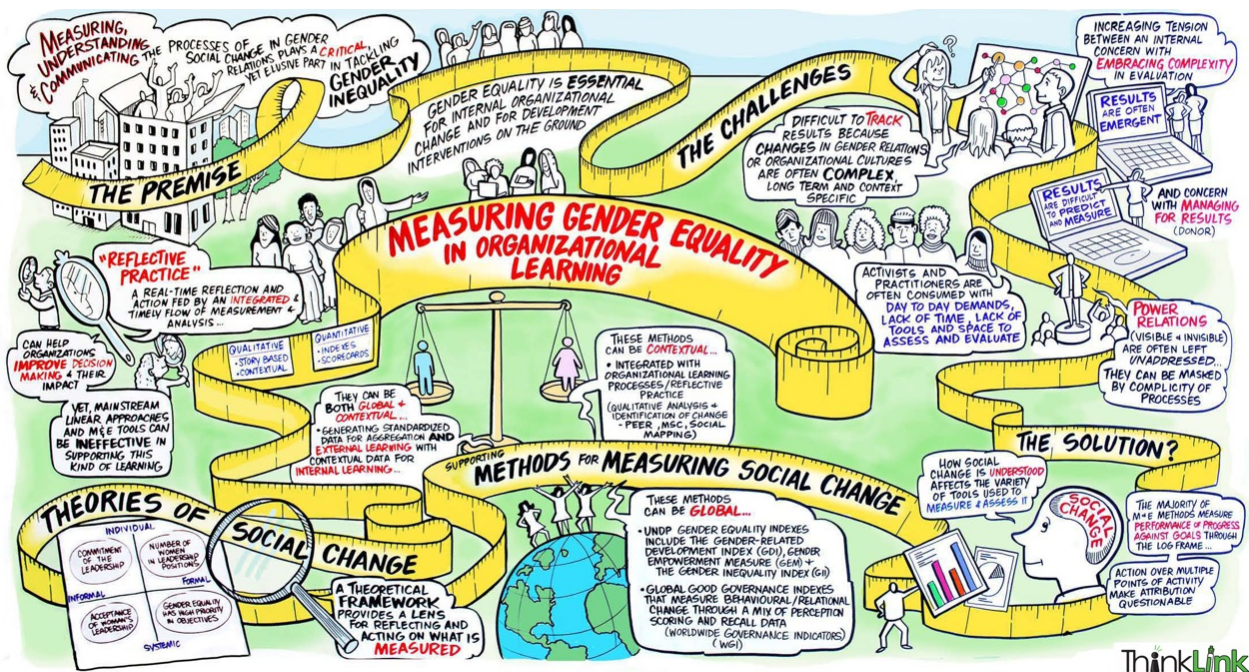


# Measuring Gender Equality Initiative

## Workshop Report

Launched on 19<sup>th</sup> and 20<sup>th</sup> May 2011  
 At Dunford House, West Sussex, UK  
 Gender at Work and IDS



ThinkLink  
 GRAPHICS  
 thinklink@graphics.com



## Table of Contents:

1.	<a href="#">Summary</a>	4
Day One		
2.1	<a href="#">Challenges and Outcomes</a>	7
2.2	<a href="#">Presenting the Playing Field: Measuring Gender Equality in Organisational Learning</a>	9
2.3	<a href="#">Small Group Work: discussing self-identified challenges</a>	11
2.4	<a href="#">Closing comments at the end of day one</a>	12
Day Two		
3.	<a href="#">Addressing key interests through an open agenda</a>	13
3.1	<a href="#">Making Sense of SenseMaker®: plenary discussion with Jeremy</a>	13
3.2	<a href="#">Self-selected clinics</a>	14
3.2	<a href="#">Closing summary</a>	16

## Annexes:

<a href="#">One</a>	Participants and their contact details	19
<a href="#">Two</a>	Visual landscape of Background paper	20
<a href="#">Three</a>	Presentation of Background Paper	21
<a href="#">Four</a>	Gender at Work Analytical Frameworks	32
<a href="#">Five</a>	Suggested reading on SenseMaker®	35
<a href="#">Six</a>	Useful websites	48

## List of Visuals:

[2.1:](#) Participants' photo, and L-R opening comments with Rosalind Eyben (IDS), Madhavi Kuckreja (Gender at Work), Jennifer Laing (The Ant), Michelle Higelin (WYMCA), Jeanette Kloosterman (Oxfam NOVIB)

[2.2:](#) Graphic landscape of meeting background paper by Holland and Sheppard

[2.3:](#) Small group work: L-R; Isadora Quay (IDS), Hannah Sheppard (Gender at Work), Michelle Higelin (WYMCA), Aruna Rao (Gender at Work), Geetanjali Mishra (CREA), Alex Pittman (AWID), Rosalind Eyben (IDS), Michel Friedman (Gender at Work), Jenny Bell (Justice and Women)

[3.1:](#) Group exercise, Tai Chi led by Michel Friedman (Gender at Work)

[3.3:](#) Participants working in clinics on day two, L-R: Erica Zwaan (Cordaid), Jenny Bell (Justice and Women), Michelle Odayan (Indiba Africa Group), Geetanjali Mishra (CREA), Alex Pittman (AWID), Rachel Moussie (Action Aid), Esther Benning (Oxfam NOVIB), Ratna Sudarshan (ISST)

## **1. Summary:**

Changes in gender relations are often complex and context specific and organizations seeking transformative changes in gender relations can struggle to describe, track and communicate adequately the results of their work.

The Measuring Gender Equality Initiative addresses this by bringing together two pressing areas of work which define Gender at Work's unique focus – (i) Association for Women's Rights in Development's (AWID) leading work on understanding, enhancing and deepening feminist evaluation techniques, and (ii) the Institute of Development Studies' (IDS) work on participation, power and social change – in the sphere of measuring changes in gender relations in organizational and development learning.

This initiative was launched in collaboration with the Participation, Power and Social Change (PPSC) Team in the Institute of Development Studies at the University of Sussex, and funded by NORAD, the Global Fund for Women and UN Women. It took place during May 18-19, 2011 and brought together 26 participants, from 12 countries South and North, representing organizations working at all levels of international cooperation for social change from women's organizations, university settings and donor agencies (see [Annex One](#) for list of attendees).

At the heart of the initiative was a desire to initiate a 'community of learning' by bringing together donors and activists at all levels of the system. In this collective space, the initiative aimed to review and examine frameworks, methodologies and instruments to measure changes in gender relations in organizational development and learning.

To this end, participants collectively identified and interrogated specific challenges, alternative methodologies, institutional constraints and power dynamics which affect organization's capacities to both capture and learn from changes in gender relations both internal to the organization and as a result of its programming work.

The meeting's starting point was a set of challenges and tensions that needed to be addressed:

- When compared to other sectors, it is particularly difficult to track results because changes in gender relations or organizational cultures are often complex, long term, and context specific.
- There is an increasing tension between an external (donor) concern with "managing for results" and the evaluation community's concern with "embracing complexity" in



evaluation, by engaging (rigorously) with complexity and accepting that results are often emergent rather than predictable and easily measurable.<sup>1</sup>

- Activists and practitioners are often consumed with the day to day demands of making change happen so lack the time, space or tools to assess what has worked and why. It is very difficult for organisations to integrate monitoring, evaluation and learning in a culture of rigorous evaluative practice.
- Despite the difficulties, practitioners also express the conviction that measuring and communicating results more effectively can help us better understand how to bring about change and make better decisions about where to focus our efforts. To become more successful, we both need and want to improve our monitoring and learning practices.
- The promise of monitoring, learning and evaluation methodologies is nevertheless driven by many different theoretical, practical and organizational interests and priorities. Knowledge building and sharing, accountability, governance, communications, funding, strategies, alliances, all shape our expectations of this endeavour.

Participants' essential reading included, *Measuring Gender Equality in Organisational Learning: A Background Paper* (Holland and Sheppard, Gender at Work, 2011), [Capturing Change in Women's Realities: A critical overview of current monitoring and evaluation frameworks and approaches](#), (Batliwala and Pittman, AWID, 2010), [Assessing and Learning for Social Change – a discussion paper](#) (Guijt, IDS, 2007) and chapter 3, 'Methodological Issues in Measuring Empowerment' in [Empowerment in Practice – from analysis to implementation](#) (Alsop, Bertelsen and Holland, World Bank, 2006).

From this shared starting point, the meeting's key asset was the diversity and creative tension in the room. Through its open agenda, participants created opportunities to develop a new network, share challenges with peers, and appreciate differing perspectives as well as to call upon expertise in the room to address the issues identified.

The diversity in the room revealed itself in manifest ways and highlighted that this is not a depoliticised process –each having different perspectives and varying capacities to engage with the methodological and political agenda. Offsetting the differences within the room, there was a common interest among participants in reviewing methodologies and exploring

---

<sup>1</sup> "Evaluation Revisited: Improving the Quality of Evaluative Practice by embracing Complexity", A Conference on Evaluation for Development, May 20-21, 2010, Utrecht, The Netherlands  
<http://evaluationrevisited.wordpress.com/>

more deeply the tension between organisational learning and external accountability. There is also interest in how to build organisational learning itself through M&E processes.

Two invaluable in-depth sessions on SenseMaker® and Most Significant Change (MSC) brought concrete examples from peers' experiences, enabling participants to explore the implications of power and exclusion in the design, implementation and communication of learning and measuring gender equality and organizational learning.

A virtual community of practice has tentatively been instigated. In part the meeting confirmed the importance of peer-learning mechanisms as a space to work through issues, yet, looking ahead, without the funding for a second follow up meeting as initially planned, participants have begun instead to build upon the learning already contained within AWID's extensive [wiki](#) which contains reviews of more than 50 commonly used tools and methodologies with critical analyses of their strengths and limitations, and the potential for people to add their adaptations and experiences with the aim to create an online learning community to enhance and strengthen future and ongoing practice.

The meeting galvanized a desire amongst many participants to deepen their understanding of the terrain, and experiences of pilots have subsequently been posted online, and discussion topics opened for debate among practitioners and evaluators – in synergy with AWID's work and online space for personal development.

Participants have given overwhelmingly positive feedback; the open and the participatory claiming of agendas through clinics were highlights as process high points as well as other spaces and moments of connection.

In relation to participants' own continuing work in measuring gender equality and organizational learning; there was a new awareness to the power dynamics in M&E, and participants gained ideas for looking at its application within their own organizations. This is part of the broader process of participants connecting M&E to 'the bigger picture' and its links with advocacy, policy and community learning and reflective practice.

## 2. Day One

### 2.1 Challenges and Outcomes



*L to R: Initiative participants; and discussing the self-identified challenges.*

Gender at Work's facilitators, Michel Friedman and Ray Gordezky, led the group through an exploration of their self-identified challenges in assessing the complex terrain of measuring and assessing changes in gender equality.

These disclosures and subsequent conversation on the **challenges** faced confirmed AWID's most recent findings from its work with gender activists questioning the measurement and accountability methods used for capturing their contribution to change in women's rights and gender equality.

These included:

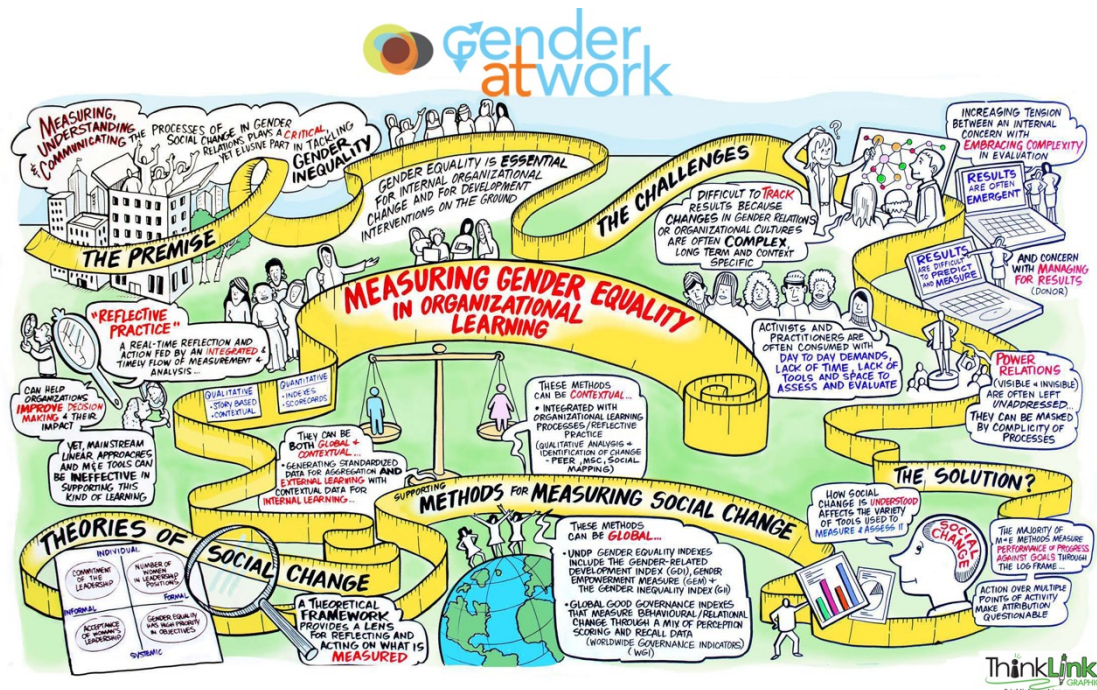
- **The importance of locating yourself re other actors:** who is doing what and where to question M&E practices in ways that transform unequal power dynamics? Expressing a desire to build a community of learning beyond this meeting.
- **Understanding more about how to change the culture of learning:** explore the tension between M&E as a learning exercise vs. making NGO's accountable; and how to better use M&E for organisational learning.
- **Informational needs around specific methodologies:** what are the most effective evaluation approaches and methods to assess the complexity of gender equality and capture non-linear processes of change? How can participatory methods be used to establish indicators at local levels with women's groups? How to access the deep, embodied stories/underlying rules that influence what gets said, when? Specifically in the terrain of cultural issues which are unsafe to speak about and are what we want

to change/measure? Pointing to the dangers of aggregation, what are the changes in attitudes in relationships and power dynamics behind the numbers?

