitative tool to assess the degree to which UN Women’s activities across the triple mandate are contributing to long-term change in the lives of women and girls, boys and men. The tool should be customized for each project in a participatory approach through adapting success criteria / indicators for each building block indicating early pre-progress, early progress, intermediate progress or advanced progress as illustrated below. Each progress stage is scored and the sum of scores across all work areas (normative, cooperation, operational) provides the overall assessment of the programme.

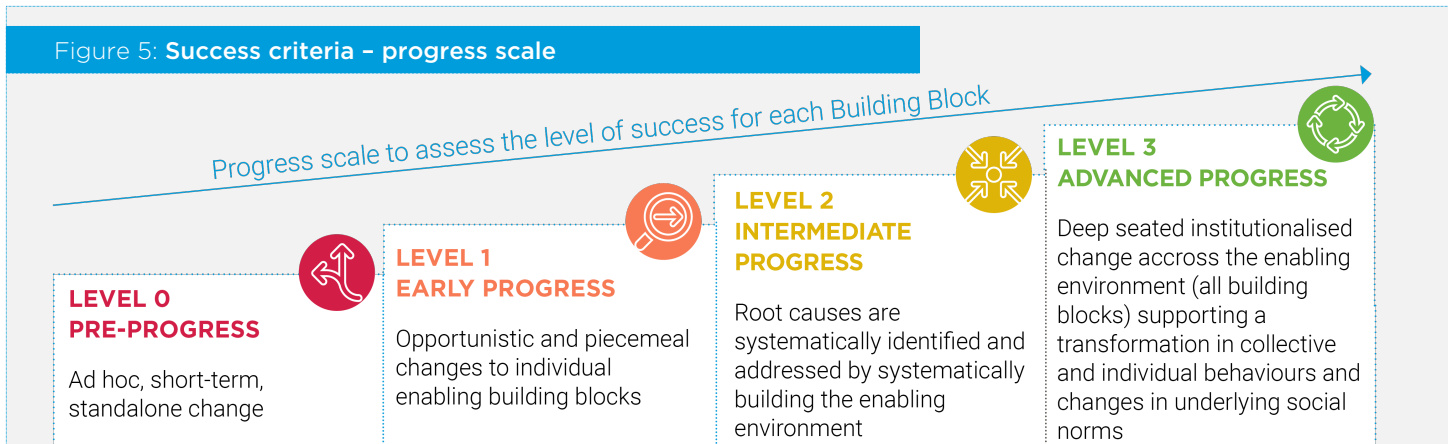
Box 8: Impact Evaluation of Lebanon Women’s Access to Justice, Oxfam GB*

IE of activities implemented by the Lebanese Council to Resist Violence Against Women (LECORVAW), as part of a multi-country project Women’s Access to Justice Programme: Phase 2. The project used a mix of normative, coordination and operational activities. It aimed to improve access to justice.








Column 1 lists the Building Blocks. Column 2 provides a short summary statement of corresponding activities and briefly describes the methods used to collect data. And column 3 presents key findings of the combined activities. The example demonstrates the use of mixed methods to enable an evaluation team to probe and understand the interconnectedness of activities, across the building blocks that influence outcomes.

*Lombardini, S., Garwood, R., & Hassnain H. (2019). Women’s Empowerment in Lebanon: Impact Evaluation of the project. Oxfam.








	ACTIVITIES (Methods)	FINDINGS
NORMATIVE	Advocacy with policy makers (Methods: life history interviews analyzed using Qualitative Comparative Analysis (QSA) and Predictive Analytics (PA) to identify why women start a litigation)	<ul style="list-style-type: none"> Women in the intervention scored more highly on women’s empowerment indicators with three key factors identified as key building blocks: i) awareness raising; ii) self-confidence; and iii) confidence in the Lebanese Council
COORDINATION	Coordination of women’s rights organizations, legal and policy stakeholders (Methods: life history interviews analyzed using Qualitative Comparative Analysis (QSA) and Predictive Analytics (PA) to identify why women start a litigation)	<ul style="list-style-type: none"> The project had a positive and significant effect in increasing trust in reporting problems to the police and organizations supporting women, and towards seeking legal advice
OPERATIONAL	Peer-to-peer support groups; training paralegals; community awareness raising (Methods: quasi-experimental design to assess the impact of awareness raising sessions; life history interviews; development of women’s empowerment index used to measure impact)	<ul style="list-style-type: none"> For the women who did take legal action, they found no difference in satisfaction with the result of the legal outcome/verdict, or on women’s perception of fairness of the courts. However, the intervention had a significant positive effect on women’s experience of the legal process.










