Women's Empowerment Impact Measurement Initiative (WEIMI)

Part II: Defining Measurement Elements of the Theory of Change

Content adapted from: *Picard, M. and Gillingham, S. (2012) Women's Empowerment Impact Measurement Initiative Guidance*

Table of Contents

Introduction: Women's Empowerment Impact Measurement Initiative (WEIMI)

What is WEIMI?

WEIMI was launched by CARE USA and from **2010 to 2012** provided **technical support** to select country offices (COs) **to operationalize key women's empowerment and gender-sensitive high-level indicators**, in the **context** of their **long term programs (LTPs).** The initiative was designed to build on CARE's Strategic Impact Inquiry (SII) and the P-shift efforts. The initiative worked with technical teams in 6 COs to support efforts to **build coherent systems** which **link** rigorous **initiative-level monitoring and evaluation to impact measurement systems**.

WEIMI COs

- Bangladesh
- Burundi
- Egypt
- Mali
- Niger
- Tanzania

What is the WEIMI Guide?

The WEIMI Guide is a harvesting of the lessons learned and good practices of the WEIMI CO teams. A few important things to note about the content in the WEIMI Guide are:

- In view of the impact groups of the six WEIMI COs, the **experience** presented **relates more to women than to girls**.
- The guide only goes as far as the WEIMI COs progressed by the end of the initiative. Hence, the **experience** is **richest in Parts I & II**, while **Parts III & IV** draw more on **external sources**.
- The guide **on its own is not sufficient** to enable a CO to create their impact measurement system – it is **simply a tool for COs to reference** when creating an impact measurement system with the **support of individuals with** the **expertise** to fully support that process.
- The guide provides information on the approach that was taken by the COs as part of the WEIMI initiative and **illustrates one of many ways to approach the development of an impact measurement system** (i.e. it is not "the" guide to creating an impact measurement system for women's and/or girls' empowerment programs).

Who is the WEIMI Guide for?

The target audiences for this guide are:

- **CO teams already engaged in similar work** in women's and/or girls' empowerment or the promotion of gender transformation
- **CO teams that are planning** to develop similar programs
- **Regional teams** working with CO teams to build similar programs

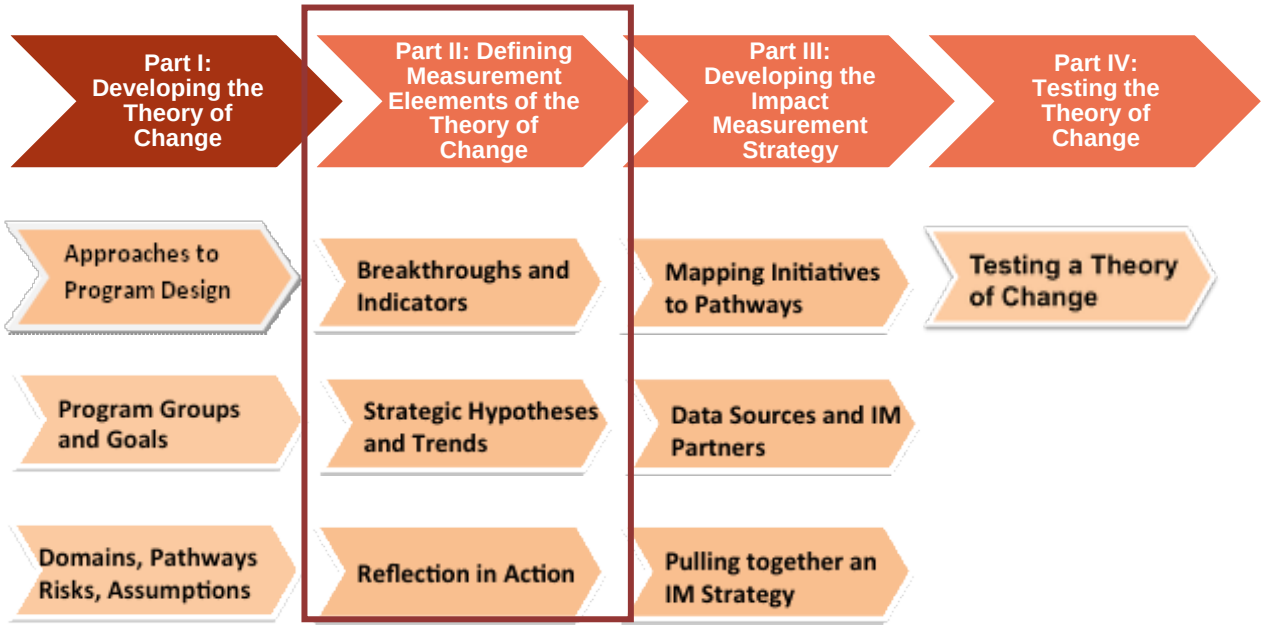
What you will find in the WEIMI Guide?

In the WEIMI Guide you will find the following information organized across four parts:

- **Part I Developing the Theory of Change**: Here you will find the basics of a theory of change (TOC), key considerations for developing a TOC for a women's and/or girls' empowerment program and some tips on how to ensure the quality and robustness of your TOC. This section also includes 3 briefs that illustrate how different CO's approached the development of their LTP; guidance on how to conceptualize impact, sub-impact, target, and stakeholder groups; and how to define domains of change, pathways, assumptions and risks.
- **Part II Defining Measurement Elements of the Theory of Change**: Here you will find ideas on how to conceptualize "impact", the challenges of measurement, areas to focus on for measurement, and key considerations of developing an impact measurement system in the context of women's and/or girls' empowerment programs. This section also includes 3 briefs that provide guidance on defining breakthroughs and indicators; developing hypotheses and monitoring trends, assumptions and risks; and reflection in action.
- **Part III Developing the Impact Measurement Strategy**: Here you will find good practices for measuring your TOC. This section also includes 3 briefs that describe how to map initiatives to pathways; identify and select data sources and partners; and the planning and implementing of a impact measurement strategy.
- **Part IV Testing the Theory Change**: Here you will find lessons from WEIMI COs on testing the TOC. This section also includes 1 brief that speaks to what it means to test your TOC; how to sequence the testing of your TOC; and guiding questions for validating your TOC.

How is the WEIMI Guide Organized?

The WEIMI Guide is organized like a roadmap consisting of 4 Parts with discrete Briefs within each as shown here:



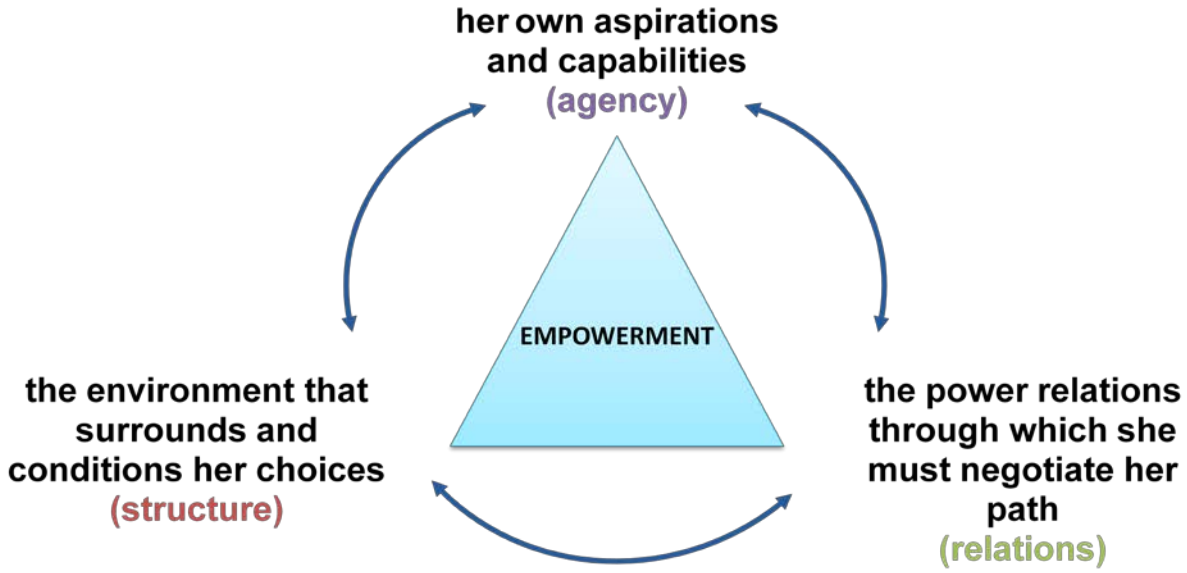
This section of the WEIMI Guide will focus upon **Part II: Defining Measurement Elements of the Theory of Change.**

Part 2: Defining Measurement Elements of the Theory of Change

What is "Impact" in the context of Women's and/or Girls' Empowerment Programs?

Impact is the highest level of change that occurs for the population group (as defined by our impact group). In the context of women's and/or girls' empowerment programs we are speaking of impact that is over a longer-term time horizon and whose meaning is associated with a social change process. Most WEIMI countries articulated an impact goal that includes the term "empower," and "women's empowerment" is defined in the SII to be:

We understand empowerment as the sum total of changes needed for a woman to realize her full human rights – the interplay of changes in:



If this is an accepted definition, then "impact" is no less than the attainment of rights.

The behavioral and structural changes amongst other actors and institutions that are crucial for attaining this goal occur at the level of pathways and domains of change. However, if gains made for women or girls under the program are incremental (e.g., improvements in rates of girls' retention in school) but not sustained by deeper social change that is grounded in changing power relations,

then rights attainment will be an elusive goal. It is also for that reason that the term "breakthroughs" was incorporated into the conceptualization of a theory of change, to mark progress over time, highlighting changes that are not easily reversed and capture both breadth and depth (**See Part 2: Brief 1: Breakthroughs and Indicators Section** below).

Challenges of Measuring Women's and Girls' Empowerment

What are some of the challenges specific to women's and girls' empowerment?

Empowerment is complex, multi-dimensional, nonlinear, reversible and unpredictable. Women may be "empowered" in one dimension but not in others. It is beyond the scope of one organization, let alone one time/resource-bound project.

There is a tension between global efforts to harmonize and standardize measures for women's and girls' empowerment and the reality that definitions for empowerment vary across contexts, time, groups and individuals. They are embedded in socio-cultural systems and communities with their own set of beliefs and values. Moreover, even over the course of a woman's life, her social position and aspirations of what is possible may change.

The Strategic Impact Inquiry (SII) on Women's Empowerment

A three-year, multi-site, participatory and rigorous review assessing the impacts that CARE's worldwide programming has (and has not) had on women's empowerment.

Empowerment is highly subjective and therefore difficult to measure, as the SII revealed. Indeed, it is said that contextual factors (e.g., the community context) are often more important in determining women's empowerment than individual level factors. As such, **social norms** may have a stronger influence on empowerment than would the abilities or capabilities or individual women/girls, especially in the public sphere. That empowerment is more a **process** than a **condition** also makes it difficult to measure. Recent efforts have sought to capture process through direct measures of decision making, control, choice, etc., which should be measured at least two points in time.

For more detailed reviews of the challenges of measuring women's empowerment and social change for gender equality see the following:

- Strengthening Monitoring and Evaluation for Women's Rights: Twelve Insights for Donors
- Capturing Change in Women's Realities: A Critical Overview of Current Monitoring & Evaluation Frameworks and Approaches

If Gender Equality is the Impact Goal, What Should be the Focus of your Impact Measurement?

The 2012 World Development Report on Gender Equality and Development highlighted five priority areas for its global agenda:

- Closing gender gaps in human endowments,
- Promoting women's access to economic opportunities,
- Closing gender gaps in voice and agency,

- Preventing intergenerational reproduction of gender inequality for which it emphasizes a strategy of investing in adolescent girls and boys, and
- Supporting evidence-based public action, for which it proposes that new information be generated, particularly on ownership of assets and on decision making within households

Information on what happens within **households** is sorely missing and key to understanding many of the gender gaps. This knowledge gap exists because households have not always been disaggregated. Accurate recording of life events (e.g., death) and property must also improve to aid the enforcement of laws.

Within CARE, guidance is now available through the Good Practices Framework that identifies 8 core areas of inquiry for a gender analysis:

All of these have relevance across the spectrum of the WEIMI CO programs and are equally applicable to all programs, regardless of whether the impact group is women or girls. Gender equality measures ought to be systematically introduced into all areas of programming and followed up with strategies to address disparities and gender injustices.

It is also worth reminding ourselves that to address the underlying causes of poverty & social injustice and gender inequality, i.e., to achieve ***transformational*** development, we will need to be measuring change in:

- **Gender roles and relationships within household, community and at a macro level**
- All three of these levels will require our attention in measuring the change, across **structure, agency and relations**.
- **The lives, networks and status of individual women, households, solidarity groups, communities**, etc.
- **Social norms around accepted gender-based roles and behaviors**. In examining community dynamics, social norms reveal themselves; these are essentially the perceptions, often incorrect perceptions, on what are typical or desirable behaviors. Social norms are properties of a social group (e.g., all men in the community, VSL groups) and are powerful when they apply to people's everyday life. They influence individual and group behaviors.

Further, attention to gender-sensitive measures, whether in a program specific to women or to girls, includes systematically **disaggregating data for both genders**. Sensitivity is also required for differences in **age groups** and **stages of the life cycle.**

These differences must be reflected in the line of inquiry as well, framing questions and employing methods appropriate to the age group and gender.

Key Considerations for Developing an Impact Measurement System

Developing an "impact measurement system" should not be viewed narrowly as a technical undertaking or putting into place the hardware to manage information. It encompasses much more, and so it is important to:

- Develop a strategy of how you plan to measure change at all levels of your theory of change
- Construct a framework for the knowledge and the results you expect to produce
- Establish processes for monitoring and reflection that become a way of working
- Continuously build capacity of staff and partners in monitoring and evaluation iteratively with acting and planning.