- **Tensions to explore:** between the easy to measure facts and the difficult to measure intangibles (power, attitudes, thoughts, behaviours, quality of life); expanding the boundaries of what is currently used to quantify social change, reconciling (fixed) predictable and standardised indicators with (flexible) unpredictable participatory approaches, and how to build a reflective, dynamic, subjective evaluation framework which is also applicable to a wider discourse.
- **Applications and adaptations:** how to adapt tools to different contexts, how to develop a framework which integrates personal stories and observations of groups, how to strengthen our approaches, and how to streamline relevant tools to use in the moment. And in the analysis, how can we bring out the relationship between quadrants in the Gender at Work framework?
- **Communication challenges:** how can we communicate more clearly, simply and effectively when stories of change are often complex, messy and multi-layered? How to share the complexities of changing gender norms more effectively in multi-cultural organisations? Who defines what we communicate and advocate?
- **Getting beyond the organisational level:** how can changes in the capacity of social movements be measured? Can some complex social changes only be measured at the donor level via the portfolio approach and what are the implications of this? What are we learning about social change processes towards gender equality that can help donors (and others?) better understand, improve and communicate their role?

Participants worked with cards and in plenary to agree on the expected **outcomes** of the meeting. These embraced the opportunity to arrive at a shared understanding of different useful – for both organisations and donors -- methodologies for measuring and analysing gender equality and of the organisational challenges in implementing those methodologies. Notwithstanding the diversity of experience and perspective of participants in the room, there was nonetheless, a call from some participants to build a ‘shared vision’ of how we can learn and move forward. This shared vision should close the gap between theory and practice so that the ‘practice of theory’ becomes something tangible and better understood.

## 2.2 Presenting the Playing Field: Measuring Gender Equality in Organisational Learning



Graphic landscape of the meeting background paper by Holland and Sheppard. See [Annex two](#) for full size graphic.

Jeremy Holland (Gender at Work) presented the prepared background paper, underscoring how measuring, understanding and communicating social change in gender relations plays a critical yet elusive part in tackling gender equality (see [Annex three](#) for slides).

A better understanding of the processes at play is essential for both internal organizational change and for development interventions through reflective practice. The issue remains that traditional linear tools for M&E are ineffective on their own in supporting this type of organizational learning as tracking changes in gender relations and/or organizational cultures are complex, long-term, and context specific. In addition, too often there is a mismatch between donors managing for results, the evaluation community's concern for complexity, and activists and practitioners who are busy with day to day activities.

Jeremy emphasized that how change is conceptualized affects how it is measured, therefore, it is important to start with a strong theoretical basis and then use appropriate methods to test the theory. He presented the Gender at Work framework which maps what change would look like along four dimensions. It captures structure and agency along two spectrums and four quadrants as a means for interpreting, reflecting and acting on what is measured (see [Annex four](#) for analytical frameworks).



For organizations to use evidence about what works best for gender equality, they need to be able to reflect on power relations (both the incentives and vested interests) that shape and maintain the status quo. An overview of the variety of methods for measuring social change was presented, ranging from global, quantitative standardized measures such as the Gender-related development index (GDI), Gender Empowerment Measure (GEM), Gender Inequality Index (GII) and World Governance Indicators (WGI) to those that are more contextual, qualitative and open ended, including Participatory Ethnography Evaluation and Research (PEER), Most Significant Change (MSC) and the story telling approach that has been used powerfully by Gender at Work in its support of organizational learning and change. Some of the methods outlined in the paper are participatory instruments which quantify relational change through scorecards – tools which are at their best when they integrate description of change with qualitative explanation of change.

In the plenary following Jeremy's presentation, the group shared current challenges and experiences with dominant assessment frameworks and approaches, and drew on expertise within the room on other research in the subject including AWID's work on Capturing Change. Alex Pittman (AWID) highlighted three challenges from Capturing Change – the inadequacy of results based framework in measuring performance, the need to shift from thinking of attribution to contribution, and the importance of tracking 'failures'.

AWID has invested considerable energy into capturing, sharing and supporting women's rights organisations to improve their learning systems and undertake more rigorous assessments of their effectiveness and impact. Since 2009, AWID has initiated several action research projects to understand the nature of the challenges faced by women's organisations and their donors and these stories are available as experiences and papers online at the [AWID website](#). Significantly, [AWID's WIKI](#) contains reviews of more than 50 commonly used tools and methodologies with critical analyses of their strengths and limitations, and the potential for people to add their adaptations and experiences with the aim to create an online learning community.

## 2.3 Small Group Work: discussing self-identified challenges

The group identified key challenges for deeper exploration in small group contexts:

- Bridging the gap between theory and practice: how can we navigate learning for organizational change versus donor accountability expectations?
- Communicating the complexity of results in changes in gender equality, e.g. who defines what we communicate and what are the implications, particularly in networks or donor agencies?
- Strategizing for how we access the deep embedded stories and power relations that influence what gets talked about and portrayed in assessments.
- Exploring tools and methods that enhance organizational learning and also provide useful information to donors.



Groups working in the grounds at Dunford House

Key issues discussed include the following:

- **The interdependent tensions between internal learning and external accountability.** These two issues are often seen as separate. The audiences for each are different and the stories you tell to each are different because for external audiences the driving motive often is to show success (and be fearful of failure), whereas for internal learning on difficult issues such as challenging power relations, successes may be few. M&E is seen as the purview of experts and generally is not integrated into organizational processes for learning although evaluations can shed light on how and why change happens. The discussion revealed that these two issues are in fact interdependent and they influence each other in subtle but important ways.

- **Evaluation processes are not neutral** because often they are embedded in unequal power relations (donor-grantee; NGO/client) and evaluations on gender equality aim to examine structural relations of power. Accuracy and openness in evaluation processes therefore requires trust but often the ideologies and motives driving the different players in this process are quite different. Moreover, policy decisions or choices are often not based on evaluations but on other sets of colliding interests (such as ideology, alliance building and/or the need for resources).
- **The need for greater investment in monitoring and evaluation processes in CSOs.** The budget line for M&E is often the last to be considered, first to be cut and is rarely integrated into a learning organisation approach. Investment in monitoring and learning is central to reflective practice towards gender equality, including investment in capacity building for partner organisations in using mixed method approaches effectively and confidently.

## **2.4 Closing comments at the end of day one:**

A productive first day covered a lot of ground exploring and discussing challenges, building towards an open agenda for day two – one which could hone in on the particular issues or open up organisational challenges and experiments to collective problem solving.

On check-out before dinner, each participant contributed their thoughts on the day, and it was during this exercise that a tangible shift in the group could be felt. This shift reflected an acknowledgement of the interplay between technical concerns around measuring gender equality, and an appreciation of the politics of M&E: of being the evaluated, not evaluators, and of being part of the power play. The group was keen to unravel and acknowledge the tensions and inequalities that exist within the M&E process, often undermining its effectiveness.

This shift was seen through the space which participants created through their open sharing and discussion, as participants began the process of owning and contributing to group and their own conceptualisation of understanding and measuring relational. This built into a willingness and keenness to deepen threads of conversation into plenary and small group work in day two.

"I wanted to say what an amazing time I had doing tai chi; I felt that it most definitely made a difference in our feeling of group connectedness--a shift that I felt was most tangible after the exercises for the second day. Measuring Gender Equality Initiative

### 3. Day Two

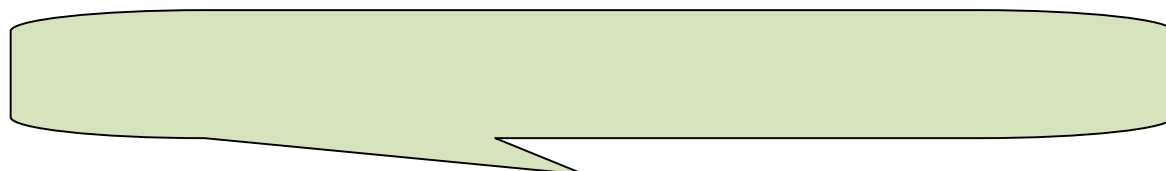
#### Addressing key interests through an open agenda

After the group energised and focused their energies through collective practice of Tai Chi on the lawn, the open agenda structure enabled groups to lead the agenda for day two, honing in on particular issues and challenges that organisations are currently grappling with or experimenting with for group discussion and problem solving. This approach enabled participants to define their own interests and then to convene short clinics to address their key interests.

#### Group exercise: Tai Chi



Tai Chi on the lawn led by Michel



Participant feedback, private correspondence

### 3.1 Making Sense of SenseMaker®: plenary discussion with Jeremy

Jeremy presented an overview of the narrative analysis software programme, SenseMaker®, whose application Jeremy is supporting in tracking societal shifts in attitudes towards – and amongst – teenage girls in Ethiopia and Rwanda. What makes SenseMaker® interesting is that it aggregates personal stories, “signified” by the storyteller, at a micro level, to enable “meta analysis” of patterns and trends, while also supporting local level, discussion, reflection and action.

A rich discussion ensued on the potential uses and strengths of the methodology, as well as its possible limitations. The group identified the positives as the signification of the stories by the storytellers themselves, the shift in understanding cause and effect from simple attribution to complex contribution, the anonymity in aggregation and potential for concurrent collection and analysis which enables real-time feedback for participants and groups on perceptions and trends. However, the methodology is laden with power pitfalls and the way one asks the question matters, highlighting the need for careful pre-testing.

Additional reading drafted by Irene Guijt on how to use SenseMaker® for Monitoring and Evaluation is attached in [Annex Five](#).

### 3.2 Self-selected clinics



Participants working in clinics on day two

The open agenda enabled the following organisations and individuals to convene clinics on the following topics:

- Invisible Power in Organizations (Jeanette Kloosterman, Oxfam Novib)
- Measuring the Impact of Gender Equality (Inga Sniukaite, UN Women)
- Most Significant Change Technique (Esther Benning, Oxfam Novib)
- Measuring the Impact of Core Support on Movement Building (Alex Pittman, AWID)
- After Dunford House, how to continue the conversation? (Rex Fyles, Gender at Work)

Those that discussed **Measuring the Impact of Gender Equality**, questioned the desire for impact assessments as there were so many factors contributing to changes, and instead discussed bringing more focus to the change process. The group identified the need for clarity about the levels of and variety of various interests – personal, collective, professional and political – in measuring impact, and the group proposed the need for meso level indicators that fall between global and local, contextualised indicators.



Those discussing the **Most Significant Change Technique** valued the technique for its ability to open discussion on theories of change to discussions at a deeper level. They looked at suggestions for how to enhance the tool, by minimising the power imbalance and its extractive nature through enhancing the feedback loop to encourage more ownership and debate, or by using peer-to-peer storytelling. And, rather than lose stories and information through the selection process, there was a discussion around codifying, categorising or ranking changes to sharpen analysis. Another level of adaption was to use the method of time-bound specific opinion driven stories to look at significant changes in partnerships – an often neglected area of analysis. This brings with it the challenge of motivating people to write, illustrating that a big part of M&E is not about M&E at all, but individual and collective preferences. By the end of the clinic, the group decided that the numerous recommendations for how to adapt the MSC technique should be collectively shared via the [AWID M&E wiki](#).

The group that discussed '**After Dunford House, how to continue the conversation?**' explored two sets of questions:

1. How do we measure changes in the different quadrants of the Gender at Work framework? Can we test methods, support peer learning? Can we develop meso-level indicators that capture changes between the macro level GDI type indices and the project level indicators? Can we build a peer network to test out practices and meet once a year to share learning?
2. How do we build cultures of learning within organizations around gender equality that makes use of the knowledge and information our measurements generate?

Some participants, such as ISST in Delhi, felt that there is a growing demand for evaluations and they are interested in building evaluative capacity which is not just about measuring impact but also about understanding the trajectories of change in specific contexts and about building a culture of evaluative thinking within implementing organizations. UN Women has supported building evaluation capacity in developing countries, notably through the African Evaluation Association and the African Gender Development Evaluation Network.