APPENDIX 1 : SAMPLE IMPACT EVALUATION FRAMEWORK FOR EACH BUILDING BLOCK (1/3)

 BUILDING BLOCKS	 ACTIVITIES	 OUTPUTS	 OUTCOME STATEMENT	 EVIDENCE OF CHANGE	 GUIDANCE NOTES	
UN Women has a triple mandate with three pillars of inter-related work: Normative, coordination and operational	Activities are informed by the TP8 ToC, and other frameworks including the RESPECT Framework for prevention and the Essential Services Framework for response	Changes in skills or abilities, or new products and services that result from the completion of activities	Institutional and behavioral changes that occur between the completion of outputs and the achievement of goals, usually requiring the collective effort of partners.	Indicators are measurable information used to determine if a programme, practice, or policy is implemented as expected and achieving desired outcomes.	Notes on indicators and data sources	Organisational capacity assessment: Recommendations for Levels 1, 2 or 3 Note: methods listed are suggested and not exhaustive. Recommend the use of mixed methods that is deemed appropriate according to the context of the work.
 NORMATIVE WORK						
1. POLITICAL FRAMEWORK	Introduce, review and/or reform a comprehensive national plan for EVAWG	National plan is inclusive of all forms of VAWG, with strategies for prevention and response, and conforms with international standards and evidence and implemented by UN and national actors in a coordinated manner	Political will and commitment across society to end harmful gender norms by taking action for the primary prevention of VAWG, multi-sector coordination, and the provision of essential services.	<ul style="list-style-type: none"> (i) Proportion of ever-partnered women and girls aged 15+ subjected to physical and or sexual violence by a current or former partner, in the last 12 months (SDG indicator 5.2.1) (ii) Proportion of women and girls aged 15 years and older subjected to sexual violence by persons other than an intimate partner in the previous 12 months (SDG indicator 5.2.2) (iii) Does the country have a national plan embedding a framework to understand the causes and factors of VAWG? (iv) # of high-level commitments to Gender Equality, violence prevention and essential service provision 	Monitoring and Evaluation (M&E) Guidance of the RESPECT Framework provides strategy impacts, indicators and guidance notes.	<ul style="list-style-type: none"> Level 1 Process evaluation to assess implementation progress against the intended programme goals Level 2 and 3 Mixed qualitative and quantitative methods Example of measures and data collections: DHS survey, conducted every 5-7 years, last one was 2016, disaggregated by type of violence, age and other relevant characteristics. Survey/questionnaire of high-level officials Annual review of desk top analysis of policy commitment and implementation, i.e., Government Gazette.

APPENDIX 1 : SAMPLE IMPACT EVALUATION FRAMEWORK FOR EACH BUILDING BLOCK (2/3)