When monitoring and evaluation ceases to be episodic and is supported by continuous reflection as an internalized practice, then the gap between "implementing good initiatives" and "telling an impact story" will close.

Brief 2.1: Breakthroughs and Indicators

Developing the measurement system will require the support of monitoring & evaluation specialists, which was a mix of both in-house and external staff for all six COs. **Expertise in M&E and in gender equity/women's and girls' empowerment are both essential for the work on measurement** – to ensure rigor and to tap the experience within CARE and globally on measurement of women's empowerment.

You should expect that the design of indicators and hypotheses to be an iterative process, resulting in a final round of refinement of the theory of change so that it is specific enough for measurement purposes. Typically this process produces more 'sense-making' in the group until agreement is reached on the most feasible and priority indicators and hypotheses. It is wise to build in some slack time but not lose momentum in between reflection sessions or workshops. One way to manage the process is to use a combination of larger-scale participatory workshops with program staff alternating with core working groups.

Filter Questions

Periodically as you progress through the process of developing your measurement elements (breakthroughs, indicators, hypotheses), check to see if you are able to respond to the following questions (**See Annex A.2.1: Filter Questions for Measurement for full explanations of questions**):

- Are you designing your measurement systems with the capacity to disaggregate all data by sex (comparing males to females), as mandated by the CARE International Gender Policy?
- Do the measurement elements altogether capture changes in agency, structure and relations (CARE's empowerment framework)? It is especially important to measure change in structure and relations (i.e., not to privilege agency above the other two). See CARE's Good Practices Framework.
- Are changes in different units of analysis being measured - the individual, household, community, group level, and the aggregate (population level)?
- Are you measuring both process and outcome of women's and girls' empowerment? How are you disaggregating this and measuring equivalent measures in men and boys?
- Are you capturing both breadth (scale) and depth of impact?
- Are you capturing the essential changes among key stakeholders in government, civil society and the private sector reflected in your theory of change?
- Are you including men and boys in measurement of behavioral change at target group level?
- Are there any expected changes in the overall theory of change that are not being captured and may actually be slipping through the measurement cracks?

Defining Breakthroughs

Identifying breakthroughs can be done before or after you have worked on indicators (**See Annex A2.2 WEIMI Country Breakthroughs**). However, most COs found it helpful to think about breakthroughs after the pathways are constructed, i.e., what are breakthroughs as critical measures of progress in relation to an individual pathway or a group of pathways?

The intent behind the breakthrough concept was to be able to note critical turning points on a 15-year path of social change / empowerment for women and/or girls. Not all of these points in time are predictable but some can be anticipated, i.e., predicted. For example, , in retrospect, over the nearly 20-year history of women's savings & loans groups (MMD) in Niger, the following can be considered breakthroughs and might be considered possible opportunities for other COs to consider building into their pathways:

- Legal recognition of MMD groups
- The first MMD women became elected officials
- An organized network of MMD groups formed across the country
- The first Congress on the political participation of MMD women

Breakthroughs can exist as one-time events (structural, policy or precedent change) or as a critical threshold for an incremental change.

Breakthrough

A change that represents a significant leap forward that is not easily reversed. A breakthrough represents a change that affects both the breadth of impact (increasing impact on many more people in our impact group) and the depth of impact (increasing the level of wellbeing or transformation in the lives of our impact group). The change resulting from a breakthrough is reflected in the lives of people in our impact group whom we directly work with, as well as people in the impact group outside of our operational areas.

Example of a One-Off Event	Example of a Critical Threshold for an Incremental Change
The Domestic Violence Act is passed	50% of all medical facilities have specific confidential services available for survivors of violence who seek help.

While you still need indicators for pathways, consider what might be 5- or 10-year breakthroughs for these two pathways from Egypt's TOC:

- There is acceptance of gender equity and equality at all levels
- Women and men access services that prevent GBV and support survivors

Breakthroughs themselves do not need indicators, as goals do, because they serve as an aid to measuring change in a long-term program. Thus, you may find a breakthrough useful as:

- A progress marker on your pathways (as Burundi has done)
- The key changes in public policy and/or action by key stakeholders from your stakeholder analysis (which also qualify as progress markers)
- A 5-year milestone on your theory of change (associated with one or more pathways that are being tested) (as Egypt and Bangladesh)
- A critical threshold to an incremental change which means specifying the level (depth and breadth) of that change
- A way to leverage the achievement (that a breakthrough represents) to create wider, deeper impact

In the WEIMI CO experience, breakthroughs became a critical element in the measurement system for Mali and Niger (and possibly Burundi). The concept was the most challenging of all the elements

in the measurement system. Other CO teams made optional or minimized them to two or three key ones at the 5-year mark.

Defining Indicators

Impact Indicators and Domains of Change (DOC) Indicators

Proposed Criteria for Selecting WE Indicators

- Linked to MDI+ indicators
- Linked to PRSP and national statistics
- Experience within the CO
- Cost Effective
- Feasible
- Relevant
- Reliable
- Understandable (and with clearly defined terms)
- Existing source of data

Possible additions to measuring empowerment specifically:

- Sensitive to change over time
- Tested by others

See Annex A2.3 WEIMI CO DOC Indicators Categorized to the Global MDI Indicators

At this level, we are concerned with measuring change for the impact goal and for the domains of change. While there are practical aspects to this (see below), care should also be exercised to not settle for measures of change that represent lower levels or pre-conditions to achieving impact. It should be commensurate with a structural change, altering the conditions of life for all members of an impact group or sub-impact group – e.g., a change in a law or policy, a social norm, an institutional practice (integrating services for victims of violence in health centers), employment of women in jobs not traditionally open to them, a shift in power relations. Note that behavioral change amongst specific target groups, for example, is not tantamount to impact, nor are incremental improvements in women or girls' access to services. When they are on a par with men and boys or when this trend reaches a threshold of certainty that it will not be reversed, then this is impact.

As for indicators, making the best choices is not merely a technical procedure; it is an analytical process that begins with identifying the concept or issue to be measured, then probing until we come to an agreement on what it stands for, what its constituent parts are, what is it exactly we would like to change, and what the change would tell us. Only then, after clarity and agreement are attained, should the process of selecting indicators for the concepts identified and articulated in this way begin.

There is always the risk that you will end up with an excessively large number of indicators, which is what happened to most WEIMI countries. This may be an acceptable way to start, but then proceed with these guidelines for completing the selection process:

- After you have generated a list of indicators, prioritize which of these are the most critical changes that will be a measure of the domain of change or the goal (**See Annex A2.4 CO Process to Prioritize Indicators**).
- Sometimes, after reducing down a large number of indicators to the most critical changes, another round of prioritization and reduction is needed. You may then want to use the indicator prioritization worksheet, to reduce the number to a set of around 10-12 for impact and DOC levels combined (**See Annex A2.5 Indicator Prioritization Tool**).

- To ensure sufficiency in capturing the critical changes, 2-3 indicators at goal level that capture agency, structure and relations are recommended and are population-based.
- At impact goal level, if links can be established with the MDI+ or national level statistics, consider, before agreeing on these, whether the data already exist and are being gathered by repeat national surveys. Other sources may include the Demographic Health Survey (DHS) or the UN Development Report indicators. You will want to consider the cost of measuring a particular indicator vs. opting for a proxy indicator based on a secondary source.
- The use of secondary sources will present other challenges - the year of collection, the disaggregation of the data, and the precise variables being captured, may not always correspond to the requirements of the program. What are the alternatives?
    - Continuing to seek possibilities with other organizations for inclusion of indicators in their national level surveys
    - Continuing to identify a 'best fit' indicator for which data are already being collected
    - Conducting your own baseline, while weighing options to achieve cost-effectiveness (see Mali and Niger examples)
- In the aggregate, indicators at impact and DOC levels combined should include both population-based measures using secondary sources of information and measures adapted to the local context that you will likely test in your operational area.
- To finalize indicators, define operational terms as Bangladesh has done. This makes your indicator measurement-ready.

See **Annex A2.6 Indicator Tips** and **Annex A2.7 Example of Outcome Indicators and Operational Definitions from Bangladesh** for more information on constructing good indicators.

Are there women's empowerment indicators at impact and DOC levels that are in greater use?

Since 2008, CARE has promoted a Millennium Development Indicators Plus (MDI+) approach, which reflects the Millennium Development Goals (MDGs), to measure and communicate CARE's contributions to social change. CARE goes beyond the MDIs (indicated by the "+") to capture areas of work, such as governance, not covered by the MDIs and to signal the intent to apply measurement of change to the most marginalized groups who are often excluded from national programs and statistics.

Many of the MDI+ indicators that WEIMI COs selected were relevant at DOC level of measurement.

Of all the MDI+ indicators, 6 were chosen by at least 3 of the WEIMI countries, as shown in the table below (**X** = with slightly different wording). Also of note is that three of the WEIMI COs were participants in CARE Norway's WE program (2009-2013), namely, Mali, Niger, and Tanzania. This program had in common several MDI+ indicators, some slightly adapted, that appear as an * in the table below.

The Most Frequently Used Indicators amongst WEIMI Countries

THEME	BD	BI	EG	MA	NG	TZ
% men and women reporting meaningful participation of women in the public sphere*	•	•	•	X	•	X
% men and women with changed attitudes toward gender-	•		•	X	•	•

based violence*						
% households with access to secure land tenure, by sex of the head of household	**X**		**X**	**X**	•	**X**
% men and women reporting ability of women to effectively control productive assets*	**X**		**X**	**X**		•
% women reporting meaningful participation in decision making at household level in a domain previously reserved for men	•			•		•
% households with capacity to cope with environmental shocks without depleting assets, by sex of the head of household*				•	•	•

The span of the 6 indicators is indicative of the multi-dimensionality of women's empowerment. Most of them (perhaps excluding the last one) are fundamental to changes in life conditions – in health (esp. reproductive health), education, and economic security. They also reflect agency, relations, and structure. What appears in the top six are also in accord with research on women's empowerment/gender equality indicators; one study that examines the most frequently used indicators at household and individual level mentions three:

(a) domestic decision-making,

(b) access to and control over resources, and

(c) mobility and freedom of movement.

Only the last one is neither an MDI+ indicator nor does it feature in the indicators selected by WEIMI countries (**See Annex A2.8 MDI + Indicators**)

Indicators other than the MDI+ indicators were also selected and these vary across countries.

Using the categories of the global MDI indicators (poverty reduction, food security, education, women's empowerment, child health and nutrition, maternal health, sexual & reproductive health, HIV/AIDS, environmental sustainability, governance), the majority of those that fall under "Women's Empowerment" relate to women's political participation or representation (**See Annex A2.9: Impact Indicators for WEIMI Countries**).

What kind of baseline is needed for impact and DOC level indicators?

There are different choices for conducting baselines, as exemplified in WEIMI CO experiences (**See Annex A2.10 CO Baselines**):

- Collect primary data as part of a start-up initiative or evaluation for an existing initiative of some significant scale
- As a separate exercise, collect primary data in a designated geographic zone that meets your sampling frame and takes advantage of the synergies between initiatives and programs
- Conduct a separate baseline exercise for several programs at once or for a CO-level theory of change
- Relying on secondary data sources in combination with:

- any of the above, or
- measuring indicators in all relevant initiatives at the same time

The choice you make depends on:

- The opportunities (e.g., alignment with a new multiyear initiative)
- Economics – what are reasonable costs and what you can afford
- The quality and availability of secondary data

The baseline will only serve its purpose if you conduct it with the intent to repeat data collection on the same indicators with the same sampling frame in future points in time in order to have comparable data.

Defining Pathway Indicators

Pathways, as outlined in the earlier section of this guidance on "Developing the TOC," express the "how" of achieving a domain of change. Testing hypotheses within pathways is therefore an important measurement element. This involves measuring indicators for pathways to ensure that levels of change (both breadth and depth) for the pathway are being tracked. Some CO teams substituted pathway indicators with breakthroughs (See earlier section on Niger and Mali. **Also see Annex A2.11 WEIMI CO Pathway Indicators**).

Measuring breadth of change is indicated by the proportion, number or percentage within the unit of analysis. For this, it is also important to know your denominator. Measuring depth of change pertains to the type and degree of change occurring within the unit of analysis. Hypothetical Example:

Pathway	**Girls show greater confidence, competence, and leadership**
Breadth	% of girls in the program area
Depth	Level of confidence, type and level of competence, and level of leadership
Denominator	All girls in the age group of 7-14 in the convergence zone of the program

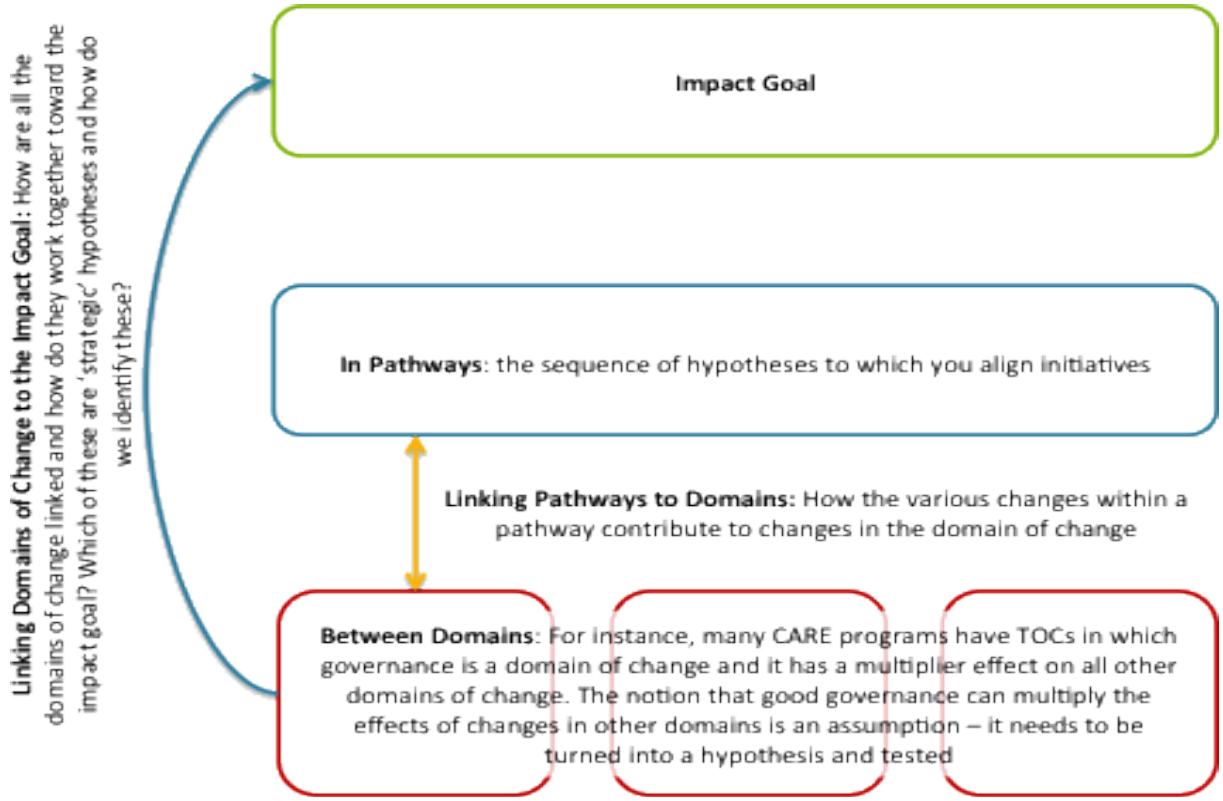
Finally, a practice that most WEIMI COs have found useful is to map initiatives to pathways at an early stage, before finalizing the measurement system. Making the link between the conceptual and the 'real world' is a nice break and will help teams bring their programmatic knowledge to bear on formulating pathway indicators. See next section for more information on this.