The group agreed that we should build a peer-learning mechanism and process for sharing experiences with measuring changes in gender relations in organizational development and learning. Gender at Work offered to explore with IDS how this agenda fits within their research themes, priorities and new funding prospects. Participants were keen on developing/testing particular methodologies for measuring changes in social gender relations and organizational change and on connecting to other on-going M&E initiatives such as those of AWID and CREA/Cordaid. However, the group concluded that given the diversity of perspectives and the lack of funding, there are prospects for sharing individual experiences

through networking in virtual spaces like wikis and webinars and for one-on-one collaborations but no collective momentum to work together as a new formal group.

Gender at Work had initially proposed that the Dunford House gathering would launch an 18 month peer learning process where participants would test methods and reconvene to share their experiences. However, it only got funding for the one meeting. Gender at Work still believes a peer learning process is the best way to deepen experience in this field and explore the intersection between methods, processes, learning and power in the organizational contexts in which we work.

Those discussing **Measuring the Impact of Core Support on Movement Building** defined core support as flexible funding which organisations can spend upon whatever they choose. Participants brainstormed what arguments would be effective to make a case to donors to provide more core support in the context of the MDG3 fund, and on some potential indicators and outcomes.

Their conversations highlighted the divide between donor agendas vs. grantee agendas as one where donors can prioritise vs. core support funding to support organisational capacity and/or movement building. A need was identified to shift donor funding and norms to support organizational capacity or movement building. It was agreed that the best way to argue with traditional donors is on cost effectiveness grounds, highlighting the inefficiency of restricted or project based funding, where every dollar must be tracked, going straight to the issue of control and power.

A donor-to-donor advocacy strategy was suggested, where they make the case for core support to each other; as well as the need to develop a culture of negotiation over funding.

With the aim to persuade different actors, the group looked at various outcomes and evidence to measure. These included: strong and sustainable organizations; diverse voices; increased strategic collaboration; building alliances across movements; engaging and influencing governments and policies; and building alternatives.

As an outcome to this discussion, AWID have added questions on core support to their 2011 global survey to better support its tracking. Based on the evidence from this survey, AWID will produce a short brief on the importance of core support and share it with other grantees, alongside any core support indicators generated for the MDG3 aggregate analysis. All these developments will be shared with the wider community through the WIKI.

Due to insufficient interest in the **Invisible Power in Organisations** clinic, Jeanette spent the time thinking through this problem for the Oxfam Novib context.

### 3.3 Closing summary

The meeting confirmed the **importance of peer-learning mechanisms** as an opportunity to create a new network, expand knowledge, deepen understanding and build upon existing agendas and expertise to develop individual and organisational capacities to measure changes in gender relations in organizational development and learning.

As participants represented a diverse set of organizations working at all levels of international cooperation for social change, there was a variety of perspectives and locations within the collective reflective space. **Alongside the generation of new ideas and thoughts, participants could value the differences in their thoughts.**

With concrete examples from participant experiences, the value of the presentations on SenseMaker® and Most Significant Change inspired groups **to consider some of the deeper issues of organizational learning and gender equality from multiple perspectives**, some with the potential to adapt and engage with these methods in the future. These discussions provided avenues to explore the implications of power and exclusion in the design, implementation and communication of learning and measuring gender equality and organisational learning.

The **diversity** in the room revealed itself in manifest ways – different perspectives, attitudes, and capacities to engage and interact with various conceptions and methods of change and its measurement. This diversity was a vital element in enabling participants to broadly map the landscape of perspectives and interest in the field and enables an articulation of common challenges. However, the diversity also made it difficult to build a shared agenda of experimentation and learning and to deepen our understanding of the multilayered power dynamics involved in instituting and managing evaluative processes – a hot-bed of issues with vying perspectives at play within the processes. Thus, while the diversity in some ways supported the Initiatives objectives, in other ways it inhibited cohesion and the building of common agendas.

There is a clear interest from the practitioners, activists and the Global Fund for Women in testing methodologies, and **exploring more deeply the tension between organisational learning and external accountability**. There is also interest in how to build organisational learning itself through M&E processes.

This diversity also illustrated that there are many players within this field, and there was **no collective agreement among participants to define a specific focus for future discussions and meetings**. However, Gender at Work has established this focus for itself,

with ideas on how to move forward that need to be further explored with potential partners and donors.

Undoubtedly, this face-to-face time was invaluable to develop trust and define and deepen understanding of new areas to explore. Yet, **due to funding restraints**, it is not sure how and when the group will reform to deepen this discussion.

Despite that, there was a **collective ethos of reflection and sharing** enabling individuals and organisations to share valuable sources (see [annex six](#)) as well as make a commitment to build upon the bonds, connections and complementarities within each others' work.

**Recognising AWID's extensive body of work in evaluating and sharing critiques of feminist evaluation to enhance current practice, and in synergy with it** – groups and individuals committed to return to this WIKI to share their experiences. A virtual community of practice will enable groups to continue to share their experiences, and continue the peer learning process.

Following the rich discussions on MSC, where **Oxfam Novib shared their experiences** and through group discussion and peer analysis a richer and deeper reflection was shared on the process, its values and pitfalls. Esther Benning and Jeanette Kloosterman (Oxfam) have shared their experiences with AWID's wiki [here](#), and subsequent to the meeting, AWID participant and WIKI author, Alex Pittman, expanded the wiki to create a [discussion section](#) to stimulate more conversations like those during the launch of the Measuring Gender Equality Workshop. The intention is that AWID's wiki will continue to deepen its evaluation of feminist techniques and pursue better practice by enabling evaluators and practitioners to pose and respond to questions, troubleshoot or share successful methodologies and adaptations.

At the close of the meeting, **participants gave overwhelmingly positive feedback**, with some expressing the wish "for one more day" to continue exploring these topics, highlighting the workshop "stimulating discussion and dialogue" and that while the analysis of challenges was complex and at times difficult, it was an "inspiring, pleasurable, and productive" meeting. The open agenda and the participatory claiming of agendas through clinics were highlights as process high points as well as other spaces and moments of connection (e.g., informal spaces, unexpected conversations and exchange of ideas and tai chi with Michel).

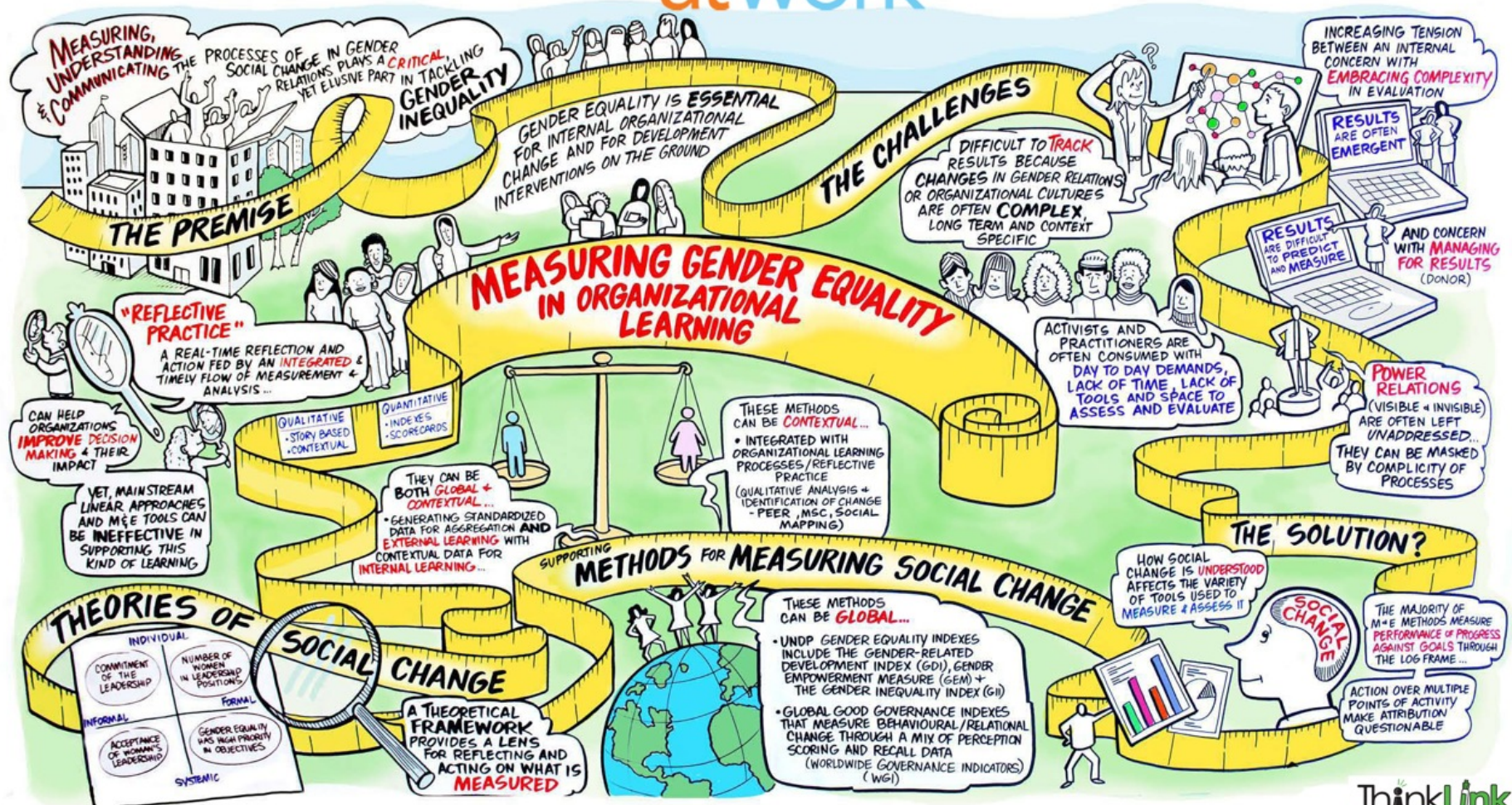
In relation to participants' own continuing work in measuring gender equality and organizational learning; there was **a new awareness** of the power dynamics in M&E, and participants gained ideas for looking at its application within their own organizations. This is part of the broader process of participants connecting M&E to 'the bigger picture' and its links with advocacy, policy and organizational learning and reflective practice.

## Annex one: Participants and their contact details

Name	Organisation	Contact
Alex Pittman	AWID	<a href="mailto:APittman@awid.org">APittman@awid.org</a>
Aruna Rao	Gender at Work	<a href="mailto:arunashreerao@gmail.com">arunashreerao@gmail.com</a>
Caitlin Stanton	Global Fund for Women	<a href="mailto:cstanton@globalfundforwomen.org">cstanton@globalfundforwomen.org</a>
Erica Zwaan	Cordaid	<a href="mailto:erica.zwaan@cordaid.nl">erica.zwaan@cordaid.nl</a>
Esther Benning	Oxfam Novib	<a href="mailto:Esther.Benning@oxfamnovib.nl">Esther.Benning@oxfamnovib.nl</a>
Geetanjali Mishra	CREA	<a href="mailto:gmisra@creaworld.org">gmisra@creaworld.org</a>
Hannah Sheppard	Gender at Work	<a href="mailto:hmasheppard@gmail.com">hmasheppard@gmail.com</a>
Inga Sniukaite	UN Women	<a href="mailto:inga.sniukaite@unwomen.org">inga.sniukaite@unwomen.org</a>
Isadora Quay	Institute of Development Studies	<a href="mailto:isadoraquay@gmail.com">isadoraquay@gmail.com</a>
Jeannette Kloosterman	Oxfam Novib	<a href="mailto:Jeanette.Kloosterman@oxfamnovib.nl">Jeanette.Kloosterman@oxfamnovib.nl</a>
Jennifer Liang	the ant	<a href="mailto:jenniferliang07@gmail.com">jenniferliang07@gmail.com</a>
Jenny Bell	Justice and Women	<a href="mailto:jaw@futurenet.co.za">jaw@futurenet.co.za</a>
Jeremy Holland	Gender at Work	<a href="mailto:jemholland@googlemail.com">jemholland@googlemail.com</a>
Katie Oswald	Institute of Development Studies	<a href="mailto:K.Oswald@ids.ac.uk">K.Oswald@ids.ac.uk</a>
Madhavi Kuckreja	Gender at Work	<a href="mailto:madhavivanangana@gmail.com">madhavivanangana@gmail.com</a>
Michel Friedman	Gender at Work	<a href="mailto:michel.friedman1007@gmail.com">michel.friedman1007@gmail.com</a>
Michelle Higelin	World Young Women's Christian Association	<a href="mailto:michelle.higelin@worldywca.org">michelle.higelin@worldywca.org</a>
Michelle Odayan	Indiba-Africa Group	<a href="mailto:michelle@indiba-africa.org.za">michelle@indiba-africa.org.za</a>
Prof. Maureen C. Pagaduan	Women's Legal Bureau	<a href="mailto:womenslegalbureau@yahoo.com">womenslegalbureau@yahoo.com</a>
Rachel Moussie	Action Aid	<a href="mailto:Rachel.Moussie@actionaid.org">Rachel.Moussie@actionaid.org</a>
Ratna Sudarshan	Indian Social Studies Trust	<a href="mailto:ratna@isst-india.org">ratna@isst-india.org</a>
Ray Gordezky	Gender at Work	<a href="mailto:ray@thresholdassociates.ca">ray@thresholdassociates.ca</a>
Rex Fyles	Gender at Work	<a href="mailto:rex.fyles@uottawa.ca">rex.fyles@uottawa.ca</a>
Rosalind Eyben	Institute of Development Studies	<a href="mailto:R.Eyben@ids.ac.uk">R.Eyben@ids.ac.uk</a>
Siv Lillestol	NORAD	<a href="mailto:siv.lillestol@norad.no">siv.lillestol@norad.no</a>
Sohela Nazneen	Department of International Relations, University of Dhaka	<a href="mailto:sohela.nazneen@gmail.com">sohela.nazneen@gmail.com</a>