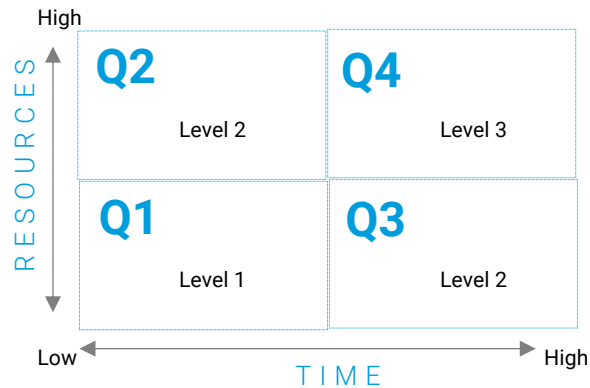
 BUILDING BLOCKS	 ACTIVITIES	 OUTPUTS	 OUTCOME STATEMENT	 EVIDENCE OF CHANGE	 GUIDANCE NOTES	
 COORDINATION WORK						
5. LEADERSHIP AND ADVOCACY	Women and girls leading and participating in community sporting events, with gender equality perspective.	Local, national and international organisations are led by women and girls as agents of change.	Responsibility for leadership, advocacy and ownership of strategies for the EVAWG is increasing nationally beyond the ministries or departments for women.	<ul style="list-style-type: none"> (i) # of advocacy initiatives led by women (ii) # women and girls in leadership roles who feel empowered (iii) # Women’s organizations leading the PVAWG and response sectors (iv) # Women appointed to senior leadership across government to implement and oversee who feel empowered 	<p>Indicators and outcomes align to SDG Target 5.5, which is to ensure women’s full and effective participation and equal opportunities for leadership at levels of decision-making in political, economic and public life.</p> <p>The impact of leadership development and advocacy interventions maybe difficult to capture holistically. A contextualized analysis of the extent to which leaders feel empowered and what women leaders do once they gain office.</p>	<p>Level 1 Process evaluation to assess implementation progress against the intended programme goals</p> <p>Level 2 and 3 Experiment/quasi-experimental not recommended.</p> <p>Mixed method quantitative and qualitative measures can assess immediate outputs and outcomes.</p> <p>Data sources may include statistics received from parliaments, administrative data based on electoral records or public administrative data available from ministries. CSO survey, disaggregated sector, setting, region or province</p>

APPENDIX 1 : SAMPLE IMPACT EVALUATION FRAMEWORK FOR EACH BUILDING BLOCK (3/3)

 <p>BUILDING BLOCKS</p>	 <p>ACTIVITIES</p>	 <p>OUTPUTS</p>	 <p>OUTCOME STATEMENT</p>	 <p>EVIDENCE OF CHANGE</p>	 <p>GUIDANCE NOTES</p>
 <p>OPERATIONAL WORK</p>					
<p>5. LEADERSHIP AND ADVOCACY</p>	<ul style="list-style-type: none"> Implement primary prevention social marketing for behaviour change campaigns and programming directed at whole of population or at-risk priority populations using different platforms Utilise community-level organisations for social marketing and awareness raising of Gender Equality, human rights and access to essential services Capacity development and training for service providers Implement essential services in urban, rural and remote geographic areas to ensure access for all. Deliver programs designed to deliver economic empowerment and opportunity for women and girls. Deliver programs for men and boys that support positive behaviour. 	<p>Service providers deliver quality, coordinated essential services to victim survivors and families, justice systems hold perpetrators accountable.</p> <p>The prevention and response sectors' capacity integrates global standards and guidelines for essential service delivery.</p> <p>The prevention and response sectors collect monitor, collect and use client and perpetrator data, in an ethical manner.</p> <p>Women, girls and other stakeholders, understand and exercise their human rights to access quality services so the impacts of violence are addressed, and perpetrators are held accountable.</p>	<p>Social norms from the individual, relational, community and societal levels support non-violence and gender equitable relationships and uphold women's empowerment as a norm and do not accept violence against children or other marginalised groups. Victim survivors of violence are aware of and confident to use sexual and reproductive health services.</p> <p>Essential service personnel do not accept attitudes that condone or justify violence, hold perpetrators to account, follow instituted practices uniformly across institutions in all geographical locations and in relation to all social groups.</p> <p>Women and girls move autonomously and safely within and beyond their communities</p>	<p>(i) # of programmes funded and implemented in primary and response sectors</p> <p>(ii) # of acceptable, quality services</p> <p>(iii) # of service providers with capacity to provide coordinated, quality services</p> <p>(iv) # of women and girls who make their own informed decisions regarding sexual relations, contraceptive use and reproductive health care</p> <p>(v) Proportion of ever-partnered women and girls aged 15+ subjected to physical and or sexual violence by a current or former partner, in the last 12 months</p> <p>(vi) rates of types of violence and % of prosecuted cases that result in conviction</p> <p>(vii) # of girls and boys aged 1-17 years who experience physical punishment and/or psychological aggression by caregivers</p> <p>(viii) % change in reported incidents resulting in assault, injury or death</p> <p>(ix) % change in women reporting feeling safe in public and reduced incidents of sexual harassment or gender-based discrimination</p>	<p>Links to SDG Target 1.4; 5.5 S1-4 Services Ensured; S1.1. Number of service providers demonstrating improved response to VAW survivors. E1-3 Environments are made safe. C1-6 Child and adolescent abuse prevented.</p> <p>Link</p> <p>Indicators (ii) and (iii) See 3. Institutional frameworks. Indicators (iv - ix) See the RESPECT Framework: T1-2 Transformed attitudes, beliefs and norms.</p> <p>Link</p> <p>Indicator (vi) See the Violence Against Women and Girls: A Compendium of Monitoring and Evaluation Indicators, sub-indicators for improved institutional response. See also the Essential Services Framework: Common characteristics of quality essential services. Effective communication and participation by stakeholders in design, implementation and assessment of services</p> <p>Link</p> <p>Level 1 to 3 capacity: Mixed methods (qualitative and quantitative) are strongly recommended.</p> <p>Level 2 and 3 Population-based survey including attitudinal and personal safety Training participant survey (pre/post) Service level surveys</p>