Brief 2.2 Strategic Hypotheses and Trends

Causal Relationships and Theories of Change

The Theory of Change is a web of causal relations.

The potential for hypothesis testing exists within and across all elements of the Theory of Change:



It is not possible to test the entirety of hypotheses in your theory of change.

You must decide on those that are vital to success in achieving the impact goal and for which there is no existing empirical base of evidence. Conducting a literature review first will help you determine which hypotheses have already been tested.

Many other relations in the theory of change will hinge on the veracity of the tested hypothesis. One way in which you are testing your theory of change is through pathways – initiatives aligned with pathways allow you to test hypotheses in your pathways. But there are many more hypotheses in the theory of change, as indicated above, some of which may already have been tested and confirmed by you or others.

The results of testing a strategic hypothesis will inform your program strategy and may result in adjusting your theory of change (**See Annex A2.12: A Tool which outlines a process for selecting and prioritizing strategic hypotheses**).

Hypothesis

A proposed explanation for a phenomenon that establishes a relationship between two concepts.

The statement of a relationship becomes a hypothesis when it is able to be tested, by reference to observable phenomena, and confirmed or disconfirmed.

It is typically expressed, using formal logic, as an "if X, then Y" statement, whereby "X" is the independent

Be aware that this is also an iterative process that will require team input and dialogue. Like indicators, analytical thinking lies at the core of this process and requires the perspectives of gender specialists, M&E, and technical or sectoral experts. At any point in time, you may have 1-3 strategic hypotheses. Depending on resources and opportunities, more than that may not be realistic.

Strategic hypotheses take the form of if-then statements, such as:

If (independent variable):	Then (dependent variable):
Women have the capacity for collective action	Their social and political status will increase
Women improve their economic power	They will experience less social marginalization
Men are engaged in social change	The social and political status of women will be strengthened
Life conditions of women and girls are improved	The life conditions of households will improve

For examples of WEIMI Country Hypothesis selection see **Annex A2.13 Examples of WEIMI Country Hypothesis Selection**.

The strategic hypotheses developed by WEIMI COs have some commonalities, in particular for the relationship between economic strengthening and women's decision making in the household and/or the public sphere (**See Annex 2.14 Strategic Hypotheses Developed by WEIMI COs**).

Why do we need to test strategic hypotheses?

- **Hypotheses are the central piece to testing your theory of change.** This goes beyond tracking indicators to measure progress against your theory of change.
- You need to be vigilant of causal relations embedded in the TOC which, if investigated and proven false, may be a **critical barrier to success**. Sometimes the most tacit assumptions

are the ones that really need to be tested (e.g., does women's economic empowerment always lead to improved status in the household?).

- **Improving the quality of your program will depend on what you learn from hypothesis testing**. Otherwise, your theory of change is nothing more than a conceptual framework that can help guide your work but is not necessarily subject to questioning.

Be aware that hypothesis testing will only tell you whether a causal relationship holds true or not; it will not tell you whether your interventions or strategies are flawed or inadequate.

Monitoring Trends, Assumptions and Risks

Following on from the **identification of assumptions, risks and trends found in Part I of the guide**, you are reminded that these are a constituent element of measuring the theory of change. The analysis you have done should generate a list of trends to monitor periodically. You need to review your trends when there is an external trigger to repeat the analysis (e.g., a major political event) or an internal trigger, such as an annual program reviews, AOP, or design of a new program.

Use **Annex A2.15 Context Analysis Matrix** that lists trends and conduct the following inquiry each time:

- If it is an external trigger, what are the implications for your program and the theory of change?
- Do you need to update your analysis or scenarios?
- Has anything else changed amongst trends? How can these be explained?
- How is the impact group being affected or will it be affected in the future?

It is recommended that trends monitoring be incorporated in your impact measurement system, as some WEIMI countries have done.

Brief 2.3 Reflection in Action

Core Questions for Transformational Development Work

How do we approach the task of measurement for transformational development work?

The use of logic models for projects has often caused us to simplify reality and apply linear thinking to causal relations in ways undeserved. Reality is more complex and environments more unstable and unpredictable than we would sometimes like in seeking to achieve our goals. Aspiring to transformational development at the heart of which is a change in power relations increases the chances of unpredictability, non-linearity in change pathways, emergent and unexpected outcomes, and significant changes in empowerment that we may not perceive if we are not consciously tuned in to the felt changes of our impact group.

How do we make sure that our own theory of change is not going to blind us to the real process of social change taking place or to the signs of increasing tensions in social relations that may accompany positive changes in a woman's financial independence or a girl's continuing education?

Conventional monitoring and evaluation methods will not suffice. Paying exclusive attention to the measurement elements of the theory of change in the section above will also not suffice. In fact, the theory of change enables us to spell out and make transparent our assumptions about change processes so that we can learn just how social change does take place in reality, in the context in which we are operating. The theory of change can become a trap if our conviction in the possibility of a better future becomes dogma – that our own paradigm is the only viable and desirable one.

Reflection in Action

Thus, what is required to use our theory of change effectively in order to deepen our understanding of social change, be able to tell an impact story, or scale up good practices?

One approach is about developing a practice of "reflection-in-action," constantly evaluating what you are doing through your interventions while engaged "in action." It challenges us to be in a state of inquiry of testing and questioning our own assumptions while acting as a steadfast observer to what is emerging. Metaphorically, this approach is akin to a mesh that supports and underlies the data collection from the use of the indicators, breakthroughs, and hypotheses and, at the same time, catches the bits of data that stem from our own observations, dialogue with participants and partners, and the myriad "information spaces" we encounter as part of our engagement in social change work – e.g., taking part in alliance building, listening to stories in the field, organizing advocacy events, attending any number of meetings not originally in our plan. These bits of data contribute to our knowledge base and are part and parcel of the data that gets "put on the table" during a regular schedule of group reflection.

"Reflection" is an important skill or practice that needs to be developed. We know that an empowerment approach requires a high level of sensitivity to facilitate changes in cultural norms, practices, and behaviors. What are we learning as we go about this work?

It is widely noted in CARE (and other development organizations) that much of the knowledge that resides in the organization is tacit rather than explicit. This phenomenon results from staff acquiring knowledge through their own experiences or practice without an imperative to reflect on or share it. Explicit knowledge refers to knowledge independently recorded and documented.

Usually, converting tacit knowledge to explicit requires an intervention, either a facilitator or a tool that can guide the knower through a process of inquiring reflexively on his or her experience. Tacit knowledge can also be in effect when performing tasks that are based on 'know-how' we developed over time and through experience. Thus, reflection opportunities are aimed at surfacing tacit knowledge, as well as emergent processes of change that occur in the field / within the framework of our program; they are neither predicted (by indicators or hypotheses) nor prescribed.

Reflection and Knowledge Management

Reflection as a practice in establishing a knowledge management approach should be institutionalized in two ways:

- As systematic, planned events that take place on a regular basis
- As a practice or skill embedded within the way that staff and partners work

The project approach in CARE typically relied on evaluations (midterm and final) as its primary source of knowledge. Evaluations are commonly accomplished by consultants in collaboration with staff. By contrast, in a program approach, evaluations are an integral part of impact measurement but are insufficient for achieving levels of "intelligence" and "wisdom" in CARE. Reflection sessions are pivotal to this process, as they harvest a broader pool of knowledge residing in people's heads. They enable collective / shared learning belonging to a team of people. With this comes a sense of ownership around the knowledge generated from these moments of collective sharing. Subsequently, the documenting and systematic capture of the knowledge from these reflection events ensures the knowledge belongs to the organization and rises to the level of organizational "wisdom." Moreover, over the 15-year program cycle, the knowledge management cycle makes it possible for the program (CARE with partners) to engage in triple-loop learning, the process we expect to see in reviewing and adapting the theory of change. Thus, one imperative is that the reflection process be systematized and structured (See **Annex A2.16 WEIMI COs Making Space for Reflective Learning**).

Reflective Practice

A set of abilities and skills to indicate the taking of a critical stance, an orientation to problem solving or state of mind. In essence, it is a readiness to constantly evaluate and review your practice in the light of new learning (which may arise from within the context of your professional practice).

Jennifer Moon, *Learning Journals: A Handbook for Academics, Students and Professional Development.* London: Kogan Page, 2009.

Being the reflective practitioner and a knowledge broker clearly creates some tensions that can be mitigated with heightened awareness of the following:

- Our own assumptions and presumed ways of working / behaving through a process of self-reflexivity. For more information see Heron, B. (2005) Self-reflection in critical social

work practice: subjectivity and the possibilities of resistance. *Reflective Practice: International and Multidisciplinary Perspectives, 6*, 3 (341-351)
(Web link: http://dx.doi.org/10.1080/14623940500220095)

- The effects that our 'dominant' position in a set of power relations may be having on the relationships we cultivate with partners, client groups, and others
- Whose voices are being heard and reflected in processes of collecting, analyzing and presenting information, and making sure that, in our aim to measure impact, we do not fail to continuously elevate the voices of our impact groups. This means going beyond an impact measurement approach in which impact groups are merely respondent groups. For more information see Jupp and Ali (2010), Measuring Empowerment? Ask Them, *Sida Studies in Evaluation*.
(Web link: http://www.oecd.org/countries/bangladesh/46146440.pdf)
- Making claims to knowledge. Knowledge is power, and transparency is imperative – in who contributes to creating knowledge, giving credit for knowledge created, and deciding with whom you will share knowledge. CARE may legitimately carve an identity out of the knowledge, expertise, and experience that exists in the organization, but, as capital, does not belong exclusively to CARE.

Recommended Practices for Reflection and Learning

In conclusion, here is a set of recommended practices for reflection and learning processes:

Schedule a reflection event at least once a year

- include a core group who come regularly (the WE program staff, partners, impact group representatives, program quality and M&E staff) and others whose input and perspective may be especially relevant to a particular reflection session.
- Or, you may want to take advantage of an already scheduled programming event and invite participants who would not normally be present (CI Members, other program staff, etc.).

Use as material for review at your reflection session

- Data collected for indicators and hypotheses.
- Any relevant studies or reports developed by your partners that specifically relate to your impact group. This evidence may also inform your theory of change or supplement / triangulate findings from your own work.
- Any other recent, relevant material developed by your Country Office. Use your judgment.
- The experiences, observations and perspectives of those engaged in the program. It is best to encourage staff and partners to keep a diary, take field notes or even to use photos or other media as a 'memory jog;' but in their absence, rely on the 'knowledge in their heads'.
- Impact stories which can be audio recordings, print, or other media, if and when available.

Share and vet the agenda with invitees prior to the event

- This allows staff and partners to add material if they wish and to come prepared to the meeting.

Designate a facilitator who will remain a part of the process from one session to the next

- This ensures there is a common thread of inquiry.
- This does not have to be staff nor the same staff each time; it can be a consultant or partner.
- The quality of facilitation must remain constant for all sessions.

Be systematic in the inquiry you apply to each reflection session

- This may mean piloting the process before you systematize.

Appoint 2-3 notetakers

- Make use of a pre-set format.

Document each reflection session

- Develop a reporting format that is regularly applied to each session.

Your report on the reflection session, if well prepared, may be the basis for your annual program report.

ANNEXES for Brief 2.1 (**Annexes A2.1-A2.10**)

Annex A2.1: Filter Questions for Measurement

- **Objective:** To check the thinking around breakthroughs, indicators and hypotheses across program development and action.
- **Participants**: Program design team

Are you designing your measurement systems with the capacity to disaggregate all data by sex (comparing males to females), as mandated by the CARE International Gender Policy?

Here it would be best to consult the CI Gender Policy and review it before initiating your measurement system. Gender disaggregation should apply to all units of analysis. Be sure to disaggregate within the household and not merely at the level of the household head.

Do the measurement elements altogether capture changes in agency, structure and relations (CARE's empowerment framework)? It is especially important to measure change in structure and relations (i.e., not to privilege agency above the other two).

This information will help you understand the Empowerment Framework and its relation to measurement. Information has been adapted from Christina Wegs and Christine Galavotti's powerpoint presentation on "Measuring Women's Empowerment & Gender Equity for Health Outcomes (SRMH & Nutrition)."