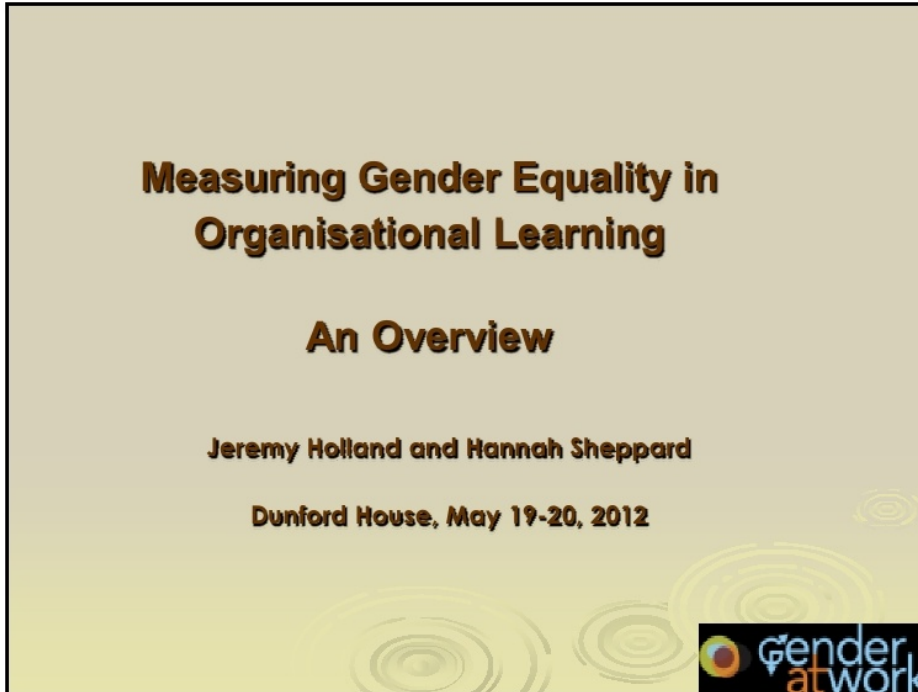


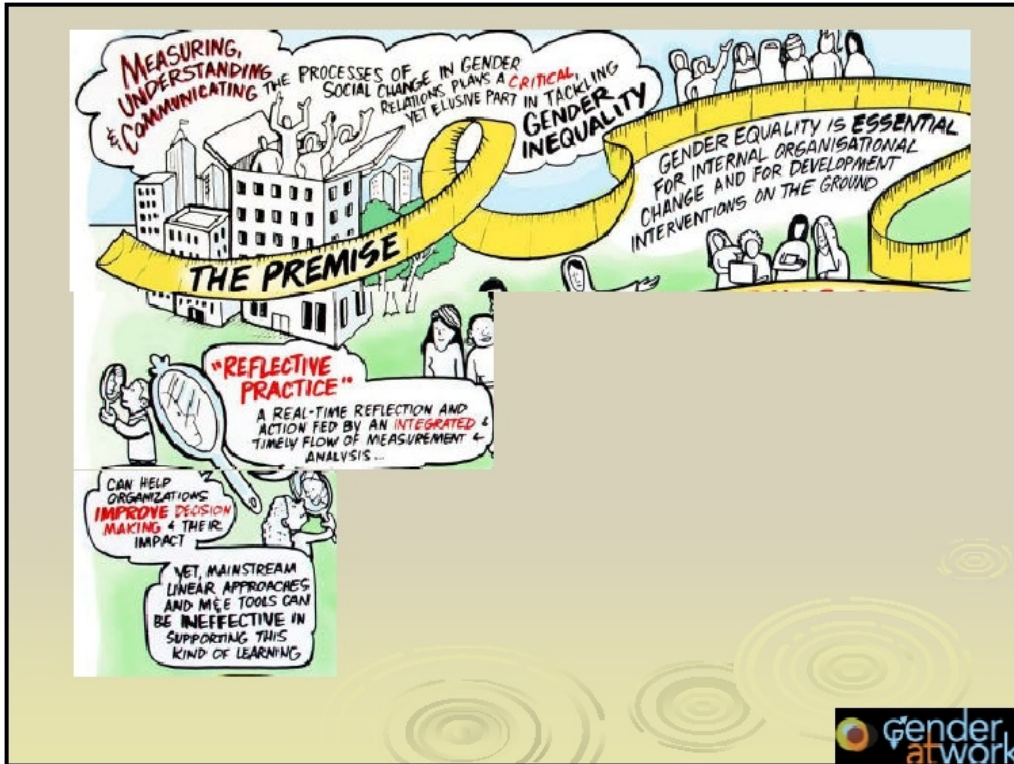
# Annex two: Measuring Gender Equality Graphic





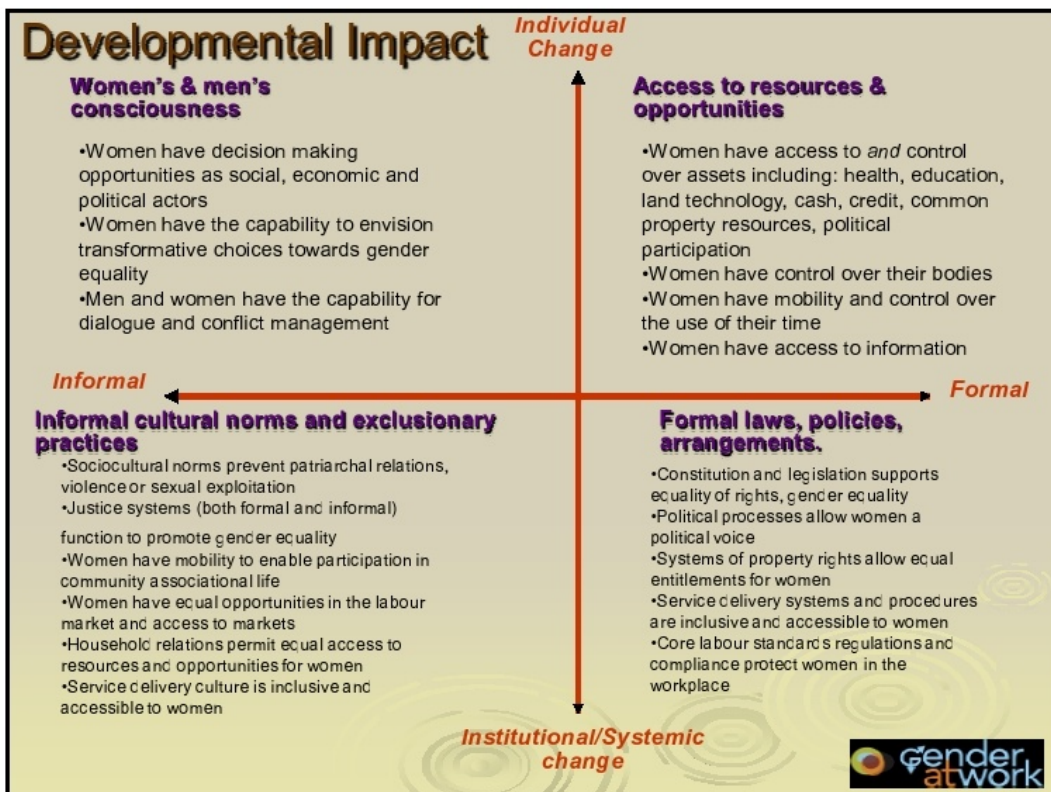
# Annex three: Presentation of background paper



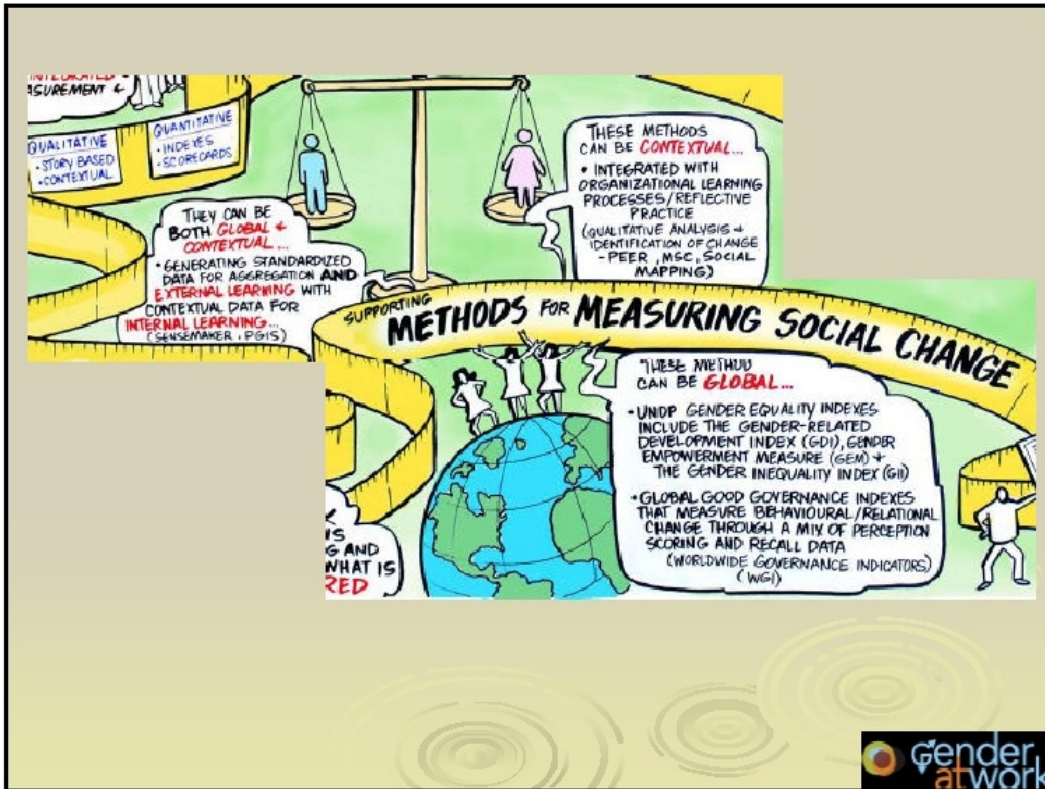










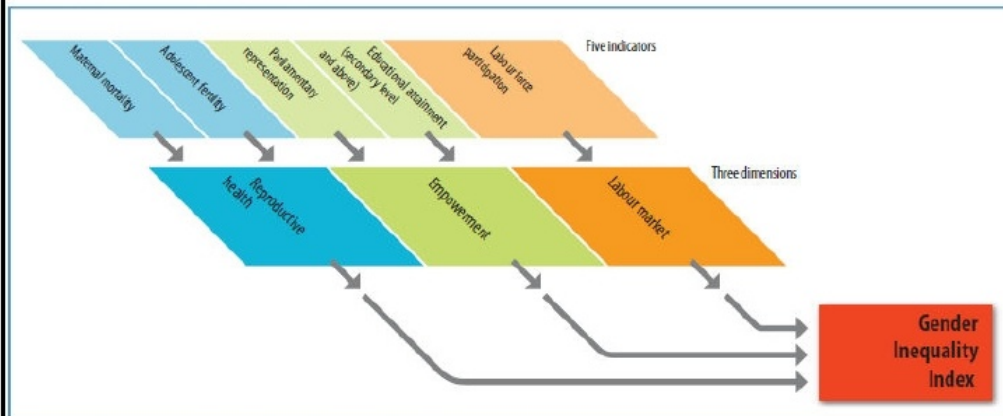


## Global Gender Indexes

FIGURE 5.3

### Components of the Gender Inequality Index

GII—three dimensions and five indicators



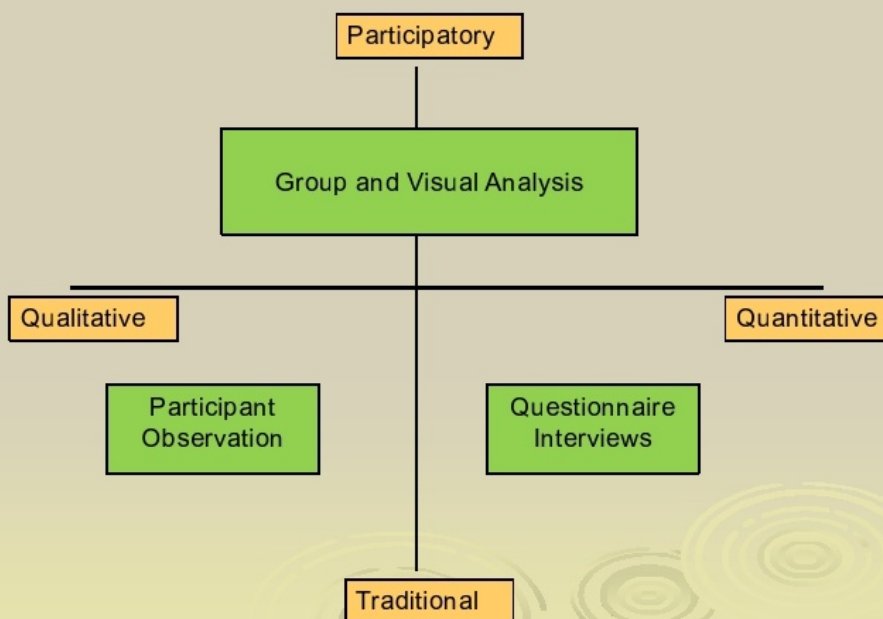
Note: The size of the boxes reflects the relative weights of the indicators and dimensions.