APPENDIX 2 : ORGANIZATIONAL CAPACITY FOR EVALUATION (1/2)

The graph below shows four quadrants. Each quadrant has a high/low scale on the vertical and horizontal axes. The vertical axis is RESOURCES. The horizontal axis is TIME. Answer each question and rate your response against both axes. Considering both the resources and time, and place an X anywhere in the quadrant you think most accurately reflects your position.



Example: A five-year project, covering a population of over 50 million women, has allocated funding for one M&R specialist. This project would receive a low/high rating against resources and time in Quadrant 3.

You will place an X for each question on the graph giving you a cluster of X. Looking at the distribution across the quadrants, make a judgment about your level of evaluation capacity.

- **Level 1 - Quadrant 1:** low across resources and time
- **Level 2 - Quadrant 2 and 3:** varying levels of resources and time
- **Level 3 - Quadrant 4:** consistently adequate resources & time

QUESTIONS

1

LENGTH OF PROGRAMME
How long is the programme of work funded for?

1 year = Low
2 to 3 years = Medium
4 to 5 years = High

2

FUNDING FOR M&E ACTIVITIES
Funding available for monitoring and evaluation activities (i.e. data collection and monitoring systems) for the programme duration?

UN Women evaluation policy recommends 2-3% of programme expenditure for evaluation; and 8-10% for monitoring)

1% ← → 10%
Low High



3

RESOURCES FOR EXTERNAL CONSULTANT

Resourcing external consultants for baseline, mid-term and a summative Impact Evaluation? Are resources sufficiently high to fund an external evaluator to do an in-depth analysis using appropriate methods to generate rigorous evidence of impact of the programme? Has there been previous evaluation or research work undertaken?

No evaluation work can be undertaken = low resources and time

Programme evaluation at end-point only = high resource / low time

Yes, but selective about baseline or midterm = low resource / high time

Yes, resources and time is adequate = high resources and time



4

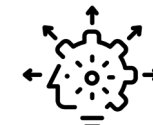
M&R SPECIALIST

Does the CO have a M&R specialist for the duration of the programme of work?

0 Full-time equivalent staff = Low

2 Full-time equivalent staff = High

Do you have the IE expertise to manage the programme of work at the country office? (If no, the resource axis is low). Does the programme of work require more than one person in a full-time capacity to manage stakeholders and participants? Are the timeframes of the programme of work able to accommodate recruitment and procurement processes to secure country-level expertise?