Components of the framework:

- Agency: A woman's own aspirations and capabilities
- Structure: the environment that surrounds and conditions her choices
- Relations: the power relations through which she must negotiate her path

Women often internalize their subordinate status; may have diminished sense of their own value, rights and entitlements, so, at the agency level, empowerment involves a fundamental shift in women's own perceptions or "inner transformation."

Empowerment involves critical examination and deconstruction of unequal gender relations that are often seen as natural, but individual "transformation" and agency is not enough. It is not meaningful without the ability and power to implement and realize alternatives.

- A woman needs access to resources (social, political, financial) to enable her to exercise choice, power and control over her life.
- A woman's exercise of individual capabilities/assets is mediated by social norms that define rights, roles, responsibilities and entitlements of women and men. So we must work to shift social norms that govern social relations.
- Familial and household relations play a key role in perpetuating women's disempowerment. Over her lifetime, there may be a relative shift in power relations (i.e.

from young woman to mother) but this is not enough. What is needed is a fundamental shift in unequal power relations among men and women in society.

Here are examples from **CARE's SII** work on how outcomes are classified:

Agency	Structures	Relations
1. Self-Image; self-esteem	11. Marriage and kinship rules, norms and processes	19. Consciousness of self and others as interdependent
2. Legal and rights awareness	12. Laws and practices of citizenship	20. Negotiation , accommodation habits
3. Information and skills	13. Information and access to services	21. Alliance and coalition habits
4. Education	14. Access to justice, enforceability of rights	22. Pursuit, acceptance of accountability
5. Employment/control of own labor	15. Market accessibility	23. New social forms: altered relationships and behaviors
6. Mobility in public space	16. Political representation	
7. Decision influence in household	17. State budgeting practices	
8. Group membership and activism	18. Civil society representation	
9. Material assets owned		
10. Body health and bodily integrity		

Are changes in different units of analysis being measured - the individual, household, community, group level, and the aggregate (population level)?

Women's empowerment is a social change process that is manifest at many different levels. It is not sufficient, for example, to merely refer to change at individual level (often, agency indicators). Change is almost invariably embedded in relations which a woman has with household members, community, with other women in a collectivity, and as part of society. These must all be taken into consideration.

Research on measuring women's empowerment tells us that in fact few standard indicators measure the interaction between individuals/households and larger community, district or state. Also most of the case studies on women's collective action for social change are descriptive with few standard measures.

Population level measurement often uses proxy indicators for WE, such as participation in education, labor force, age at first marriage, political representation.

Are you measuring both process and outcome of women's and girls' empowerment? How are you disaggregating this and measuring equivalent measures in men and boys?

Both process and outcome are equally important to understanding achievements in and threats to women's and girls' empowerment. And often the "process" to achieving specific outcomes will differ for the population or sub-population group. The risks (e.g., of increasing women's economic empowerment) may be higher for some groups than for others. At the same time, this is what makes measuring women's empowerment and the reason it is important to apply a mix of qualitative and quantitative data gathering methods.

Women and girls' empowerment, however, cannot be measured exclusively; it needs comparison with men and boys' empowerment for gender equity to be achieved.

Are you capturing both breadth (scale) and depth of impact?

Even with reporting on indicators, evidence can sometimes be anecdotal with reference to a specific success case or possibility of change, for example. It is important that programs demonstrate systematic measurement and that means being able to show breadth or scale of the social changes – for how many women is this true? AND depth – how much improvement? What is a significant level of change?

This data is essential to demonstrate that a specific model of change works or to convince policymakers that the impact achieved is truly significant. Efforts at policy change must proceed from credible evidence.

Are you capturing the essential changes among key stakeholders in government, civil society and the private sector reflected in our theory of change?

Where reference is made in your theory of change to any of these actors, check to see that you are actually measuring changes in their behaviors and decisions. And do not neglect the private sector, as CARE is more accustomed to working directly with civil society and government.

It may help to construct a matrix showing what the essential anticipated changes are for each category of actor.

Are you including men and boys in measurement of behavioural change at target group level?

Inclusion of men and boys is pivotal to success in achieving gender parity. As such, they become very important target groups in any program. Try to identify within your particular country context which men and which boys are important to target and the behaviors for your behavior change strategy.

Are there any expected changes in the overall theory of change that are not being captured and may actually be slipping through the measurement cracks?

This is a final check on your theory of change, after you have selected breakthroughs, indicators, hypotheses, and macro trends. Go back to the theory of change and ask yourselves what you have neglected.

Click here to return to Brief 2.1: Breakthroughs and Indicators

Annex A2.2: WEIMI Country Breakthroughs

CARE Burundi

Burundi reframed breakthroughs as progress markers, milestone events that enable you to know that you are moving in the right direction. As a result the CO has defined a set of 23 breakthroughs at pathway level for the four DOCs of the WE program which include a combination of breakthroughs referring to one-off events (legislation drafted and adopted) and attainment of critical thresholds (e.g. the breakthrough referring to the implementation of Nawe Nuza at scale). The SMIC document proposes that these breakthroughs will be monitored based on the use of participatory methodologies with the impact and target groups. The draft format developed for the operationalisation of this part of the SMIC has not yet been operationalised. However, the breakthroughs have been used by the CO for strategic planning purposes in the development of the FY 2013 and 2014 AOP.

CARE Mali

The team developed a plan for measuring breakthroughs that included the initiatives contributing to them, the changes they expect will be induced by breakthroughs, the operational definition for these, the data sources, the method for data collection, the frequency and who. See table below for illustration of how this is done. They have 7 breakthroughs for the PEF program and these are included in the measurement strategy. In their formulation of breakthroughs, the anticipated change would need more precision as a critical threshold (all means of production, for example?).

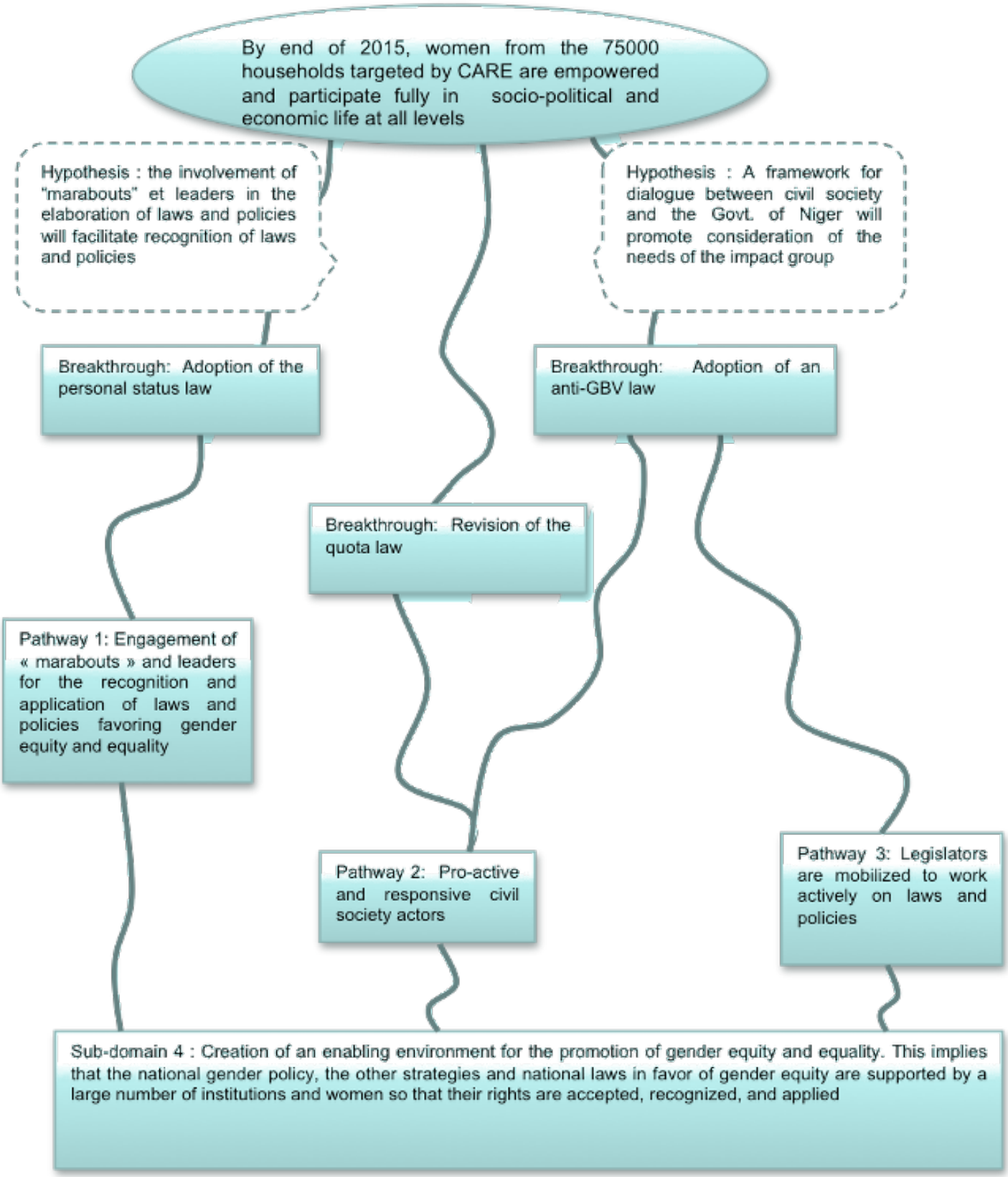
Mali has a matrix with a calendar of when breakthroughs are expected to be attained. Achieving a breakthrough is through pathways. This achievement is not as linear as several pathways may contribute to achieving a breakthrough.

CARE Niger

Initiatives (projects) are linked to the theory of change via the breakthroughs, i.e., their contribution to breakthroughs. The team plans to annually review progress of initiatives in aligning with and achieving breakthroughs. Niger identified 1-3 breakthroughs for each domain of change in the LEFF program cutting across several (sometimes all pathways), although, at a Dec. 2011 workshop on its CO impact measurement system, participants seemed to understand that the formulation of many of the current breakthroughs do not fit the definition. (One example of a breakthrough is women own and control productive resources, but this does not have the level of specificity to know when you have reached the breakthrough).

Niger intends to build evidence around these breakthroughs through the accumulated contributions of initiatives. Each initiative has been assessed for its level of contribution (if any) to a breakthrough, ranked as weak, medium or strong. The CO has developed a tool with IM standards that includes monitoring the alignment of initiatives to breakthroughs. Each year, the individual program will assess this and at CO level, the same will be assessed for the CO TOC.

See diagram of pathways, breakthroughs, etc. below:


By end of 2015, women from the 75000 households targeted by CARE are empowered and participate fully in socio-political and economic life at all levels
Hypothesis : the involvement of "marabouts" et leaders in the elaboration of laws and policies will facilitate recognition of laws and policies
Hypothesis : A framework for dialogue between civil society and the Govt. of Niger will promote consideration of the needs of the impact group
Breakthrough: Adoption of the personal status law
Breakthrough: Adoption of an anti-GBV law
Breakthrough: Revision of the quota law
Pathway 1: Engagement of « marabouts » and leaders for the recognition and application of laws and policies favoring gender equity and equality
Pathway 2: Pro-active and responsive civil society actors
Pathway 3: Legislators are mobilized to work actively on laws and policies
Sub-domain 4 : Creation of an enabling environment for the promotion of gender equity and equality. This implies that the national gender policy, the other strategies and national laws in favor of gender equity are supported by a large number of institutions and women so that their rights are accepted, recognized, and applied

Breakthrough	Changes induced by breakthrough	Key elements to arrive at the changes	Key interventions
Women have secure access to means of production (land, equipment), actes de cession, titles, etc. **Relevant Initiatives: Advocacy, engagement of men and entrepreneurship**	• Application des conventions • Mobilization initiative of GMJT/RMJT for access to the means of production • Funds invested in the purchase of inputs for members • Increase in production • Food security • Improvement in income • Improved nutrition • Vulnerable groups have sufficient and balanced food • Strengthening of woman's decision making power in the household, community, reproductive health, management of household goods, etc. • Lessening of prejudices (inheritance, etc.)	• Engagement of men (community leaders, property owners…) • Mobilization of GMJT/RMJT who need access to land • Les GMJT invest in access to inputs • A critical mass of women, impact group members of the CBOs participate in decisions made at municipal, regional and national levels (X%), example : Joint Committee on Land Management	• Watch for synergy with other CARE programs linked to means of production • Watch for synergy with other interventions linked to means of production • Investments related to the RA results, with production plots (gardening, farmland, etc.) • Training / information on questions of land tenure and application of the laws
Data sources	**Method for data collection**	**Frequency**	**Responsible Person**
MJT groups and network	• Self assessment of engagements in the conventions • Monitoring of mobilization initiative for access (# of groups, types of access, areas acquired, etc.) • Monitoring of the use of income from GMJT/RMJT	Quarterly	Junior experts and Area Coordinators