Source: HDRO.

## Measuring changing relationships

- **Observation data:** What is happening?
- **Recall data:** What do you remember happening?
- **Perception data:** What do you think about what is happening?

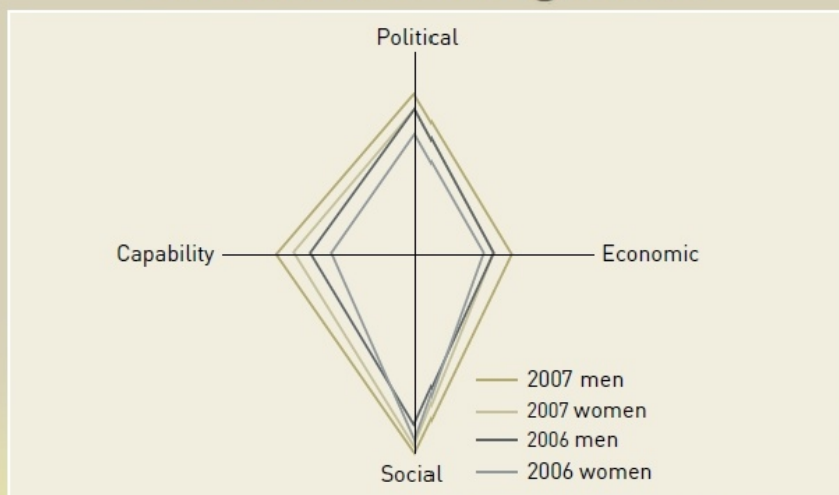
## Dimensions of methodology



## Scoring service delivery and outcomes in Rwanda

- I am able to access information about my entitlements to agricultural extension services
- I am able to access agricultural service X (e.g training, Agric inputs)
- I am able control the use of the agricultural assets in the household
- The services provided are of good quality
- The services provided are relevant to my needs
- I am satisfied with the attitude and behaviour of agricultural extension officers
- I am able to participate fully and effectively in local decision making processes around agricultural services
- I have an equal opportunity to access agricultural services when compared with male farmers
- Household relations have improved as a result of agricultural extension services

## Quantifying qualitative changes: Measuring Empowerment in a Social Movement in Bangladesh



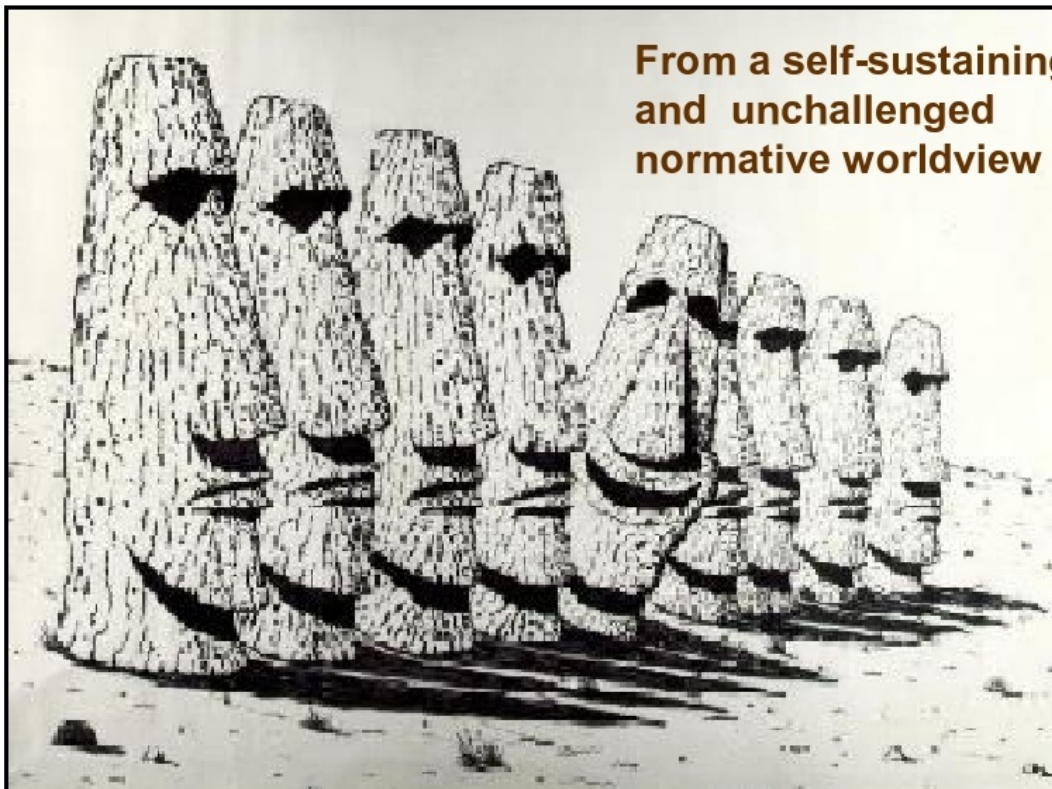
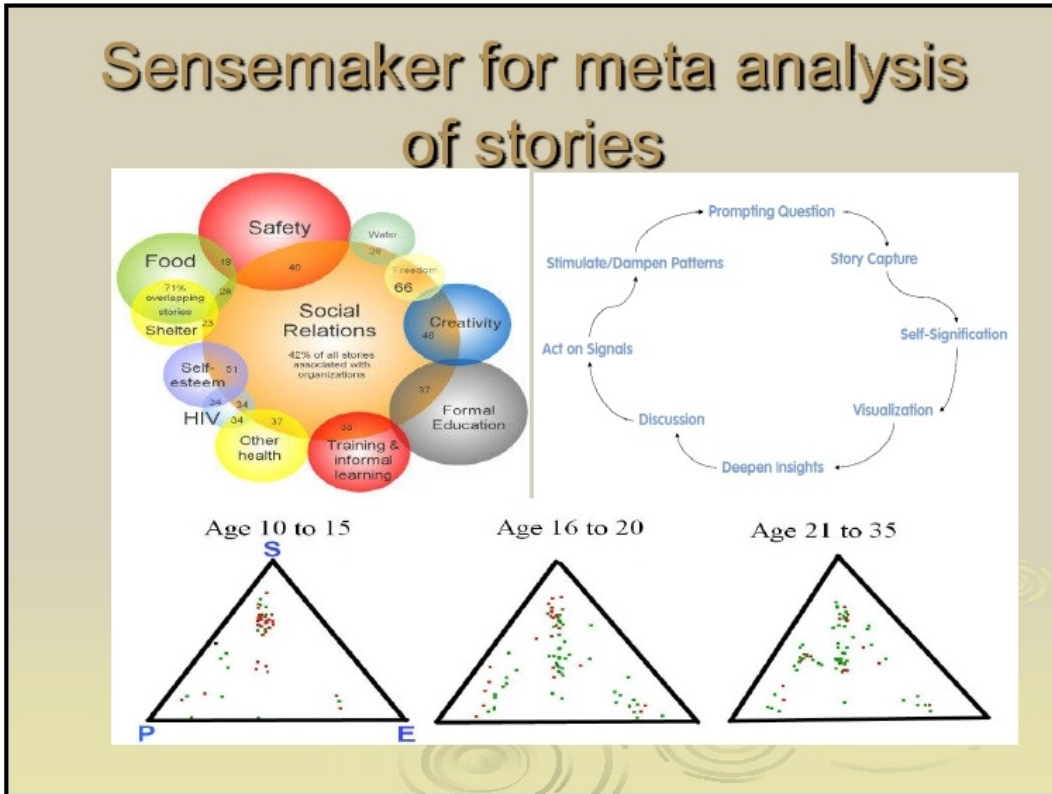
## Samata: Empowerment in Three Acts



## Storytelling

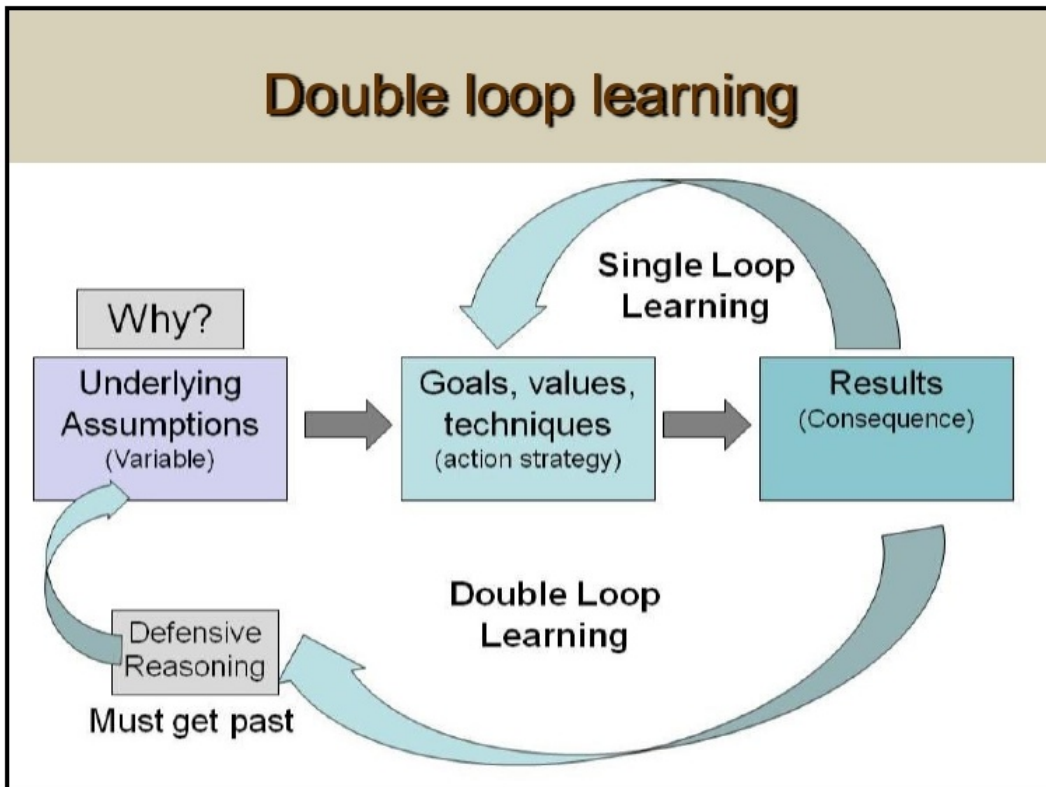






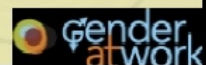
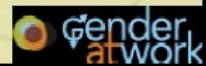
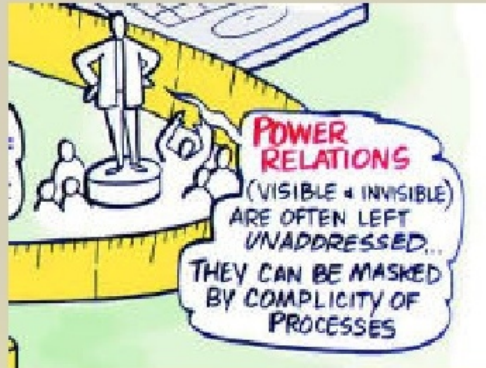


... to reflective practice for organisational learning and change





## ... and reflecting on power: Interests and incentives



# Annex Four: The Gender at Work Framework

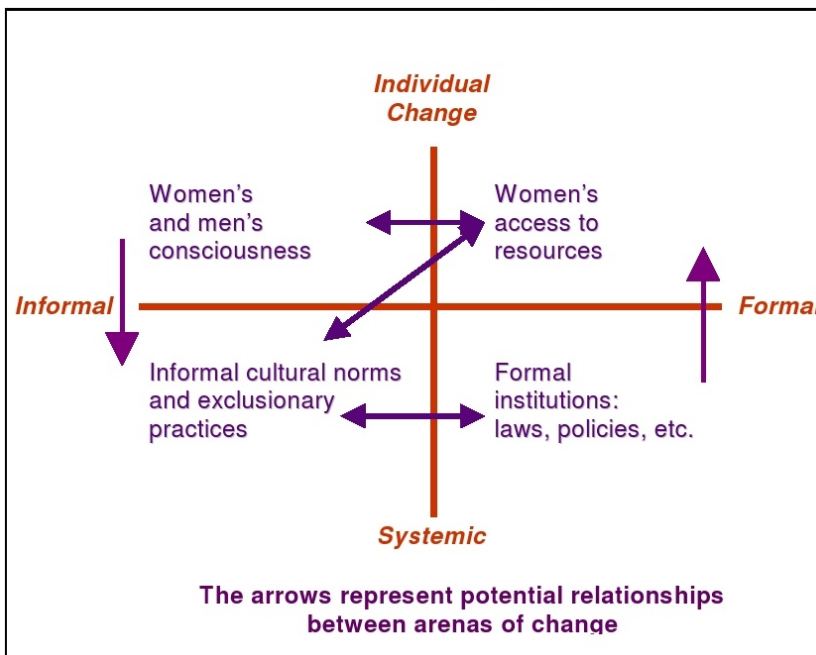
The following conceptual frameworks illustrate Gender at Work's approach to gender equality and institutional change. This focus of the frameworks is on transforming institutions in a process of social change.