5

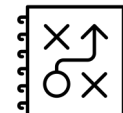
EVALUABLE PROGRESS

(relationships with stakeholders, donors, political sensitivities, budget, execution of programme on track, evaluation plans, monitoring systems)

Not on track, budget is limited = low / low

Commenced (on time), but key deliverables cannot be met due to unforeseen circumstances (preventing field work, stakeholder engagement low, participant recruitment etc.) = high / low


On track, on budget and meeting all project milestones = high / high



APPENDIX 2 : ORGANIZATIONAL CAPACITY FOR EVALUATION (2/2)

RESULT	LEVEL	DESCRIPTION OF LEVEL
Q1	1	<ul style="list-style-type: none"> ▪ Programme of work is short-term ▪ Funding for evaluation is insufficient ▪ Specialist evaluation oversight is low ▪ Capacity to recruit external evaluator is low ▪ Programme progress is low <p>Programme of work is not suitable for Impact Evaluation however; some evidence of change may be found. Evaluation activity may only generate evidence of proximal change, that is immediate and/or incremental outcomes that may eventually lead to desired long-term outcome but cannot be concluded with this level of time and resource.</p>
Q2 or Q3	2	<ul style="list-style-type: none"> ▪ Programme of work is short to medium term ▪ Funding may not be sufficient to establish evidence of impact ▪ CO evaluation capacity is low, impeding execution of evaluation plans ▪ Capacity to recruit external evaluator ▪ Programme of work is on schedule <p>Programme of work may be suitable for Impact Evaluation if enough evidence can be generated to show the project's impact is consistent with the planned impact. Over time, evaluation activity can report progress on prioritized programmes of work or with a well-designed and executed end-point evaluation. A multiplicity of performance indicators may be used. The evaluation activity may be participatory. Is likely to be mixed-method, and non-experimental in design. Results have some external validation by third-party perspectives, such as key stakeholders or programme partners, or external evaluator.</p>
Q4	3	<ul style="list-style-type: none"> ▪ Program of work is medium to long-term, and more likely to show impact over time ▪ Funding is sufficient for evaluation and monitoring across the program of work ▪ CO evaluation capacity is sufficient ▪ External evaluator can be procured ▪ Programme of work is on schedule <p>The programme's timeframes and resources are sufficient to fund evaluation research that demonstrates how the programme of work has produced long-term, desirable change. Monitoring systems are well-developed and are implemented to monitor progress on a continual basis over a medium to long-term timeframe. Multiple data collections provide comprehensive coverage with clear, measurable, and meaningful indicators. High levels of participatory engagement with the programme of work and evaluation to validate the findings.</p>

APPENDIX 3 : MEASURING PROGRESS USING THE BUILDING BLOCK FRAMEWORK*

 BUILDING BLOCKS	OUTPUTS of the program of work you are evaluating	OUTPUT SUCCESS CRITERIA				SCORE
		0 Pre- progress Ad hoc, short-term, standalone change.	1 Early progress Opportunistic and piecemeal changes to individual enabling building blocks.	2 Intermediate Progress Root causes and enabling factors are identified and addressed	3 Advanced progress Deep seated institutionalised change across the (all building blocks) supporting a transformation in collective and individual behaviors and changes in underlying social norms.	
Political framework		Ad hoc	Influential champion(s) identified to advocate, lead and coordinate for change	Political will demonstrated through systematic changes to the legal & institutional framework	Committed, effective & accountable leadership at all levels including commitment to resourcing	
Legislative & policy framework		No framework thematic priority or ToC	Legislation & policies reviewed and reformed to align with international standards	Intervention strategies, programmes, policy and action plans developed through participation of all key stakeholders	Comprehensive GEWE laws, policies, strategies & plans in place, endorsed and harmonised	
Institutional framework (including norms & standards)		Institutions and services lack capacity to implement and integrate legislation & standards	Institutional frameworks systematically integrate best practice (e.g., codes of conduct, roles & responsibilities)	Guidelines integrating standards are in place	Capacity development programmes (e.g. employee training) ensure GEWE informed practices	
Financial framework		Legislation, essential services & initiatives not implemented due to insufficient resourcing	Finance & development ministries involved in review of funding allocations for the thematic priority field	Feasible financing es are developed with appropriate budget allocations	Sustained funding sources in place, implemented, monitored and tracked (e.g. oversight mechanisms)	
Leadership & advocacy		Ad hoc and intervention-specific leadership	UN Women lead and coordinate UN work and advocate for national and local leadership in companion and correspondent fields	UN Women is the trusted partner of choice & its leadership ensures coherence, alignment with norms	UN Women systematically supports national/local leadership, ownership, championing & advocacy	
Partnerships & networks		No partnerships, networks or collaboration and stakeholders operate in silos	Opportunistic partners leveraged between government and key stakeholders (e.g., human rights)	Sustained coordination mechanisms provide forums/spaces for engagement, dialogue and advocacy	Coordination/partnerships institutionalised for multi-sector/stakeholder implementation of plans	

*Based on an approach developed for the "Pacific Resilience Standards: A Guide for Pacific Resilience Practitioners: PIFS/World Bank" (Selby, 2020) – awaiting publication.