Click here to return to Brief 2.1: Breakthroughs and Indicators

Annex A2.3: WEIMI CO Impact and DOC Indicators

Category	Indicators
Poverty Reduction	**Burundi:** • % of women (per IG members and NON IG members) living in a household with less than $1 per day **Mali:** • # of households with resilience capacity in the face of crises and conflicts • # of vulnerable and marginalized women who control their own means of production **Tanzania:** • % of households that do not rely solely on agriculture for their livelihood by sex of the head of household **Egypt:** • Engagement in economic activities- formal sector (gender-disaggregated) • Employment in the government and private sectors (gender-disaggregated) • Ratio of female participation in the labor force (disaggregated by governorate) • Engagement in economic activities – vocational (gender disaggregated) **Niger:** • % of households living under the poverty threshold (MDI+)
Food Security	**Mali:** • % of young adults from vulnerable households who are improving their food security, particularly the nutritional status of children under 5 years in their respective households
Education	**Mali:** • Enrollment rate of girls in primary school compared to that of boys (MDI+)
Women's Empowerment	**Egypt:** • % Change in women's participation in CSOs • % of CSOs led by Women • Prevalence of marital violence (any, ever) by governorate for all upper Egypt • Law on domestic violence is passed and implemented • Personal status law is modified **Burundi:** • % of women (per IG members and non IG members) who report a quality participation in decision making structures at community level • % of community structures with effective representation of at least 30% of women • % of women (per IG members and NON IG members) who claim their basic rights • % of women (per IG members and NON IG members) who cover their basic needs with dignity • % of women (per IG members and NON IG members) who report having access to quality financial services • % of women reporting an improvement in their psychosocial well being **Niger:** • % of women elected in their communities relative to the total number of men and

	women elected **Mali:** • # of elective positions in communities and civil society occupied by women or girls • # of trained girl mothers who are practicing profitable socio-professional activities • % of women and girls satisfied with NGO services of the program **Tanzania:** • Average number of hours per day spent on housework and in retention to the duration of the working day, by sex and age **Bangladesh:** • % of IG women reporting meaningful participation in collective action initiatives undertaken by solidarity groups at the community level and beyond • % of men reporting positive attitudes towards the collective action initiatives of women's solidarity groups at the level of the community and beyond
Child Health and Nutrition	
Maternal Health	**Niger:** • Proportion of births attended by skilled birth attendants (MDI+)
Sexual and reproductive health	**Niger:** • % of couples making informed joint decisions regarding sexual and reproductive health
HIV/ AIDs	
Environmental Sustainability	
Governance	**Tanzania:** • # of effective spaces in which the IG participates meaningfully (planning, budgeting, monitoring) **Burundi:** • Coordination between the government and civil society in favor of poor and vulnerable women • Number of laws and policies adopted / ratified (conventions, protocols, resolutions, declarations) which promote women's rights **Mali:** • # of NGO consortium members who have an operational management system on performance, talents, and knowledge

Click here to return to Brief 2.1: Breakthroughs and Indicators

Annex A2.4: CO Process to Prioritize Indicators

CARE Tanzania

"Just as we learned several lessons with regard to the development of the TOC, so we learned several lessons with regard to the development of the impact measurement and learning system. Originally, we selected about 64 indicators that seemed right and suitable for our Mwangaza [WE] program. But when we started to think deeply about the way we were going to operationalize them, we realized that it was going to be too complicated. Instead, we decided to select 20 indicators. But WEIMI helped us to see that even 20 indicators were just too many. So, WEIMI helped us choose 7 indicators from CARE's list of Millennium Development Indicators plus (MDI+) for our impact goal and broad outcomes."

The group struggled with finding an appropriate impact indicator. It was felt that an indicator associated with food security would be the most meaningful, but neither of the two MDI+ impact indicators for food security – proportion of the population below minimum level of dietary energy consumption, which is extremely difficult to measure and use, and the indicator, prevalence of underweight children under 5, was a clear fit for its impact goal Marginalized and vulnerable women and girls, at critical life stages, in rural under-served and environmentally restricted areas are empowered to live sustainable, healthy and secure lives. The group settled for the latter, as this would also speak to assumptions about changes in availability of food and income (and women's knowledge), improved access to health services, changes in women's decision-making at household level, and sustainable use of natural resources in pursuit of food and income security.

CARE Niger

This indicator was chosen by CARE Niger in relation to the hypothesis that "an improvement in income would mean economic empowerment and self-efficacy of women."

Niger chose two indicators at impact level, for its LEFF program goal—social equity is improved for women and girls of 75,000 households by their increased capacity of expression and organization, the equitable distribution of capital and assets.

To see how it measured self-efficacy, a description follows. It should be noted that self-efficacy measures agency; it does not measure relations or structural change.

Operational definition: a woman is assumed to have personal efficacy if she demonstrate socio-political and economic efficacy. Economic efficacy was defined as the ability to have an income level sufficient to (i) be able to make viable economic choices for her and her household, (ii) satisfy her economic needs and contribute to those of her household, (iii) develop her capacity to cope with economic shocks (resilience). Socio-political efficacy is recognized in (i) the visibility of a woman's leadership, (ii) the ability to propose, organize and help implement actions for the development of the community and (iii) the ability to make informed political choices.

Calculation of the indicator: the indicator is a composite and is assessed on the basis of an overall score, using criteria (variables) each ranked low, medium or high (1, 2, 3, respectively). A total score was calculated as the sum of all scores on all criteria. The level of personal efficacy is based on the following scale:

Low	Medium	High
8-13	14-18	19-24

The indicator was calculated on the basis of statements and perceptions made by women about themselves and by their husbands and their perceptions of what level of self-efficacy their wives had attained.

Factor analysis performed on responses by both men and women yielded an extraction of two factors with a total of 8 variables. The two factors cumulatively accounted for 56.4% of the variance in the entire set of variables in the case of women's responses and 51.39% for men's responses.

Factor 1 : economic security capacity of households (IGA and income)	**Factor 2 : participation in decision making, in public life**
• Technical skill for undertaking IGA • Profitability of IGA • Capacity to anticipate risks linked to IGA? • Level of satisfaction of economic needs from income (own and contribution to household needs) • Which strategy are you adopting to secure IGA when confronted with economic shocks?	• Level of information regarding major events in the community? • What is your level of participation in forums and level of influence on decision making in your community? • Choice of candidates during elections

CARE Niger concluded that both economic efficacy and socio-political efficacy were validated as definitional for self-efficacy. It confirmed to CARE Niger the need to prioritize this indicator within the women's empowerment framework, primarily through the LEFF program. CARE envisions a target of 10% increase by 2015 for the high and medium levels of self-efficacy. The advantage of factor 1 is that it can be measured more objectively than relying upon perceptions (of women or men) and is thus a more reliable predictor variable for self-efficacy.

Click here to return to Brief 2.1: Breakthroughs and Indicators

Annex A2.5: Indicator Prioritization Tool

- **Objective**: Prioritize indicators at goal and DOC levels, particularly once a brainstorming session has generated too many indicators
- **Participants**: Program design team

CARE Burundi selected a set of 10 program-level indicators for measuring changes at the impact and domain levels of the TOC for the WEP by scoring all potential indicators against a set of agreed criteria. This process was considered by program staff to have been very useful as a way of reaching agreement on a manageable/ limited number of higher level indicators. "Options" have been added for this guidance.

It is a form of a "values clarification exercise" that helps the team rank the indicators.

1. Develop criterion for each 'column':
    - Weight the criteria if you consider one to be of greater importance than another. In this case, assign a weight to each criterion of 1 to 5 (low to high importance).
2. For each indicator, mark a score in each of the columns using the criteria at the top of the column – allowable scores are 0-3, where 3 is the very highest/strongest possible.
3. For every place where you rate an indicator at "3" (highest level), include some comments about why you rate it that high in that column.
4. When done, total across the rows, and identify the highest indicators at goal, and domain (use the pathway level ratings to check if any of the pathway indicators might be better than the existing domain indicators).

*Other criterion to consider: existing source of data, i.e., does not have to be gathered by CARE.

Explanation of the column headings in the worksheet						
Column 1	**Column 2**	**Column 3**	**Column 4**	**Column 5**	**Column 6**	**Column 7**
Link with the MDIs/ MDGs	**Link with CSLP/ PRSP**	**Experience within CO**	**Feasibility**	**Relevant**	**Reliable**	**Understandable**
The indicator already in the MDI list? Or could it be extracted from one of the MDI indicators	Linkage to indicators in the national poverty reduction strategy	Has anyone in the office, or any project, used this indicator so that they know something about any special issues in collecting data for it or in analyzing it	Cost effectiveness- can it realistically be done in the local circumstances?	How well does it demonstrate what is happening in this goal or domain?	Can it be done by different persons over the life of the program and get comparable results?	Is it simple enough to be used by/ with communities; and can the results be explained in a meaningful way and used by various stakeholders?

The example below is from Burundi's Children Empowerment Program – group 1 (indicators 1-15). This is a partial list; the full set was 7 pages of indicators.

Logical hierarchy/ chain	Objectively Verifiable Indicators	1. Link with MDIs/ MDGs	2.Link with CSLP/ PRSP	3.Experience within CO	4.Feasibility	5.Relevant	6.Reliable	7.Understan-dable	Total Score
Impact Vision: By 2015, orphans & children vulnerable to abuse, to denial of rights, aged from 6 to 18, living in economically vulnerable households, particularly girls, effectively enjoy their rights, live in an enabling environment, & see their future with hope & confidence	Level of fulfillment of OVC rights	1	1	2	1	2	1	1	9
	% of OVC who claim considering their futures with hope and confidence	0	0	1	0	2	0	0	3
	% of households hosting OVC & living under $1/day16	3 (written in the MDGs)	2	1	2	3	3 (calculate according to the revenue of the HH)	2	16
	% of children (disaggregated per sex) who have easy access to quality services	2	2	2	3 (appreciation of facilities provided)	2	1	1	13

Click here to return to Brief 2.1: Breakthroughs and Indicators

Annex A2.6: Indicator Tips

- **Objective**: Tips on how to construct good indicators, to balance rigor in measurement with practical considerations. Indicators should be sparse and the choice strongly rationalized.

Use existing indicators to the extent possible; avoid inventing new ones

Coming up with good indicators is a difficult task that requires various technical skills and expertise. It requires the person designing the indicator to be familiar with methods to operationalize concepts; methods to weigh variables in parsimonious models; and so on. Because CARE is not a research institute but an implementing development organization, we often do not have the expertise—nor do we need—to design indicators in-house. Instead, we should employ existing, proven indicators developed by experts, rather than try to invent new ones ourselves. If we use existing indicators, we will benefit from the expertise and experience of those who know how to develop them and have already done the work; we will save ourselves the trouble of having to test the indicators; and we will avoid having to provide lengthy explanations and justifications when communicating our work externally.

Sometimes, we are so focused on the details of our own work that we think we cannot find existing indicators to help us measure what we need. However, that is rarely true. If we make a commitment to exhausting the options of existing indicators first, we will not need to invent new ones—and, ultimately, our work will be much higher quality and defendable as a result. For example, assume that we are working with an impact group that is exploited by selling advance labor during seasons when the price of labor is low, then having to perform the labor during seasons when its price is high. This exploited group obviously does not benefit from the seasonal increase in the price of labor, and we are interested in measuring this economic exploitation. We could start developing a very complex index, or a composite indicator, including various aspects of economic exploitation. Or, we could select a very simple indicator that has been already developed and tested numerous times by economists: seasonal wages. We need to measure wages in the different seasons for our impact group and for a group that does not experience the same exploitation (e.g., nationally, for a higher caste, etc.). The difference in wages between those who sell their labor expensively when its price is high, and those who sell it cheaply even when its price is high, is an excellent indication of exploitation.

Use simple indicators. Do not think that a complex concept requires a complex indicator.

Like many answers to questions, often the simplest makes the most sense. Similarly, the best indicators are the simplest. They tend to be straightforward, easy for everyone to understand, and minimize the potential measurement error. They also enable us to focus on analysis and learning rather than on complicated measurement. If we encounter a complex concept or problem, as we often do when we work on marginalization, empowerment, exploitation, rights and so on, we need to break those complex concepts down until we can find simple indicators for them. It is always better to simplify a thought and attach a simple indicator to it than to try to design complex indicators to measure complicated thoughts.

Composite indicators are inappropriate for analyzing complex processes. When we are dealing with a complex problem, such as economic exploitation or marginalization, we are tempted to design an

index to measure this problem. In general, indexes are useful for predicting macro-level trends. For instance, if the government wants to predict the movements of the economy in the next few months, it will have its statistical institute calculate a composite index of many different indicators, including employment, manufacturers' production, value chains, financial changes, interest rates, and consumers' behavior. This approach is not appropriate for our purposes.

To continue with the previous example about vulnerable people exploited by village moneylenders, imagine that we wanted to measure economic exploitation. This is a very complex and multi-faceted problem. Instead of doing the conceptual thinking to simplify what we need to measure, as we did in the example above, we could start to develop a very complex indicator for economic exploitation, which would likely take the shape of a composite index. A composite index of indicators combines various indicator components to produce one overall number that can be tracked over time. Our index of economic exploitation could contain an indicator measuring whether those borrowing money are extremely poor; one that measures whether they s ell advanced labor; one that measures whether they are aware of the prices of labor each season; and still another that measures whether they take loans from the village moneylender. We could then attempt a complex calculation to aggregate all these variables into a composite index measure. As the following explains, using composite indexes of indicators is not a good approach for our work for a variety of reasons:

☑ First, a composite index approach requires statistical technical expertise that we do not have to conduct the work rigorously.

☑ Second, the approach requires a lengthy and expensive measurement.

☑ Third, the approach would not tell us anything that a simple indicator like the one discussed above – interest rate on loans – does not already tell us.

☑ Fourth, composite indexes are not useful for explaining complex processes, such as what we do when testing our TOCs and analyzing our contributions to social change. If one thinks of indicators as once-removed from reality (e.g., there is "real life," then there is an indicator to reflect real life on paper), then indexes are twice removed from reality--think of them as reflections of reflections.

☑ Fifth, constructing a complex indicator such as an index is difficult, demanding statistical work – one cannot simply decide to add a few indicators together. Even the selection of the components, which will be part of the index, needs to be tested statistically. The calculation of measures to produce the final number on the index is difficult statistical work as well.

In summary, complex indicators such as composite indexes are to be avoided. If you must use one, do so with extreme consideration and care, and always use an existing one. However, it is best to analyze and break down a complex problem or issue to its component parts, then assign a simple indicator to each part. A complex problem is best captured and analyzed with simple indicators.

💡 ***Use one indicator per change you seek to measure. Do not use multiple indicators for the same concept.***

This is why you must be clear what the critical changes are in your goal or domain of change. Identify these before you begin your work on indicators.