The frameworks identify four interrelated clusters of change that can be used to analyse and assess changes towards gender equality, both in terms of organisational change and developmental outcomes.

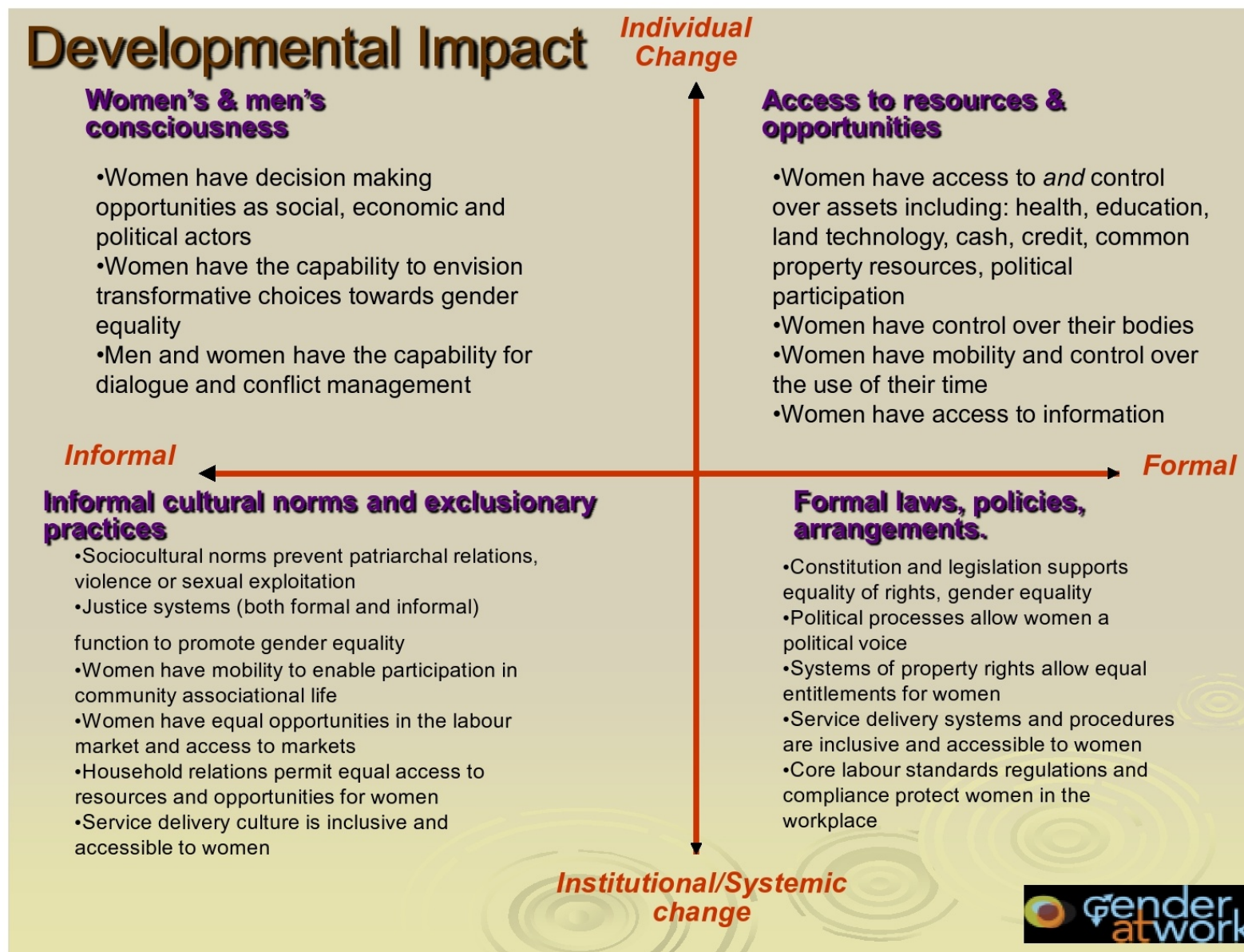
In each diagram, the top two clusters are individual (changes in measurable individual conditions – resources, voice, freedom from violence, access to health) and individual consciousness (knowledge, skills, political consciousness and commitment to change toward equality).

The bottom two clusters are systemic. The cluster on the right is of formal institutional rules as laid down in constitutions, laws and policies. The cluster on the left is the informal norms and cultural practices that maintain inequality in everyday practices.

Change in one of these quadrants is related to change in the others and the arrows show possible directions of relationship:



Gender at Work's approach to gender equality and institutional change is available in full here: [http://www.genderatwork.org/sites/genderatwork.org/files/gw\\_approach.pdf](http://www.genderatwork.org/sites/genderatwork.org/files/gw_approach.pdf)









## Annex Five: Suggested reading on SenseMaker®

---

SenseMaker® is a story collecting monitoring tool that allows organisations to create a set of thematic questions – or a ‘signification framework’ -- to facilitate storytellers to make sense of their own stories. In this way, the interpretation of any story of change, and what is most significant about that story, is left to the story teller herself or himself.

Using the signification framework SenseMaker® software generates quantitative data on themes that can be quickly aggregated allowing ‘meta’ analysis of hundreds or even thousands of stories across populations and over time. SenseMaker® was not developed with international development in mind, but Irene Guijt (Learning by Design) saw the potential for application to monitoring and learning around development interventions and has been working recently with the SenseMaker® team and with development agencies to adapt and implement the methodology. In this paper Irene explains and illustrates the use of SenseMaker® for M&E of development interventions.

## Hearing Need and Seeing Change through Story Cycles: Using SenseMaker® for Monitoring and Evaluation

Irene Guijt, based on work with John Hecklinger, Marc Maxson and Dave Snowden

Draft v2  
January 2011

### What is SenseMaker® ?

**“Creating security.** Many people leaving in slums have lived in fear due to lack of peace and lack of criminal cases many have been raped but to day we can have something to enjoy and SenseMakerile at although there is still little to be said to be add on it. There is a certain group in our community called “*vijana amani pamoja*” (VAP) who have really provided security among the community they have held different workshop educating people to leave in peace and respect each others property they have also advise the youths to involve in various money generating activities which will keep them busy and get money to cater for their instead of stealing other people property.” (Original story shared by 16-20 male in Nairobi for SenseMaker pilot project)

Imagine collecting thousands of stories like the one above from citizens, community organizers, and NGO staff about what really matters to them. Or perhaps your context is one that includes district water engineers, water policy makers, municipal governments, and water users. Now imagine looking through a prism at these stories to find patterns and compare and contrast patterns between organizations, themes, geographic areas, stakeholders, age groups, and more. And imagine getting a continuous flow of stories that allows you to see needs as they emerge and change as they manifest. SenseMaker is a way of thinking and going about collecting, analysing, debating, and sharing large numbers of stories on a continuous basis. But what does this add to existing evaluative practice? How does it challenge common understandings of monitoring evaluation? And where do we still need to know more?

This document describes the logic and logistics of SenseMaker-based evaluation methodology. It draws on an experience in Kenya, implemented in 2010 with [GlobalGiving](#) and funded by [Rockefeller Foundation](#). This experience was the first application in the development sector. And it suggests that much is possible (see Box 1). [The lack of quick feedback continues to hinder development work](#). Understanding change as it emerges and making real-time adjustments is key to meeting people’s needs efficiently. But we also need to seek surprise to challenge ourselves.

#### **Box X. The Essence of SenseMaker**

- focus on *shifting patterns of impact* as perceived by different perspectives, including beneficiaries
- generate databases, actually *people’s life libraries*, that allow – if facilitated and linked to decision makers - *creating evidence-based policy*
- generate *rolling baselines* to continually update evidence base
- enable *cross-silo thinking* (including across organisations, professions, themes) and move away from narrow understandings of attribution of efforts
- seek *surprise* explicitly away from only viewing people’s lives through our own concepts
- more grounded and more diverse *feedback* to donors and therefore more local autonomy
- generating *actionable insights*, based on very concrete needs, via peer to peer sharing

### **1.1 Recognising complexity means accepting new ways of knowing**

SenseMaker as a methodology is based on the recognition that change is often 'complex'. But what does this term mean? Is it only a concept or nice metaphor or also useful for practice?

The interest in complexity has emerged from growing recognition of the disconnect between unpredictability of change processes and linear ways of working in much of development. Amidst growing pressure to show impact, this tension has becoming glaringly obvious. On the one hand, those in development generally acknowledge that improving people's lives is not a linear change process. It requires multiple efforts by many at different levels and has surprising twists and bends on route. On the other hand, those funding this work are asking for unambiguous statements of attributable impact, in the search to know 'what works' and what does not. Many aid organisations use the achievement of pre-defined outcomes to assess performance, rank and fund.

Complexity science offers several terms and ideas that may shed light on these dilemmas and offer ideas on how to rethink our ways of knowing. For example, it suggests that we are dealing with interconnected and interdependent elements and dimensions. In these systems, feedback processes promote and inhibit change so we need to know what we can leverage. *It also says that* change is non-linear and sensitive to initial conditions (so very time and context specific). People, organisations converge around certain parts of the change process ('attractors'). Long term periods of stability can be interrupted by sudden upheaval, requiring a new dynamic equilibrium to be sought. But there are patterns in this.

In practice, it means we need to be able to deal with: real-time emergence – that which we did not predict and cannot control but is happening now that requires a response now. And we need diversity of insights, because many perspectives interact on the situation being changed and all perspectives matter to know what to do. Change efforts resemble an iceberg: what is hoped for is explicit and clear for all to see, but much of what happens or is at play is hidden under the surface. Diversity, emergence, real-time and actionability are core ideas that become operational with SenseMaker.

### **1.2 Core design features that challenge evaluative practice**

SenseMaker is based on several design features, some of which are quite challenging for mainstream evaluation practice. While often associated with [specific software](#), it has a [specific methodological logic and process](#).

1. **Stories** (*micro-narratives or fragments*) about *people's experiences are the starting point – and generate quantitative data*. People's stories matter. People learn by sharing a multitude of anecdotes in many conversations, mingling facts and opinions. They act on stories. However, 'fragments' can also be newspaper clippings, photographs, audio or video recordings, or sections of reports. Each story is converted into a set of mathematical points that enable numerical patterning. This integration of hard and soft data addresses a key challenge in mainstream evaluation practice.

#### *What is SenseMaker?*

- a methodology, based on a set of principles and design features that uses narratives to understand trends
- pattern detection software that facilitates analysis of large quantities of narratives

#### *What part is patented?*

- The generic signification of meaning into a geometric shape, so both the methodology and software.

Mixed methods are common in evaluation processes, with people collecting stories and experiences, as well as statistics. But each is a separate data set, and often only one or the other forms the evidence base, with separate streams of inquiry with difficulty of integrated analysis and problems related to both. SenseMaker integrates both.