APPENDIX 3 : MEASURING PROGRESS USING THE BUILDING BLOCK FRAMEWORK

BUILDING BLOCKS	OUTPUTS of the program of work you are evaluating	OUTPUT SUCCESS CRITERIA				SCORE
		0 Pre- progress Ad hoc, short-term, standalone change.	1 Early progress Opportunistic and piecemeal changes to individual enabling building blocks.	2 Intermediate Progress Root causes and enabling factors are identified and addressed	3 Advanced progress Deep seated institutionalised change across the (all building blocks) supporting a transformation in collective and individual behaviors and changes in underlying social norms.	
Mobilisation & capacity of national actors		Key stakeholders are not engaged	Opportunistic mobilisation & outreach to key stakeholders (e.g., women's health)	Systematic mobilisation/engagement and development of capacity development strategies & plans	Capacity development institutionalized	
Mobilisation & capacity of subnational actors		Awareness raising and outreach is ad hoc or one-off	Piecemeal mobilisation and capacity development but women and girls increasingly aware of rights	GEWE systematically mainstreamed into all curricula, employee training & outreach activities	Increased awareness, knowledge, understanding & commitment to gender, human rights and service availability	
Planning & design		Thematic priority is targeted in small-scale, fragmented, standalone activities	Thematic priority is incorporated into development planning processes & tools	Thematic priority is systematically integrated into intervention design with involvement of women & girls	Institutionalization of joint planning of multi-sectoral/stakeholder programmes mainstreaming GEWE	
Implementation		Relevant legislation, programmes & services not implemented due to a lack of human/financial capacity	Capacity of services to deliver quality services is piecemeal but increasingly coordinated	Support services systematically provided in all geographic areas & EVAWG programmes implemented	Women/girls understand rights, have access to a range of services and these are widely used in all areas	
Data collection, M & E		There is no critical mass of data, reporting or lesson sharing	Collection and management of data to build knowledge	Decision making is more systematically informed by VAWG data, which is stored centrally & accessible	Policies & programmes are informed by quality evaluation providing knowledge and evidence of intervention effectiveness according to context	

Key: Scoring progress using success criteria >

Level of assessment	PRE-PROGRESS	EARLY	INTERMEDIATE	ADVANCED
One building block	0	1	2	3
One work area*	0	1-4	5-8	9-12
All three work areas /building blocks	0	1-12	13-24	25-36

*e.g. normative, cooperation, operational

APPENDIX 4: METHODS FOR IMPACT EVALUATION AND ADDITIONAL RESOURCES

Approaches, ToR and methods for impact evaluation, European Regional Development Fund

<https://www.interact-eu.net/download/file/fid/4996>

Impact Evaluations in UN Agency Evaluation Systems - Guidance on Selection, Planning and Management, UNEG

<http://www.uneval.org/document/download/1875>

Impact Evaluation in Practice, Second Edition, World Bank Group*

<https://elibrary.worldbank.org/doi/pdf/10.1596/978-1-4648-0779-4>

Handbook on Impact Evaluation Quantitative Methods and Practices, The World Bank*

<https://elibrary.worldbank.org/doi/pdf/10.1596/978-0-8213-8028-4>

Impact Evaluation Series, UNICEF

<https://www.unicef-irc.org/KM/IE/impact.php>

The Mystery of the Vanishing Benefits: An Introduction to Impact Evaluation, The World Bank Economic Review**

<https://academic.oup.com/wber/article-abstract/15/1/115/1682017?redirectedFrom=fulltext>

A step-by-step guide to impact evaluation, ILO

https://www.ilo.org/wcmsp5/groups/public/---ed_emp/documents/instructionalmaterial/wcms_627314.pdf