If we survey the different areas in which we work, we will find hundreds of indicators for each. For example, we can find numerous indicators for education: enrollment rates, completion rates, ratios of teachers to students in the classroom, and many others. This does not mean, however, that we need to measure all of them. Using as many indicators as we can think of in our data collection tool does not improve our measurement; it only makes it longer, more costly, and more burdensome. Instead, we should focus on selecting one good indicator for the underlying change we want to see, and concentrate our resources on measuring and analyzing that.

For example, our global menu of outcome and impact indicators lists five different indicators in the area of Environmental Sustainability. Our job is to figure out which one would best capture the change we are working towards in our program. We do not need to include more than one of those five indicators. A good detailed analysis of one of them will go further than lengthy measurement on all five.

Give due consideration to the financial, time and skill implications of indicators. It is acceptable to opt for a cheaper, easier to measure indicator.

When we have gone through the dos above (i.e., done our conceptual thinking, identified potential existing indicators, thrown out the complicated and selected the simple ones, and tried to narrow them down to one), we are sometimes still left with choices among good indicators. How do we select one among many equally good options? We should seriously consider indicators based on whether they are less expensive to measure than others, would take less time, or would require fewer specialized technical skills. For example, imagine that we need an indicator for food security and that we have two options: caloric intake and dietary diversity, both disaggregated by sex. Which one do we choose? Caloric intake is extremely difficult to measure and calculate, as it requires difficult precise recall on the part of survey respondents. It also requires an expensive survey done by people with special technical skills to gather the data and do the necessary calculations. And it would also require a long lead time to organize the resources and people needed. Dietary diversity, on the other hand, is a good proxy for food security: people who have higher dietary diversity are less threatened by food insecurity, because those who diversify their diet over time have escaped the threat of food insecurity. Dietary diversity is also easier to measure for both survey enumerators and respondents: the recall involved here is much less demanding than would be the case in trying to calculate calories. Because this is a simpler indicator, it would also require less skill, time and money to measure. Based on these practical considerations, we choose dietary diversity as our indicator, and leave the measurement and calculation of caloric intake to the DHS and other national surveys done by research institutes.

Focus on analysis; do not think that an indicator, however good, is a substitute for good judgment and good thinking.

Indicators are designed only to facilitate our analysis and thinking, not to replace them. Even if we select the perfect indicator but then do not analyze the data we collect on it, reflecting on what it is telling us, then we have not done our job. Indicators are not short cuts for thinking. They are simply a tool to help us collect the data and do the analysis that should inform our thinking. Further, indicators only indicate. That is, they point in the direction in which change has taken place. They are not a substitute for our work – we need to do the analysis of whether the change is positive or negative, expected or unexpected, as described by our TOC or not, large or small. Indicators are

good for data collection, but the core of our task with impact measurement is analysis. We have not done IM until we complete our analysis, interpret the information derived from the analysis, and draw out its implications for our work.

To illustrate how collecting data on an indicator is different from analyzing and interpreting data, consider the following example:

Imagine a project to reduce violence against women and girls, and that some of the main interventions include raising community awareness and encouraging people to report instances of such violence and seek legal redress. The number of instances of reported violence against women and girls could be an output indicator for this project. It is very likely that, between the baseline and the mid-term, the number of instances of reported violence would increase; our data collection would show the numbers going up. If we simply collected the data and ended our effort before properly analyzing those data, we would have no idea what this increase meant. It could mean that more violence against women and girls is taking place, or it could mean that the amount of violence has not changed, but the increased community awareness means more instances of violence are actually reported, rather than going unrecognized, as they did before awareness was raised. It could also mean that the amounts of violence and awareness have not changed, but more people had been empowered to report instances of violence. We cannot know which of these changes has taken place simply by saying that the number of instances of reported violence has increased. We will understand the change that has taken place only if we analyze our findings in their context, compare them to findings on other indicators, and draw conclusions from this analysis.

Having a good indicator does not, in and of itself, replace taking the time to analyze the collected data, which reveal very little without an analytical context. The key, therefore, is to do the in-depth analysis to understand the meaning of the findings and act accordingly.

Click here to return to Brief 2.1: Breakthroughs and Indicators

Annex A2.7 Example of Outcome Indicators and Operational Definitions from Bangladesh

Domain of Change: Exercise of greater choice affecting decisions in their lives

Indicator: % men and women reporting meaningful participation of women in public sphere.

Operational Definitions

Public Sphere: This is defined as women's involvement in both formal and non-formal organizations at, especially, community level.

Examples of formal organizations include:

- School Management Committees;
- UPs and UP standing committees;
- Registered organizations (clubs, community based organizations);
- Local chapters of political parties;
- Religious groups and committees;
- Natural Leader Organisations.

Examples of non-formal organizations include:

- Village Development Committees;
- Solidarity Groups/Self-Help Groups;
- Community Support Groups;
- Market Management Committees;
- Savings Groups;
- Other community groups for supporting development initiative activities (including EKATA groups).

In addition to involvement in organizations, public sphere includes women's participation in public forums, such as Salish and community meetings.

Meaningful Participation: This is defined as women's physical presence in committees/meetings AND their participation in organizations and processes, as demonstrated by the following steps:

1. Raising their voices in meetings;
2. Influencing the views of others;
3. Adding value to decisions by shifting them in favor of the Impact Population.

Other considerations: will need to also ascertain men's description/opinion of women's involvement and participation, using these 3 questions.

Click here to return to Brief 2.1: Breakthroughs and Indicators

Annex A2.8 MDI+ Indicators

MDI+ Indicators Selected by 1or More WEIMI Countries

The ✧ indicates a use of slightly different wording. Indicators not used by any countries are shaded.

	MDI INDICATOR	TAN-ZANIA	EGYPT	BANG-LADESH	NIGER	BUR-UNDI	MALI
1	Proportion of the population below $1 a day					✔	✧
2	Prevalence of underweight children under 5 years of age	✔			✔		
3	Proportion of the population below minimum level of dietary energy consumption						✧
4	Net enrolment ratio in primary education						
5	Proportion of pupils starting grade 1 who reach last grade of primary school						
6	Literacy rate of 15-24 year olds, men and women						
7	Ratio of girls to boys in primary, tertiary and secondary education			✧			
8	Share of women in wage employment in the non-agricultural sector						
9	Change in women's self-efficacy				✔		
10	Under 5 mortality rate						
11	Infant mortality rate						
12	Maternal mortality ratio			✔			
13	Proportion of births attended by skilled health personnel either at home or at a health facility				✔		
14	ANC and PNC (at least 1 visit and at least 4 visits)						
15	Contraceptive prevalence rate				✔		
16	HIV prevalence among the population aged 15-24 years						
17	Adolescent birth rate						
18	Unmet need for family planning						
19	HIV prevalence among population aged 15-24 years						
20	Condom use at last high risk sex						
21	Proportion of population aged 15-24 years with comprehensive correct knowledge of HIV/AIDS						
22	Proportion of population using an improved drinking water source				✔		
23	Proportion of population using an improved sanitation facility						
	MDI + INDICATOR						
1	% people whose net income generated within target value chains has increased, by sex						
2	% households with access to secure land tenure, by sex of the head of household	✧	✧	✧	✔		✧

	MDI INDICATOR	TAN-ZANIA	EGYPT	BANG-LADESH	NIGER	BUR-UNDI	MALI
3	% households that do not rely solely on agriculture for their livelihood, by sex of the head of household	✔					
4	% households with capacity to cope with environmental shocks without depleting assets, by sex of the head of household	✔			✔		✔
5	% change in dietary diversity (to measure both nutrition and as a proxy for biodiversity)						
6	% children 0-23 months who are underweight (weight for age), by sex						
7	% households with food reserves, by sex of the head of household				✔		
8	Primary school completion rates by formal (gov't funded)/non-formal schools, by sex						
9	Primary education programs gross enrollment rates by formal (gov't funded)/non-formal schools, by sex						
10	Student-to-teacher ratios by formal (gov't funded)/non-formal primary schools						
11	% women reporting meaningful participation in decision making at household level in a domain previously reserved for men	✔		✔			✔
12	% men and women reporting meaningful participation of women in the public sphere	✧	✔	✔	✔	✔	✧
13	% men and women with changed attitudes toward gender-based violence	✔	✔	✔	✔		✧
14	% couples making informed joint decisions regarding sexual and reproductive health	✔			✔		
15	% men and women reporting ability of women to effectively control productive assets	✔	✧	✧			✧
16	% women reporting an improvement in their psychosocial wellbeing					✔	
17	Average number of hours per day spent on house work, and in relation to the duration of the working day, by sex	✔					
18	% children exclusively breast-fed within the first 1 hour after birth and up to age of 6 mos.						
19	Proportion of 1-year-old children immunized against measles						
20	% children age 0-23 months who slept under an insecticide-treated bed net in the previous 2 weeks						
21	% women with met need for emergency obstetric care						
22	% women attending 4 ANC visits at a health facility						
23	% women reporting satisfaction with the quality of care received						

	MDI INDICATOR	TAN-ZANIA	EGYPT	BANG-LADESH	NIGER	BUR-UNDI	MALI
24	% people whose need for family planning services is met, by sex						
25	% people making informed decisions about their contraceptive use, by sex						
26	% adolescents with access to contraceptive methods						
27	% people with correct knowledge of the 2 major ways of preventing the sexual transmission of HIV/AIDS, by sex						
28	% adults with more than one partner in the past 12 months reporting the use of a- condom during last sexual intercourse, by sex						
29	% people with advanced HIV infection receiving antiretroviral combination therapy, by sex						
30	% people utilizing HIV-associated clinical services such as VCT, STI, TB, by sex						
31	% local actors with meaningful participation in productive natural resource management at community level, by sex	✔			✔		
32	% of population using an improved drinking water source				✔		
33	% of population using an improved sanitation facility						
34	% of population with access to emergency warnings						
35	%of population using improved fuel sources						

Click here to return to Brief 2.1: Breakthroughs and Indicators

Annex A2.9: Impact Indicators for WEIMI Countries

Bangladesh	Burundi	Egypt	Mali	Niger	Tanzania
IMPACT GOAL					
The 'most' socially, economically, politically marginalized women are empowered	By 2025, poor women, aged from 18-50, vulnerable to gender-based violence, from rural households with less than 0.5 hectares of land property, without access or control over other sources of income have regained their dignity and fully enjoy their basic rights	By 2025, poor women in Upper Egypt are empowered, enjoying a better quality of life attaining their rights	By 2015, vulnerable and marginalized women and girls (15-49 years) in the regions of Mopti, Segou, and Timbuktu will have increased their economic power, their socio-political status and participation in decision making while maintaining social cohesion at regional and national levels	Social equity is improved for women and girls of 75,000 households by their increased capacity of expression and organization, the equitable distribution of capital and assets (gender-sensitive curriculum at school, a positive perception of the roles and status of women and girls, household capacity in adapting to climate change, food availability, efficient supply systems and management of food stocks, continuous access to good quality water)	Marginalized and vulnerable women and girls at critical life stages, in rural underserved and environmentally restricted areas are empowered to live sustainable healthy secure lives
IMPACT INDICATORS					
Maternal mortality reduced significantly in extremely poor and most marginalized women	% of women (per IG members and non IG members) reporting improvement in their psychosocial well being	% men and women reporting meaningful participation of women in the public sphere	# of women and girls who are empowered in their choice and decisions affect the quality and the security of their livelihood	Prevalence of underweight children under 5 years of age	Prevalence of underweight Children under 5 years of age

At least 30% of marginalized women completing secondary level of education	% of women living on less than a dollar (PPP) a day	% of women first married by age 18, for women age 25-49 and for women age 20-24	# of social norms, policies and laws that have changed in favor of promoting the rights of women and girls	Change in women's self efficacy	
Chronic malnutrition among most marginalized women reduced by 80%	Maternal mortality rates	# of women in parliament	vulnerable and marginalized women with an income above the economic poverty threshold ($ 2/ day)		
		# of girls in Upper Egypt in age group 0-17 who have been circumcised			

Click here to return to Brief 2.1: Breakthroughs and Indicators

Annex A2.10 CO Baselines

CARE Mali

Mali benefitted from the fact that its long-term CARE Norway program "Empowering Mothers and Girls" was being developed at the same time as the WE program TOC and measurement system. This a 7-year initiative with major contributions to the PEF program with a 7-year impact goal.

Mali conducted a baseline by focusing on "common convergence areas" contributed by at least three CARE interventions. It took a representative sample of villages that will serve as the panel villages to be followed up during the life of the program. The choice of these villages is random and is based on the statistical method of two-stage sampling: the choice of villages which is proportional to the size of households in the villages and the selection of households that are to be drawn randomly after the census household in the village.

CARE Burundi

Burundi plans to do a program-level baseline as a separate/ distinct data collection exercise from any project-level baseline. The timing of the proposed program baseline has yet to be finalized but it is expected to take place at some point over next fiscal year, with follow-up survey to track impact and domain-level changes at 5 year intervals. The baseline and follow-up surveys are expected to be funded by development of cost-pool in new project design processes topped up by UNR and/or specific research funding.

In the meantime, the program-level indicators are being used to inform the design of project reviews and evaluations, and new projects are being designed around the TOC, incorporating the pathway and domain-level indicators as appropriate.

CARE Niger

Niger collected primary data by teaming up with the National Institute of Statistics. The study took baseline measures for all three programs which compose the CO theory of change. Thus, it included the WE program for which there are 2 impact indicators. It also included hypotheses to be tested. The baseline study was a survey with a sample of 1660 men and women and 66 interviews with community leaders in one of two target geographic areas.

Click here to return to Brief 2.1: Breakthroughs and Indicators

Annex A2.11 WEIMI CO Pathway Indicators

CARE Tanzania

The team, with support from WEIMI, chose either MDI+ indicators or ones closely associated with the PRSP for Tanzania, for all levels of the theory of change. Each pathway has one corresponding indicator.