2. **Many narratives** (*‘mass capture’*) show diversity of perspectives. Large numbers of stories – for GlobalGiving we had almost 3000 – help reduce biases and open up many new aspects of the issues being analysed. All these fragments add up to a set of multi-faceted impressions about the situation being looked at. Purpose and random sampling can be used to structure data collection. Narrative gathering can and has been done by children as well as experts.
3. **The storytellers self-signify**. In standard evaluation practice, experts interpret the stories being shared, bringing in their own cognitive and cultural biases. The outsiders tag, code, make sense of personal experiences. SENSEMAKER asks the storyteller to code his or her own story, thus adding additional information to the story being shared. External signification is also possible, for example, of newspaper clippings or report snippets. This approach has been used on its own but can also be used in combination with self-signification. External signification inevitably introduces some external bias, and provides less additional information as the signifier is not familiar with the context in which the narrative or story was generated.
4. **A shared signification framework generates consistent data**. Each fragment is signified through the same question framework. These are not specific indicators but more meta-level variables and issues. The question framework is created based on overarching objectives but also values and beliefs that are driving the program or project being looked at.
5. **Analysis is aided through pattern visualization**, which allows actionable insights to be constructed. The software takes the many stories and allows for very quick and flexible probing to detect correlations that might matter. Patterns are generated by asking questions about the full set of narratives or comparing sub-sets. It is only after interesting patterns are detected, that people then zoom in on the stories related to that cluster and discuss, make sense and plan actions.

#### **Clarifying when SenseMaker might be an option**

As with any method, SenseMaker needs to be appropriate. What can and can it not do? And when is it useful? First it is important to distinguish the approach from the software. The innovative type of questioning on which SenseMaker , with its triads and dyads (see below) can certainly brings novel insights. However, it really only becomes useful when dealing with many stories. The software becomes optimal from around 100 or 200 stories upward. Less than that and the human eye and brain can detect many patterns itself.

SenseMaker is well suited for those situations in which there is great diversity of possible outcomes and/or great diversity of perspectives that need to be compared and contrasted. For example, a trust fund focusing on domestic violence reduction that has 1000 grantees, each with its own outcomes, could identify the patterns are present in that diversity. It can also be used when many experiences from beneficiaries are needed to inform what has and has not changed. This is the case of the GlobalGiving experience (see Box 2).



It is also well suited to describe systemic level changes over time. In some situations where it is as yet unclear precisely what kind of interim results will occur, but where it is clear what the long term intentions are. The IRC is working on improving water service delivery in Uganda and Ghana and each year will see a new set of activities being initiated. Hence, unanticipated changes at the local, district, national and international level will need tracking and patterning.

**Box 3. GlobalGiving asks itself ....**

- Are we helping organizations learn to be more effective?
- Are GlobalGiving organizations different from others?
- What is the risk profile of the 1000+ organizations on the GlobalGiving platform?
- Are we delivering on our value prop to donors and organizations alike?
- What can GG do differently to respond to/foster community efforts that are not getting the attention they deserve?

For a narrowly defined accountability evaluation, SenseMaker is not the best approach. It is not well suited for establishing unequivocal attributable impacts, nor for precise hypothesis testing.

However, much depends on the design of the question framework. One can, in theory, make it more connected to specific outcomes or to test certain hypotheses.

### **Working with SenseMaker**

Once clarity exists about why SenseMaker might be an option, four tasks follow:

1. Identify what you want to ask
2. Collect stories and other kinds of narratives
3. Analyse the stories for significant patterns
4. Make sense of the patterns and act.

For each of these tasks, it is essential to ask who should be involved.

#### **2.1 Decide what to ask – steps and skills to design a signification framework**

SenseMaker differs considerably from a classic evaluation approach in three ways. First, a classic evaluation requires describing what you want to check is happening, the performance questions, and the indicators to be measured. Developing a signification framework means letting go of this logic – but not entirely. Indicators are absent and explicitly ambiguous questions are designed. Second, it is a frugal approach. The GlobalGiving question framework contains only 15 questions. A minimalist mindset is essential, not a survey mindset. And third, changes are not questioned directly. Rather comparisons of story sets over time allow for seeing any shifts that occur. In this sense it could be used for results-oriented monitoring.

Start by developing a **prompting question**. This single question is what triggers people to tell a story they find meaningful. For example, 'What specific moment or event made you feel discouraged or hopeful about rural water service delivery?'. This prompting question will be used for all people from whom you will collect stories. So think about who these people are. Make sure at this stage that it is clear that all information will be and remain anonymous.

**Identify the main domains of information or themes.** Questions that can be useful to discuss and then map out in thematic clusters are:

- What are the main information needs?
- Who wants to know what and why? Think about the different users of the information and what they need or would be curious about.
- What are the main objectives of the organisation/project/program/portfolio? Is it about seeing shifting gender relationships? Or more coordination among water service providers? Or empowerment of farmers to engage with markets?
- What the values, beliefs, capacity shifts underpin the changes being sought?

Here are a **few practical tips** to keep in mind. Do not be tempted to formulate *evaluation questions* at this stage. Keep it general, for example, ‘we want to know about service provision improvement’ or ‘we want to know if people’s needs are being met and if there are many unaddressed problems’ or ‘we want to know about participation in decision-making’. Also, *less is more*. What are the really non-negotiable issues here, and weed out the interesting but unnecessary elements. Make this process *interactive*. Have as many perspectives as you can in the room – program managers, water engineers, people working at different levels, etc.

Translate the **information domains into a set of questions**. SenseMaker uses three types of questions. The *triads* and *dyads* (or polarities) are what people will use to give additional meaning to their stories. These are called the ‘signifiers’. They are very different from current question types used in evaluations and they are very unique to SENSEMaker. The other question type is the traditional *multi-choice questions*. These are the so-called ‘modifiers’ and help create comparisons. An example of the full signification framework can be found on the [GlobalGiving story website](#).

- *Triads* are explained in Box 3. Stories are usually a mix of elements. Requiring people to tick one box or another, would lead to an excessive simplification of the stories. Triads allow people to say, for example, ‘my story is mainly about social relationships, and quite a bit on economic opportunities but not at all on improving my living conditions’. The storyteller then places a mark in the triangle to represent that distribution.

**Box 3. Two kinds of triads**

1. Three variables are the points of the triangle. These are evenly balanced labels: all positive, negative or neutral, with the centre being a balance with all three being equally present. The storyteller marks in the triangle to indicate how the stories relates to each three labels. For example: this community effort is about social relations, economic opportunities, physical wellbeing. **INSERT VISUAL EXAMPLE**
2. Take a polarity and add an opposing pole. The extra pole is not part of a continuum between the two poles but relates to another factor. For example, ‘Your story of starting at university is about... enjoying the freedom to decide your own study hours, enjoying the structure given by classes, disliking making friends’. **INSERT VISUAL EXAMPLE.**

- *Dyads or polarities* are useful to construct patterns in relation to a certain quality, issue or result area. It is a sliding scale between two extremes, one of

excess and the other of deficiency. The ideal situation – for example, certain behaviour or a certain impact – lies in the middle. A specific value, such as trust or ownership, would be suited for a dyad. But it could also be degrees of success or degrees of failure.

- *Multi-choice questions* are a traditional question form. They allow for secondary analysis of patterns in the stories. Essential is basic demographic information about the story teller and story is needed. For example, age and gender, story location and if it is a personal experience or a heard experience would be included here. But multi-choice questions would also include thematic options. For example, the different factors that determine quality of water service delivery can be a multi-choice question, or the core human needs. Other useful questions are about story intensity ‘How do you feel about your story?’ or ‘How likely are you to remember these stories?’. Remember to list the appropriate options when constructing the multi-choice question. And to indicate if one or more options are possible.

The process of creating these questions is iterative because they function as a set. As you sketch ideas for dyads and triads, think how you might combine them. What kinds of combinations of multi-choice questions or demographics with which dyads and triads will allow you the kinds of comparisons you want? In this process, overlapping questions can be identified and removed.

Decide if you can work with **one or multiple signification frameworks** for the different sources of stories. If you are working with considerable heterogeneity, then you may need to have more than one. Or you might have two frameworks, with partial overlap and sharing of some questions. For example, you might need a different set of questions for water users and for all professionals working to ensure water delivery.

**Pre-test** the questions, first with peers to iron out the obvious problems and inconsistencies. And then **field test** the questions with the intended story users. Make adjustments and then the Collector website can be built for your application. For a recent application in water, we pretested the questions over 24 hours with 30 people. The fieldtest was organised to collect 75 stories in two weeks from different groups of people internationally, in Ghana and Uganda.

### ***B. Collecting***

As for any evaluation process, decide on the process of collecting information, which in this case, are stories:

- from whom the stories will be collected;
- what technology will be used;
- who will collect;
- with which frequency and how (interviews or anecdote circles; help or no help with signifying).

**List all the groups of people whose stories you will seek.** In some cases, these story tellers may also be the users of the analysis, for example community volunteers to understand how their organisation’s work compares to other similar organisations. This issue is important, as you need to consider which incentives will help to encourage people

to share. Will this be ensuring structured feedback of the analysis, or by hearing specific story examples? Or some other reason?

**Collection options.** Stories may need to be collected in different ways from people, depending on cultural preferences, technological options, and quite simply, the resources you have. The main issue is to ensure that stories are collected in a suitable and feasible way that maintains high accuracy, and that they are uploaded into the Collector website (see below). Collector is like a library where all the stories over time come to be stored and can be accessed for analysis.

- *Basic pen and paper.* This option requires someone to ask, perhaps document, and hand in the stories that are written on paper. This could involve volunteers or field staff or children. If stories are to be collected on paper, guided by the [question structure](#), then you will also need to ensure typists who enter the stories and their signification into Collector. GlobalGiving collected over 3000 stories in 10 weeks and then needed another month to upload the paper-based stories. Transcription requires some basic training for the typists and spot-checking to ensure accuracy of transfer.
- *Direct web capture.* For those people with internet access, the URL can be sent to the Collector website for direct entering and signifying. See the GlobalGiving website for an example of a Collector website.
- *Smart pens or mobile phones.* Although not used in the first GlobalGiving experience, Cognitive Edge is working on a mobile phone application that allows stories to be shared and signified and then sent to the Collector site. Smart pens are also being tested but not yet available for wide use, and are a relatively costly option.

**Pilot test story collection.** While field testing the stories, also pilot the story collection process. In GlobalGiving, much testing of technologies, collection modes, scribing modes, training modes, communicating with story modes and even incentives was undertaken. This year will see more of that experimentation. Do look at 'The Real Book' (Maxson et al 2010), section 1.6 which discusses the many aspects that needed considering and the many unexpected that we encountered, including language and electricity.

**And then [collect the stories](#)** – lots of them. And keep collecting them. The power of SenseMaker occurs when you can start comparing, over time, the stories and the patterns of change they reveal.

**Get Collector up and running.** The stories, fragments, narratives all end up in Collector. This is a website that Cognitive Edge designs based directly on your signification framework. It is set up for open access so people can enter their stories and signify them. Cognitive Edge needs to receive a copy of the signification framework. They require two weeks to program the draft site, ideally without needing to edit the signification framework. Once the site is up, it needs reviewing for spelling and layout, like proofreading a document. And it needs testing by entering and signifying sample stories. All sample stories are deleted once Collector goes live. Setting up the website requires a contract with Cognitive Edge.

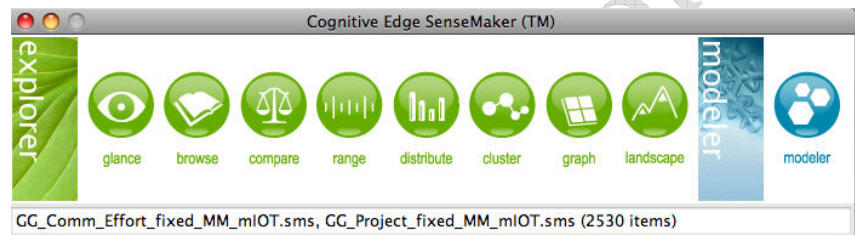


### C. Analysis for Patterns

Analysis consists of two phases. Phase 1 uses the software to show what kinds of patterns are strong and of potential interest. The second phase is when people discuss ‘so what does this mean for us’ (see under D below).

Before you do any analysis (and this is best done before transcribing if you are working with paper), **filter out the junk**. Only pick out the stories that are obviously junk, such as ‘A wheelbarrow is a useful tool...’ or ‘One day I saw two giants fighting in the forest’ (real examples). In the GlobalGiving work, only 9% of the stories were filtered out in this process. Now you have a cleaned data set, ready for analysis.

**Get prepared** to use Explorer, the analytical SenseMaker software. Download the file you receive from Cognitive Edge onto computers. With the full data set of stories, open Explorer (see Figure) and play with the stories to detect patterns of interest. The first and second times are quite difficult. The seeking of visual patterns, understanding the potential of the software require some skills. Some initial coaching is ideal but plenty of playing around will also get you there.