Evaluability assessment tools:

<https://genderevaluation.unwomen.org/-/media/files/un%20women/gender%20evaluation/handbook/tool%205%20-%20how%20to%20conduct%20an%20evaluability%20assessment.pdf?la=en&vs=2521>

<https://www.vera.org/downloads/publications/cultivating-evaluation-capacity-sexual-domestic-violence-guide.pdf>

Checklist for EVAWG

https://vetoviolence.cdc.gov/apps/violence-prevention-practice/sites/vetoviolence.cdc.gov/apps/violence-prevention-practice/files/evaluation-plan-checklist-508_0.pdf

Base line studies:

1. <https://openknowledge.worldbank.org/bitstream/handle/10986/11075/643810BRI0Reco00Box0361535B0PUBLIC0.pdf?sequence=1&isAllowed=y> (World Bank)

2. <http://www.oecd.org/dac/evaluation/dcdndep/37671602.pdf> (OECD)

* World Bank eLibrary subscriber-only content

** Accessible with an Oxford Academic account

Literature List

Abramsky, Tanya, Karen M Devries, Lori Michau, Janet Nakuti, Tina Musuya, Nambusi Kyegombe and Charlotte Watts (2016). "The impact of SASA! a community mobilisation intervention, on women's experiences of intimate partner violence: secondary findings from a cluster randomized trial in Kampala, Uganda." *Journal of Epidemiology and Community Health* 70(8). February 2016.

Acumen. "Acumen and Unilever launch Lean Data How-to Guide for Measuring Gender Impact." June 05. 2019. <https://acumen.org/blog/acumen-and-unilever-launch-lean-data-how-to-guide-for-measuring-gender-impact/>

BOND (2015). "Impact Evaluation: A Guide for Commissioners and Managers"

Buller, Ana Maria, Amber Peterman, Meghna Ranganathan, Alexandra Bleile, Melissa Hidrobo, and Lori Heise (2018). "A Mixed-Method Review of Cash Transfers and Intimate Partner Violence in Low- and Middle-Income Countries." *The World Bank Research Observer*, 33(2), pp.218–258.

CARE (2018). "Counting the Cost: The Price Society Pays for Violence Against Women." Report.

CCSL (2014) *Lessons in the Theory of Change: Monitoring, Learning and Evaluating Knowledge to Action*; and DFID (2012) *Violence Against Women and Girls, CHASE Guidance Note Series*.

Duvvury, Nata, Patricia Carney, Nguyen Huu Minh (2012). "Estimating the Costs of Domestic Violence Against Women in Viet Nam." *UN Women*.

Ellsberg, Mary, Diana J Arango, Matthew Morton, Floriza Gennari, Sveinung Kiplesund, Manuel Contreras, Charlotte Watts (2015). "Prevention of violence against women and girls: what does the evidence say?" *Lancet* 385: 1555-66.

Fearon, James and Anke Hoeffler (2014). "Benefits and Costs of the Conflict and Violence Targets for the Post-2015 Development Agenda: Post-2015 Consensus." Copenhagen Consensus Center. *Conflict and Violence Assessment Paper*.

Fulu, Emma and Xian Warner (2020). "Literature Review: Ending Violence Against Women and Girls." Paper prepared by The Equality Institute for the Australian Department of Foreign Affairs and Trade. January 2020.

Gibbs, Andrew and Kate Bishop (2019). "Combined economic empowerment and gender-transformative interventions." *What Works Evidence Review*. September 2019.

Gibbs, Andrew and Laura Washington. "Stepping Stones and Creating

Futures: An evidence-based intervention to prevent violence against women and improve livelihoods. October 2018.

Glennerster, R, C Walsh, and L Diaz-Martin (n.d.). "A Practical Guide to Measuring Women's and Girl's Empowerment in Impact Evaluations." J-Pal.

Gómez, Margarita Puerto, Manuel Contreras-Urbina, Brian Heilman, Amber Hill, Anne von Au, Jennifer Zelaya and Diana J. Arango (2016). "Community-Based Approaches to Intimate Partner Violence: A Review of Evidence and Essential Steps to Adaptation." *The Global Women's Institute and World Bank*. May 24, 2016.