Example:

Pathway: Cultural and religious norms supports gender equality

Indicator: % couples making informed joint decisions regarding sexual and reproductive health

CARE Burundi

Burundi has 30 indicators for the 16 pathways. These pathway-level indicators include indicators in use by existing initiatives, which have been mapped in terms of their alignment to the pathways. There has already been some harmonization of initiative-level M&E systems with the WE program, in that some initiatives are using common tools/ formats (e.g. Kirumara & Girijambo for monitoring VSLA, also work with community activists) albeit with their own project-specific indicators. To date however, these project-level tools are not yet linked to/ contributing to measurement of program indicators, in part because most of the current portfolio of initiatives was designed prior to agreement of the program indicators.

Examples of pathway indicators:

Pathway	Indicator
Engagement of men and boys in women's empowerment	• % of men who have changed their attitudes in favor of women's empowerment
Putting in place an accountability mechanism for the allocation of resources for poor and vulnerable rural women	• # and quality of mechanisms and/or government structures established for poor and vulnerable rural women
Diversification of economic opportunities	• # and type of economic initiatives developed in favor of women • % of women who have diversified their revenue sources • A gender-sensitive national policy on small and medium-sized enterprises and initiatives adopted

CARE Egypt

While at DOC level, data is primarily obtained from secondary sources, at pathway level, initiatives contribute data to the pathways. Egypt has not yet begun to collect data and its measurement system is yet to be tested.

Pathway	Indicator
CSOs build, support and reflect a constituency of poor women	• # of new initiatives designed and implemented by CSOs responding to poor women needs and challenges
There is acceptance of gender equity and equality at all levels (HH, community, national)	• # of men and boys who are involved in addressing GBV and discrimination against women

Click here to return to Brief 2.1: Breakthroughs and Indicators

ANNEXES for Brief 2.2 (**Annexes A2.12-A2.14**)

Annex A2.12: Tool for Selecting and Prioritizing Strategic Hypotheses

This process is based on the lessons learned from the experience of WEIMI countries (primarily Tanzania and Mali) and an adaptation of the prior instructions prepared by the Program Impact team.

Steps

1. Examine, one at a time, the relations between:

(a) Pathways and domains of change

(b) Domains of change

(c) Domains of change to impact goal

Guided by these questions:

- What are some of the major assumptions you are making here?
- How are they supported or not by empirical evidence in your programming or other literature?
- Is the evidence substantive enough for you to continue to make this assumption without presenting a risk to your program? If not, then continue with the process.

2. List the hypotheses from step 1 that are critical to the success of your program and which ought to be tested.

3. Break each hypothesis down further to reveal any sequence of cause-effect, making sure you are not simply drawing on your own logical thought process but on your knowledge of what has been attempted and/or observed in programs – this can apply to CARE's own or others.

Example from Mali:

Supposition2: The capacity of collective action by women increases their social and political status

Hypothesis: if a critical mass of networks are strong, while women are mobilizing deliberately in favor of their rights, if women mobilize deliberately for their rights, then they will influence the socio-cultural barriers and standards, if they influence barriers and socio cultural norms, then they have more [space] in the spheres of decision making and local community.

4. Do another review of your hypotheses using this set of questions:

(a) Are you certain that the hypothesis in all its constituent parts has not been tested before? List sources of information that would furnish evidence. If you are not sure it has been tested, you may need to first start with a literature review and come back to this hypothesis later with greater precision of what needs to be tested. When you consult other sources, make sure to include sources from within your country context, as what is true in one locality may not be true for another.

(b) When you are sure the hypothesis has not been tested before, consider again whether the

hypothesis has strategic importance to your theory of change. Imagine the effect it might have on your program if the hypothesis is disconfirmed when tested, e.g., if "collective capacity of women did not increase their social and political status."

(c) In the scenario that evidence does exist to confirm your hypothesis, you may actually be more concerned with the magnitude and pace of change. In this case, you may want to simply monitor progress as you would in tracking change for "within-pathway" hypotheses. To wit, it may not require a piece of research to test your hypothesis.

5. Once decided that a set of hypotheses are indeed "strategic" and need to be tested, do another final review to make them ready for testing:

(a) Be specific in your hypothesis, as it applies to your empirical context – what institutions, what issues, which policies, etc. Specificity is necessary to be able to measure your hypothesis.

(b) Identify for whom within your impact group this is a relevant hypothesis, i.e., to whom does it apply? (this is also part of being specific)

(c) Does it need to be more context-specific – a population belonging to a particular geographic location?

(d) Does your entire team agree on the formulation of the hypothesis? Do you now have a sound rationale for testing this?

6. Drop any hypotheses from your list that do not meet the review criteria above. With your final list of hypotheses that ought to be test, you will now prioritize these according to these guiding questions:

(a) Which are urgent for the program, having relevance to the next five years?

(b) What are opportunities to conduct the research (in a project evaluation or baseline; or a special study funded out of multiple initiatives)?

(c) What are then potential sources of information for this research (e.g., your own initiatives)?

(d) Do you have the funds, skills, and resources to do this OR is another organization better-placed? Will you pay them for conducting the study?

7. Once you conclude which hypotheses are feasible and meet the criteria, then you need to develop a funding / resourcing strategy and a timeline for the research. It is also wise to try and interest other organizations in co-funding and benefiting from the endeavour. Decide whether you need additional expertise to develop the research framework.

Below is an excerpt from CARE Tanzania's workshop to select hypotheses. Out of a list of 5 candidate strategy hypotheses, the team prioritized this one for testing:

Hypothesis 1: WE and Economic Development: Women and adolescent girls in VSLA groups build the self-esteem, social and economic capital that enables/shifts in gender roles in HH and community, which makes them to raise their voices and participate in the community decision-making.

Discussion:

This is an active assumption in much of CARE's work – but has it been proven?

Seems to qualify as a research hypothesis.

In this context, we're looking at the generation of cohesion that leads to women's ability to make decisions in community and HH. (e.g. "which gender roles" are we looking at? Division of labour, control over assets?) Will need to unpack this further – what are the critical roles/norms that we're shifting? Need to narrow it down to items that are comparable and compatible. We may actually be talking about "gender relations" than "gender roles" since we're talking about the relationships within the HH/community.

In Burundi, VSLA groups are an entry point for other work – we start with the VSL groups to mobilize women/bring them together

Another dimension: VSL groups are not uniformly focused – e.g. a focus on money, a focus on building solidarity, a focus on social issues - does the difference in focus lead to difference in outcomes? EG in Bdi, they call the group a Solidarity Group.

This might be a reason to focus on "decision-making" because the decision may be context-driven, but it's the change in women's decision-making power/role that we're interested in.

The hypothesis asks: does participation of women & girls in VSL build social capital and self-esteem? Does participation enable them to play a different role in HH, community, society? In some of the VSL groups, there's been increase in GBV – so, if participation in VSL builds self-esteem, social capital, economic capital, we need to understand why. If there are negative effects (e.g. GBV), we also need to understand why.

There might be a need for us to look out for other organizations collecting information on similar hypotheses. There's an opportunity to join with others.

If we're able to collect useful data and analyze it, one element we want to look at is: so what?? How does this help build impact at scale? How widely is that information shared – how much influence will it have on our work and the work of others.

After another workshop, the team formulated the hypothesis, as follows, and then proceeded to test it as part of a project evaluation:

Women and girls in VSL groups build the self-confidence, social and economic capital that enables changes in gender relations in household and community, which makes them to voice their concerns and opinions and participate in the community decision making.

Click here to return to Brief 2.2: Strategic Hypotheses and Trends

Annex A2.13 WEIMI Country Hypothesis Selection

CARE Tanzania

"[In terms of process], we first identified all of the assumptions embedded in our TOC, some coming from the relationship between pathways, some from the relationship between pathways and broad outcomes, others from the relationship between broad outcomes, and still others from the relationship between the broad outcomes and the impact goal. In all of this, we did get assistance from WEIMI. Most of the work was done in specific WEIMI meetings/workshops. The first was an 8-day meeting in late March and early April in 2011. Then, several meetings convened specifically around TOC followed and other meetings convened specifically around assumptions and hypotheses followed between April and December 2011.

The selection of hypothesis was almost as long as the formulation of the TOC. We started identifying assumptions in August and worked on developing hypotheses until September. After that, there was still some confusion on identifying and defining hypotheses. So, we asked for advice from WEIMI. In November, we spent 5 days with WEIMI team and consultants on formulating the correct hypotheses."

1. Tanzania had undertaken a rigorous and detailed process for identifying hypotheses. The steps it undertook, with WEIMI guidance, are as follows:
2. Identify all possible pre-conditions and assumptions in the theory of change (e.g., within pathways, between pathways, between pathways and DOCs, between DOCs, and between DOCs and goal)
3. Through a group think process, identify a set of critical assumptions (12 were identified. See example below).

For each assumption, explain the logical flow and then translate the assumption to a hypothesis.

To select hypotheses for testing from among the 12, pose these filter questions:

(a) Is there information available about it already?
(b) Is it strategic to the program TOC?
(c) Is it feasible?
(d) Reasons for selection

One of the learnings from CARE Tanzania is that the process was too time-consuming and the guidance herein has been adapted with this in mind. Example of converting assumption to hypothesis:

Assumption	Logical Flow	Draft Hypothesis
Institutions responsive and accountable to the needs and rights of IG leads to IG's participation in and benefits	Responsive and accountable Institutions lead to realization of the needs and rights of the IRG in NRM	If institutions become responsive and accountable, then the needs of the IG will be met and their rights will be realized.

from improved sustainable NRM (DOC3-DOC4) (12)		
	Realization of needs and rights of IG leads to transparent interaction and information sharing	If the needs of the IG are met and their rights are realized, then the interaction between the institutions and IG will become transparent and information will be shared between them.
	Transparent interaction and information sharing with responsive and accountable institutions lead to IG's participation in and benefit from improved sustainable NRM	If interaction between the institutions and IG becomes transparent and information is shared, then the IF will participate and benefit from improved and sustainable NRM.

CARE Mali

Mali's intent in developing hypotheses was to test the TOC and improve quality. The team identified hypotheses based on where evidence was needed. They refer to hypotheses at different levels (domains and the breakthrough level) broken down into four types:

- Type 1: those with evidence based on causal relationships
- Type 2: those who do not have clear evidence, but are key to our theory of change and shown to be rigorously tested
- Type 3: those who have no evidence, but are key to our theory of change and will be tested by appropriate methods that take into account cost and effectiveness
- Type 4: those who do not have clear evidence, but will only be followed periodically by contextual analysis

Obvious hypotheses do not require extensive research to demonstrate the changes. On the other hand, other types of hypotheses will be verified rigorously for the life of the program using both quantitative and qualitative methods. Thus, based on the assumptions below the hypotheses have been formulated.

Supposition1: Improving the economic power of women reduces their social marginalization. ***Hypotheses***: (1) if there's more economic opportunities (financing, to productive resources, knowledge and markets) for women, then they are more likely to seize these opportunities (2) if they are more likely to participate in these economic opportunities, while their contributions to household income will increase, (3) if their contributions to household income increases, then they will have the esteem of household members, (4) if they have the esteem of household members, they shall be consulted and participate more in decisions of the household.

Supposition 2: The capacity of collective action by women increases their social and political status
Hypothesis: if a critical mass of networks are strong, while women are mobilizing deliberately in favor of their rights, if women mobilize deliberately for their rights, then they will influence the socio-cultural barriers and standards, if they influence barriers and socio cultural norms, then they have more [space] in the spheres of decision making and local community.

Supposition3: The engagement of men for social change strengthens the social and political status of women
Hypothesis: if the government implements policies favorable to the rights of women (Family Code), if the standards change, if men know women's rights, and men and women's roles and responsibilities, then the women will be empowered in taking decisions relating to aspects of her own life.
To be test-able (or measurable), any hypothesis needs to be broken down into constituent parts so that all causes and effects are apparent and can be researched, as Mali has done.

CARE Niger

Niger developed several hypotheses (12) and tested one of these in the baseline study. The hypothesis is:

- Joint and informed decision making on sexual and reproductive health among couples could increase women's access to quality SRH services, notably births attended by qualified medical staff and the use of modern contraception.

In the baseline, hypotheses are linked to indicators. Two of the indicators measured in relation to the above is % of births attended by qualified medical staff and the contraceptive prevalence rate.
Some cannot be tested until some change has occurred, such as:

- If the conditions of life of women and girls are improved, then the conditions of life of households will improve.

Most of Niger's hypotheses are between pathways and the domain of change.

Niger intends to test other hypotheses as opportunities and funding present themselves.

CARE Tanzania

After generating a short list of testable hypotheses, the team agreed, in its final meeting, to only test one hypothesis, which reads as following:

Women and girls in VSL groups build the self-confidence, social and economic capital that enables changes in gender relations in household and community, which makes them to voice their concerns and opinions and participate in the community decision making.

The team felt it was better to test one and do it properly, to learn from this experience. It tested this hypothesis as part of the CARE-Norway funded Women and Girls Empowerment Mid-Term Review. Results are not yet available.

Hypothesis on Male Engagement

"Men's constructive engagement" is one of components of its WE program. The program conducted action research, using an innovative and participatory approach to identify, engage and motivate men called "agents of social change and behavior." These men serve as "models" and help others to behave in ways favorable to the emancipation of women and girls particularly with regard to access and control of farmland and decision making related to use of reproductive health services. See M. Aliou Bah, 2011, Responsable Initiative Men Engage. Rapport Monographique: Recherche Action sur les normes et pratiques sociales relatives aux femmes / filles. CARE Mali.

Burundi developed a strategic hypothesis and series of research questions relating to the link between men's engagement and women's empowerment. Some of these questions have since been explored in a recent study of the CARE Burundi Abatangamuco men engage initiative carried out by CARE Norway in partnership with a local policy research centre. See REF Abatangamuco report.