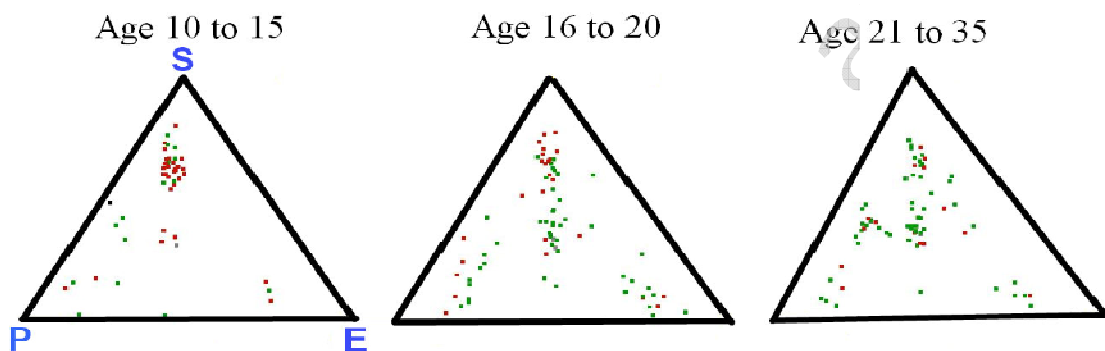
First **get familiar with your data**. Start with the ‘Browse’ function. This allows you to get an overall sense of the data, the amount, who is telling the stories, and how the answers are spread across the options. Then start to create the patterns with the dyads or triads. Use ‘Distribute’ for the dyads and ‘Graph’ to recreate the triad. See how the stories – one dot is one story – are spread between the three variables of each triad or along the dyad (see Figures X and X). **INSERT ONE VISUAL OF A DYAD COMPARISON.**

One way to help to literally **see the patterns** is by printing out all the triad and dyad patterns and [put them to the wall](#) and then look at what the pattern is. Where are extremes, where are small clusters, where do they trend towards and away from? Which patterns are revealed by zooming in on stories with a strongly positive emotion, or an intensely negative one? By iterating and comparing with the many possible combinations of variables, triads, dyads and multi-choice options, a set of significant visual patterns emerges. It is a mix of allowing the data to show you what is significant. But you can also ask specific questions of the data. For example, what is the difference between women and men’s stories in terms of access to economic opportunities? Or show me the stories on water and those on safety for young and old women in this part of Kenya compared to that part of Kenya.

After this comes the **drilling down**, when clusters of dots are marked and the software immediately allows you to look at that specific story set – one story for each of the marked dots. By reading a limited set of stories that represent a specific visual pattern, you can make meaning of what the stories as a set tell you. For example, a small set of 23 stories may all mention abuse in schools or pride about a self-built home. This is the beginning of seeing what matters to people and what needs attention. For GlobalGiving, a rich picture emerged about what people in various communities believed they needed, what services

they were getting, and what they would like to see happen in the future. The software can be used to self-generate reports.

Maxson (2010) gives one example. Men (green dots) and women (red dots) talk about education differently. Girls talk about education at an earlier age, and they couch it in terms of (S)ocial challenges and aspirations, not future (E)conomic opportunities or (P)hysical well-being. By age 16-20, girls and boys talk about education in terms of social empowerment and economic opportunity, respectively. By adulthood, women have stopped talking about education. Note that this is only starting hypothesis, and that a decision maker would then read the stories in each interesting cluster to see whether they confirm this, or actually suggest an alternative pattern.



Additional analysis is also possible without SenseMaker, by exporting it to Excel and searching for certain words.

#### ***D. Feedback loops***

Stories, if not used, are just static. Analysed patterns need to be put to work to be actionable. To do this, you need to create feedback loops in the system: feed the visually interesting patterns and comparisons back to the people making the decisions and implementing the work.

As using software is not easy for many, think through with who, when and how the analysis will take place. Will this be with direct access to the software or by printing out the patterns, taking them to a community or staff meeting and sharing impressions and ideas, reading the related stories, and figuring out what to do differently next? Of course, if the work is happening with the same people, then pattern and story analysis and decision-making occur seamlessly.

Also think through how often to do this analysis – with the software-people interaction – will take place will depend on for what product or process is needed. Is it an annual review or the bimonthly staff meeting? The rhythm of SenseMaker analysis should follow the rhythm of decision-making in the context where the stories are being analysed. For the water work, we are planning our cycles of intense story analysis to precede key review meetings, and our story collection to occur at more frequent meetings that already occur naturally.

People will need some facilitation at first to know what to do with these dot-filled triangles

and charts. Ask different people or groups to look at different comparisons or at specific story clusters. What are good patterns they want to see more of, which ones are worrying and need to be tackled. Agree on meaning - or agree to disagree. Discuss what it means for how to change or maintain current activities. That it can be powerful is clear from TYSA (see Box 4).

Feedback can also be to the general public. For GlobalGiving, as the stories are anonymous it is still exploring for ways to enable everyone to see the stories. All the stories have been geotagged and located on the [Ushahidi](#) platform to a map of Kenya. The GlobalGiving project walls also house the stories, and anyone interested in those projects can follow the emerging conversation thread. Experiments are being planned with phone text messages.

### ***E. Improve the process***

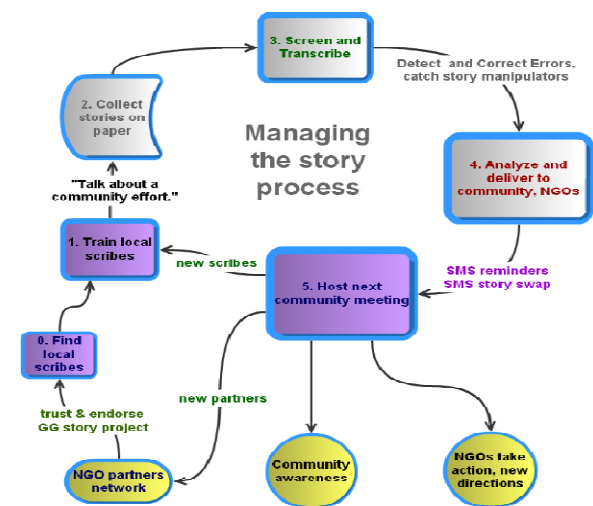
On the basis of a thorough analysis, **refine the process**, from steps A to D. Prior to a second cycle, find out if people and organisations acted on the patterns. If they did, which ones resonated, which ones were not taken up. What was the result of changes in the work - if so, which ones & what was the result - Was how we looked at it useful in the first cycle, in so far as we improving work and acted to reduce problematic patterns?

#### ***Box 4. Stories for Local Change (Maxson et al 2010)***

Trans-Nzoia Youth Sports Association (TYSA) met to discuss the 140 stories collected about their organization posted online. They identified eight specific themes and frequent mentions about three of their four projects (education, sports, and capacity training). However Francis Gichuki says that the discussion was mostly about why a fourth project (child rights and protecting children) was not mentioned much, although stories about these issues did come in. "It is interesting because we do a lot on this issue," Francis Gichuki said in an interview. "It is a gap we are seeing. A gap between our service and the community's awareness."

### **Organising the work**

For these tasks, many people need to be involved. Who will lead the design process and who else will be involve? Who will be involved in the data collection - mobilising, training, collecting, checking? This part required much creful thought and trial and error (Maxson 2010, pg 8-10). Who will do the initial analysis and who will be involved in discussing the significant patterns and the stories behind them? These roles will be distributed between different organisations, volunteers, someone with SenseMaker experience, and Cognitive Edge as it is patented software. Figure X shows how it has been organised for GlobalGiving.



### **Remaining Questions to Understand and Aspects to Improve**

Application value for evaluation. The GlobalGiving experience suggests that SenseMaker has value for different evaluation applications. It can be used to generate rolling baselines. By comparing people's stories of needs with what organisations and agencies are doing in

those areas, you can identify what else might be needed to bridge the need-support gap. It can be used for time bound exercises, such as mid-term reviews or budget support evaluations where many different perspectives along the results chain need to be heard and integrated into a coherent story. It can be used for results-oriented monitoring. And if multiple organisations use the same signification framework, then cross-organisational comparisons become possible. It can be used to create some order in an existing portfolio of projects and their reported outcomes. But each of these applications still need to be tested in more contexts with more organisations over longer time periods in order to assess the full utility.

#### Four issues of rigour and one of accuracy

Rigour has different dimensions. A first requirement is never claiming more than is possible. It is crucial that SenseMaker is understood for what it can do, not what it can't. Second is to be clearer if the sample is representative enough and if the conditions for sharing are not introducing biases in what is shared (see Maxson 2010, Part 2 on misinformation for a good overview). Another important issue is whether people truly understand the nuance of the triad and the dyad. Do they plonk their answer or deliberate and consciously locate? Are they seeing the triad as an either/or question or truly as a triad as intended? The coming year will see more work on this in Kenya. A fourth issue is that of control groups and whether we need to pursue this practice, or purposely create locally meaningful conversations. The work with IRC is intending to work with control groups to see if the story patterns differ in any way. And finally, the accuracy – how 'real' are stories? Box 5 shares initial ideas for looking at the reliability of stories. But more work is needed on this aspect.

#### **Box 5. So how do we tell if stories are not invented? Here are some guidelines. (Maxson et al 2010)**

1. Stories from many different (independent) sources are more trustworthy.
2. Stories from beneficiaries that include special details only project staff should know are untrustworthy.
3. Stories that share a similar narrative structure, and come from around the same time, but from different sources less trustworthy.
4. Stories from a source that has a long track record of submitting other trustworthy stories are also trustworthy.
5. Stories that provide unexpected lessons (perhaps a mixture of positive and negative aspects of an NGO) are more trustworthy.
6. Stories with excessive NGO self-referencing or formulaic praise are inauthentic.
7. Verification – using face to face meetings, SMS feedback – how do other people within the community react to questionable stories?
8. Stories that *djotjog* (below) are reliable, i.e. not two people say the *same* thing; rather, story *tjotjog* is when many people tell slightly different stories that reinforce some common theme among all of them.

#### What does it tell us

It should be clear that attribution is hard. But what will it tell us about contribution to development? For this, the story tellers will need to mention organisations and enough stories need to be told about them to elevate one anecdote to a pattern. The signification framework construction is crucial here. The balance between too much and too little is fine. Too many and you risk losing the interest of people to contribute stories. But too little and you might not have enough variables to drill down and see the nuanced patterns in the stories. How much is enough, how simple is too little? These questions remain open.



### Issues of power and the role of technology

Clearly the technology can lead to a distancing and an expert-dependency. For GlobalGiving this is not desirable. It is experimenting in Kenya with different ways for the participatory review of findings. One example of such an experiment has been written by [Marc Maxson of GlobalGiving](#). This step will require a bit of piloting for each context to find 'best' ways for presenting images, stories, etc. Long-distance coaching and group workshops might be ways to enable the use of Explorer more widely, particularly for projects or partnerships that are located in different places.

But power also lies in defining the question. Why can storytellers not generate their own triads about their stories, and then signify them? In 2011, GlobalGiving will test this approach called 'the story marbles approach' (see Maxson 2010, pg 7). What we need is a means to let the storyteller define the right question while also constraining the possible questions enough that we will derive useful clusters of stories with similar question frames.

### Each hammer needs a craftsman

And finally, as with everything, paraphrasing [Euan Semple](#) and applying it to SenseMaker... [it] "*...will change nothing until we use it for that purpose but the way it enables us to do so is new. What matters is that people understand it and use it. Take it seriously to shape the world.*"

This first application of SenseMaker in the South for development has led to some exciting challenges to mainstream evaluative practice. Some are summarised here but Marc Maxson discusses more in his document. He urges the abandoning random sampling and valuing locally relevant conversations that can be started with stories and analysing the patterns they co-create. Cross-narratives is discussed, a kind of peer-scribing process. This involves those from one volunteer purposively seeking stories from another, in the interest of reducing positive story bias. Maxson also discusses the importance of stopping problematic story flows, verifying stories, balancing incentives to participate per group and level. And last but not least, ensuring networked learning. Much fun lies ahead!

## **References**

GlobalGiving's webpage with much information on its experience with SenseMaker

<http://www.globalgiving.org/story-tools/>

[http://www.cognitive-edge.com/blogs/news/2010/09/final\\_report\\_published\\_by\\_glob.php](http://www.cognitive-edge.com/blogs/news/2010/09/final_report_published_by_glob.php)

<http://www.cognitive-edge.com/casestudies.php?csid=20> (and click on report files)

More theoretical information on SenseMaker concept with many videos:

<http://learningtobeprofessional.pbworks.com/w/page/22714631/From-induction-to-abduction,-a-new-approach-to-research-and-productive-inquiry>

Maxson, M et al. 2010. [The Real Book](#). Online public document. A very rich and detailed source of the entire GlobalGiving experience during its first phase in 2010.

## Annex Six: Useful websites

---

Gender at Work: <http://genderatwork.org/>

AWID M&E wiki: <http://awidme.pbworks.com/w/page/36050854/FrontPage>

AWID's M&E insights for women's organisations:

[http://www.awid.org/Media/Files/MnE\\_Thirteen\\_Insights\\_Women\\_Orgs\\_ENG](http://www.awid.org/Media/Files/MnE_Thirteen_Insights_Women_Orgs_ENG)

Better Evaluation: <http://betterevaluation.org/>

Power Cube: <http://www.powercube.net/>

Outcome Mapping: <http://www.outcomemapping.ca/>

Pelican: <http://dgroups.org/Community.aspx?c=3c4b8b5b-d151-4c38-9e7b-7a8a1a456f20>

Capacitar: <http://capacitar.org/>