INTRAC (2017). "Qualitative Comparative Analysis (QCA)." <https://www.intrac.org/wpcms/wp-content/uploads/2017/01/Qualitative-comparative-analysis.pdf>

International Food Policy Research Institute. "Women's Empowerment in Agriculture Index", accessible at: <https://www.ifpri.org/project/weai>

Jan, Stephen, Giulia Ferrari, Charlotte H Watts, James R Hargreaves, Julia C Kim, Godfrey Phetla, Linda A Morison, John D Porter, Tony Barnett and Paul M Pronyk (2011). "Economic evaluation of a combined microfinance and gender training intervention for the prevention of intimate partner violence in rural South Africa." *Health Policy and Planning* 26: 366-372.

Jewkes, Rachel, Erin Stern and Leane Ramsoomar (2019). "Preventing violence against women and girls: Community activism approaches to shift harmful gender attitudes, roles and social norms." *What Works. Evidence Review*. September 2019

Kabeer, L (1999). "Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment." *Development and Change* Vol 30 (1999), 435-464.

Lombardini, S, R Garwood, H Hassnain (2019). "Women's Empowerment in Lebanon: Impact Evaluation of the project "Women's Access to Justice in Lebanon." *Oxfam GB*.

Londsdale et al., (2015) *Transformational adaption*.

OECD DAC Network on Development Evaluation (EvalNet) (2019), *Better Criteria for Better Evaluation: Revised Evaluation Criteria definitions and Principles for Use*; <http://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>

Remme, Michelle, Christine Michaels-Igbokwe Charlotte Watts (2015). "Approaches to scale-up and assessing cost-effectiveness of programmes to prevent violence against women and girls." *What Works to Prevent Violence Against Women and Girls Evidence Reviews Paper 4*. September

2015.

Roy, Shalini, Melissa Hidrobo, John Hoddinott and Akhter Ahmed (2017). "Transfers, Behavior Change Communication, and Intimate Partner Violence: Postprogram Evidence from Rural Bangladesh." IFPRI. Discussion Paper 01676. September 2017.

Selby, S and Jiwanji, M (2016) *Risk Governance: Building Blocks for Resilient Development in the Pacific: A Policy Brief – UNDP Pacific Risk Resilience Programme*

Stedman-Bryce, G. (2017). "Women's Empowerment in South Africa: Evaluation of the Raising Her Voice Project." *Oxfam GB*.

UNEG (2013). "UNEG Handbook for Conducting Evaluations of Normative Work in the UN System."

UNEG (2013). "Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management."

UN Women (2012). "Evaluation policy of the United Nations Entity for Gender Equality and the Empowerment of Women."

UN Women (2014). "UNEG Handbook for Conducting Evaluations of Normative Work in the UN System." *UN Women Country Portfolio Evaluation (CPE) Guidance*.

Vyas, Seema (2013). "Estimating the Association Between Women's Earnings and Partner Violence: Evidence from the 2008-2009 Tanzania National Panel Survey." *Women's Voice, Agency, and Participation Research Paper 2*, World Bank, Washington, DC.

Wagman, Jennifer A, Ronald H Gray, Jacquelyn C Campbell, Marie Thoma, Anthony Ndyababo, Joseph Ssekasanvu, Fred Nalugoda, Joseph Kagaayi, Gertrude Nakigozi, David Serwadda, and Heena Brahmhatt (2015). "Effectiveness of an integrated intimate partner violence and HIV prevention intervention in Rakai, Uganda: analysis of an intervention in an existing cluster randomized cohort." *Lancet Global Health* 3.

What Works. "Impact of Indashyikirwa: An innovative programme to reduce partner violence in rural Rwanda." *Evidence Brief based on the evaluation conducted by Kristen Dunkle, Erin Stern, Lori Heise, Lyndsay Mclean and Sangeeta Chatterji*. August 2019.

White, H and S Sabarwal (2014). "Quasi-Experimental Design and Methods". *UNICEF Office of Research*.

Wilson-Grau, R. and H Britt (2012) *Outcome Harvesting*, Ford Foundation.