Click here to return to Brief 2.2: Strategic Hypotheses and Trends

Annex A2.14 Strategic Hypotheses Developed by WEIMI COs

Bangladesh	Burundi	Egypt	Mali	Niger	Tanzania
f women have access to resources and services, they will be able to influence decision making processes, reducing the frequency and propensity of violence against women	Engaging men by providing them information on SRH and changing their attitudes on masculinity will result in a change in their behavior and attitudes which enables the empowerment of women	Economic strengthening will contribute to a reduction in gender based violence, through a reduction in domestic violence (not yet finalized)	(1) if there's more economic opportunities (financing, to productive resources, knowledge and markets) for women, then they are more likely to seize these opportunities (2) if they are more likely to participate in these economic opportunities, while their contributions to household income will increase, (3) if their contributions to household income increases, then they will have the esteem of household members, (4) if they have the esteem of household members, they shall be consulted and participate more in decisions of the household.	(tested) Joint and informed decision making on sexual and reproductive health among couples could increase women's access to quality SRH services, notably births attended by qualified medical staff and the use of modern contraception.	(tested) Women and girls in VSL groups build the self-confidence, social and economic capital that enables changes in gender relations in household and community, which makes them to voice their concerns and opinions and participate in the community decision making
If non threaten environment created across all levels (hh, community, market) for women and girls that supported by men will give more		There is a relationship between economic strengthening and increasing women's involvement in public sphere (not decided	If a critical mass of networks are strong, while women are mobilizing deliberately in favor of their rights, if women mobilize deliberately for their	The re-establishment of roles attributed to women and men trigger significant changes in norms incompatible with gender	

opportunity to exercising greater choice in women's lives		yet)	rights, then they will influence the socio-cultural barriers and standards, if they influence barriers and socio cultural norms, then they have more [space] in the spheres of decision making and local community	equity/equality.	
If women are competent (knowledge, attitude, skill & education) to engage in livelihood opportunities and the support of their household then women can involve in more productive work			If the government implements policies favorable to the rights of women (Family Code), if the standards change, if men know women's rights, and men and women's roles and responsibilities, then the women will be empowered in taking decisions relating to aspects of her own life.	Improvement in women's financial resources leads to greater consideration and participation in community life and in the level of women's autonomy	
If women's mobility increased and they can engage in income generating activities, allowing households to become economically solvent and will be freed from dowry				The more women control resources, the more they are independent (self-reliant) and participate in economic life.	
Building solidarity among women can transform negative behaviors and attitudes				An improvement in income would imply economic independence (self-reliance) and the self-	

of men, engaging in increased women's engagement in both private and public spheres				efficacy of women.	
				The improvement of women's capacity in decision making facilitates their active participation.	
				Involving marabouts and leaders in formulating laws and policies will facilitate recognition of those laws and policies.	
				The construction of a common vision among CSOs around questions of gender equity and equality will facilitate the creation of a strong alliance	
				up of literate women, men and youth, then the institutionalization of good governance of natural resources at local level will be assured (tested).	
				If the conditions of life of women and girls are improved, then the conditions of life of households will improve.	
				A consultative framework between	

				civil society and government will enable the consideration of the needs of impact groups.	

Click here to return to Brief 2.2: Strategic Hypotheses and Trends

Annex A2.15 Context Analysis Matrix

- **Objective:** Conduct a contextual analysis that can be repeated periodically to monitor trends, reduce risks, and adjust assumptions affecting the theory of change
- **Participants:** Program design team

Step 1:

Identify the aspects of the context to be analyzed, keeping in mind that contextual analysis is NOT the same as UCP/V analysis. In the analysis, include implications for each of your Impact Groups.

- Regional context, plus global trends as applicable
- National context
- Local context

For each,

- Social, political, geo-political, cultural dimension
- Demographic dimension
- Economic context, including NRM
- Environmental context
- Development space context, including donors

Step 2:

Prepare context analysis matrix, summarizing findings for each of these aspects and identifying key trends (highlights of the analysis) to be watched over time. The analysis should also include considerations of future scenarios/forecasts, as relevant:

	Local	National	Regional/ Global	CARE CO: Implications for Impact Group
Political				
Economic				
Social				
Cultural				
Geo-Political				
Environment				
Development				
Trends: 1._____; 2._____; 3._____				

Step 3:

Identify critical elements of the contextual analysis and trends applicable to each program

Step 4:

Prepare individual context analysis matrix for each program reflecting the critical elements identified for that program

Step 5:

Identify trigger factors for the contextual analysis. Triggers of analysis are events or processes that signal to you that an analysis of the context, or an update of your contextual analysis, is needed again:

→ Internal to CARE

- LRSP process (review or development)
- AOP (review or development)
- Design of new program
- A key or large assessment process related to impact tracking of changes toward impact vision

→ External to CARE

- Major political change
- Major shift of strategic donors (priority, policy, focus)
- Major change in government plans, priorities, etc.
- Etc.

Step 6:

If any of these triggers of analysis occurs, conduct a contextual analysis, keeping in mind the following criteria:

- Select an appropriate method for the analysis and apply it rigorously
- The first time a trigger occurs, do a complete detailed analysis, as defined in the context analysis matrix. The second, third, etc. times a trigger occurs, focus on those areas that have experienced significant change since the last analysis, only updating those cells of the matrixes that need updates.
- Reflect on the trends you have identified for tracking
- Review forecasts/future scenarios and update as needed
- Enter the updates the analysis has produced into your matrix

Step 7:

Conduct a stakeholders analysis, considering the following questions, and reflect the summary of this stakeholder analysis in cells of the matrix as relevant:

- Actors, networks, movements
- Agendas
- Resources

Step 8:

Continue updating the matrix as needed, keeping it a living document:

- Update cells if a change has occurred
- Enter new trends
- Save a new copy of the matrix every time, instead of changing one and the same copy, so you have a record of changes in the context over time and institutional memory
- Keep the matrix to 1 page so it's user-friendly and used
- Make discussions of the context a regular feature of staff meetings

Click here to return to Brief 2.2: Strategic Hypotheses and Trends

ANNEXES for Brief 2.3 (**Annex A2.16**)

Annex A2.16: WEIMI COs Making Space for Reflective Learning

Burundi

Tradition of Learning

It has been that the case in Burundi that all studies and evaluations undertaken by the CO are required to generate lessons learned/ recommendations, which are then used to identify themes for more in-depth reflection/ study, i.e. program-level learning initiatives (e.g., a study of marriage registration and implications for WE). The CO also has a biannual joint planning process at program level, which includes mapping of evaluations and studies to ensure coordination. After each evaluation there is a restitution meeting which involves program staff in the identification of lessons learnt, and those lessons learned are compiled into a single document for grands approaches such as VSLA or Men Engage. During the biannual program planning meetings, the program team also decides when to hold learning events (2-3 each year). In the past year two they held a National Forum on VSLA and National Forum on WE, with a learning event on Community Scorecards scheduled for late May 2012. By making ready use of these existing planning and discussion forums, the WE program will have a strong foundation for developing a systematic approach for the testing of the TOC.

Another interesting offer from CARE Burundi is a Learning Plan (Plan d'Apprentissage) which includes identifying resources, activities and a timetable for 5 areas:

1. Collaboration among peers: Foster collaboration between the staff within initiatives and programs and between programs and initiatives and program support.
2. Documentation and sharing of information and key products: Developing a culture of knowledge creation and sharing it to ensure that good information is available at the right time, in the right format and to the right target (internal and external).
3. External relations and partnerships: Establish mechanisms to promote and contribute to the development of beneficial learning with other organizations.
4. Tools and Technologies: Develop / create a website, the mechanisms of online discussions and other tools and technologies to support and promote interaction and learning between people especially on topics for which it is difficult to organize meetings.
5. Skills and Behaviors: Develop and promote the skills and behaviors that are necessary for people to learn and share knowledge across initiatives, programs and organization.

Triple Loop Learning

Under the subsystem "Learning and Management of Programmatic Approaches," Burundi introduces concepts to promote individual and organizational learning. The aims of this subsystem are 1) continually learn about the actions of initiatives and approaches, 2) establish the link between situational analysis and the differences in understanding the context that we are called to further study and explore, 3) to increase collaboration opportunities for individuals and groups - exchange of tacit and explicit knowledge between individuals (project teams, including partners and neighbors), program teams, partners and other stakeholders to use the knowledge, 4) document lessons learned, best practices for availing into information (centralized management system of knowledge / evidence), 5) ensure the quality of approaches and learning models.

Characteristics of the single, double and triple loop learning (adapted from Bruce Britton March 2005, Berrett-Koehler, 2004)

	Single-loop Learning	Double-loop Learning	Triple-loop Learning
Description	The single-loop learning can be conceived as a means to be improved the way established working rules and procedures in an organization are applied in practice. It is often said that it is thinking inside the box. It asks questions about the "how" but almost never more fundamental questions relating to "why." The single-loop learning assumes that the problems and their solutions are closely related to each other in time and in the space. In this form of learning, we first consider our own action. Minor changes occur in practice and/ or specific behaviors. These changes are based on what worked or did not work in the past. This calls to do things well but not necessarily discuss and/ or challenge our beliefs and assumptions. The goal is to improve and fix that will often take the form of a procedure or a rule. The single-loop learning often ends in most cases to simple adjustments such as how to use a thermostat to regulate temperature. Are we doing things in the way recommended? In proceedings? In the rules?	The double-loop learning is often called thinking outside the box, as it allows questioning assumptions and principles underlying the rules and procedures. The double-loop learning focuses on the effectiveness of a solution. Why does such a solution work? In this form of learning we consider our actions in a framework with well-defined hypotheses and assumptions. This is the level where self-analysis process intervenes where learners become their own observers, wondering: What is working well here? What can be considered as an approach or model? Hence, we change how decisions are taken while trying to deeply understand our hypotheses and assumptions. The double loop learning works to reach changes and/ or improvements such as straightening the function or structure of an organization. Are we actually doing the things recommended? that's why it works, approaches, and models. The double-loop learning can have important consequences and can lead to so-called triple loop learning.	Triple loop learning – challenging assumptions and principles of organization that could have lost their relevance due to changes in the organization and its environment. This requires an often exchange of views and often aggressive. The triple-loop learning involves principles. Learning takes place through approaches and models adapted to the context. The result creates a shift in the understanding of the context or the way we see things. We formulate and produce new commitments and new ways of learning. This form of learning challenges us to understand how problems and solutions are inter linked even if they are separated in time and space. This form also challenges us to understand how past actions have created the conditions that guide our current problems. The relationship between organizational structure and behavior is a fundamental change because the organization learns how to learn! The learning outcomes include the strengthening of the ways to understand and change our goals, develop a better understanding of how to respond to our environment and understanding of why we do the thing we do. How do we decide what is

			correct? Why do we do this thing - principles.
Conditions	Apply established rules/ procedures Treat the symptoms of problems. To reflect the constraints.	To question the rules/ procedures Analyze the underlying causes to the problems. To reflect outside constraints.	Analyze the core values and identity. Rethink the purpose and principles. To reflect in the constraints
Expected Results	More efficient working methods. Better application of rules/ procedures	More efficient working methods Knowledge and new ideas, rules, and procedures Improved systems and strategies	New improved statement of core values and goals New identity

Mali

Mali's first reflection workshop since its baseline was guided by the question: "what are the key changes made by the program since its intervention, that is to say during the last 2 years?" This workshop took place in August 2011 in Tombouctou and Mopti and attended by junior experts and zone coordinators as well as those responsible for program components participated.

Process:

A workshop was conducted in each convergence zone (an area of at least 3 initiatives). The methodology was participatory with group exercises. The different stages of the workshop was as follows:

Step 1: The participants provided, through brainstorming, data on all key results achieved by the program and successful strategies.

Step 2: Then the results were analyzed and classified as changes or not.

Step 3: The changes were described and information gaps were identified for the group to complete by consulting the impact group.

Step 4: The changes were classified by program component.

Step 5: Qualitative interviews were conducted with the impact group to complete the missing information through an interview guide.

Advantages of the approach:

This approach allows documentation:

- To have a complete description of all changes
- To have a complete count of all changes
- Changes documented by field staff of the program

Disadvantages of the approach:

- Difficult to track unexpected changes
- And the high cost of tracking

This document produced some interesting results and each result examined by the group generated more questions for follow-up which suggests the usefulness of this type of inquiry process.

Mali has also established a series of face-to-face events for tracking change:

Quarterly meetings - all relevant initiatives will come together on a quarterly basis, allowing each initiative to share its contribution with clear evidence of changes and any impacts. The team would also analyze the strategies that were effective in achieving these changes. This is based on information collected, analyzed and documented.

Biannually there would be M&E workshops to assess changes. The team would gather and document changes by initiative. The data gathering would be based on clear evidence. This includes a learning space to take effective strategies to scale.

Program Team Meetings would also take place biannually, consisting of all team members from the ground level up to the program director. This is an opportunity to share results and progress towards the desired impact.

Annual Review Meeting with the entire program team. It allows a critical reflection on the changes made and their consistency with the theory of change and contextual analysis of these changes.

Annual Report Workshops with impact group members and other development actors - these workshops allow participation and increased accountability of partners and communities in the implementation of M & E plans and intervention strategies to ensure that real changes are tracked and highlighted. They are organized on an annual basis with impact groups on the results achieved, the changes perceived by them, failures and their reasons. This will allow organizing all information into a statistical database.

Niger

Niger's reflection forum joins up the M&E officers / Unit and the Advocacy-Communication-Gender-Partnership Unit bi-annually. Together, they plan to institutionalize an Impact Measurement working group to facilitate implementing the measurement strategy. They aim to develop a set of tools that will enable them to measure impacts and effects for the CO theory of change. They will also be developing a calendar of activities, roles and responsibilities for the M&E Impact Measurement Initiative. Still to be determined are the type and frequency of meetings and the reports to be completed for their knowledge management system. This is a work in progress.

Click here to return to Brief 2.3: Reflection in Action