Methodological reflections on the use of SenseMaker® for Impact M&E of Adaptation in the Mekong Delta (AMD)

Adinda Van Hemelrijck and Zhen Goh

December 2015

Expectations in a first pilot are often quite high and generally unrealistic.
Sponsors and managers expect magic and tend to forget that
innovation takes time and needs sufficient resourcing to mature.
They also tend to forget that piloting involves learning
in order to bring the concept to its mature stage.
Thus by expecting that innovation pilots will do magic,
one often forgets that the only real failure
is the expectation itself causing ***failure to learn.***

Table of Contents

Abbreviations 2

1 Introduction..3

2 Background 4

3 The potential of using SenseMaker® for IM&E..5

3.1 Added Value of using SenseMaker®...*5*

3.1.1 Added Value in Theory...5

3.1.2 Added Value in Practice 6

3.2 Rigour in using SenseMaker®...7

3.2.1 Rigour in Theory..7

3.2.2 Rigour in Practice 8

3.3 Summary of findings...12

4 Next steps for integrating SenseMaker® in IM&E of AMD 13

References 15

Abbreviations

AMD	Adaptation in the Mekong Delta
IM&E	Impact Monitoring & Evaluation
PAR	Participatory Action Research
PRA	Participatory Rural Appraisal
RTIMP	Roots & Tubers Improvement and Marketing Programme
SDGs	Sustainable Development Goals
SEDP	Socio-Economic Development Planning

1 Introduction

Between October 2014 and September 2015, the International Fund for Agricultural development (IFAD) piloted SenseMaker® in the AMD (Adaptation in the Mekong Delta) Programme of the Vietnamese government in Tra Vinh and Ben Tre provinces. The purpose was to demonstrate the potential use of this method for Impact Monitoring and Evaluation (IM&E) and its complementarity to other more conventional methods. In the case of AMD, this implies tracking changes in the capacity of communities and institutions to adapt to climate change, and identifying knowledge gaps and opportunities for developing such capacity through Participatory Action Research (PAR) and participatory Socio-Economic Development Planning (SEDP) at the local level, and policy dialogues at the provincial and national levels.

This paper provides a methodological account of the pilot based on conversations and reflections with the researchers who conducted the SenseMaker® pilot study. The intention is to discuss the strengths and weaknesses of this initial and relatively small pilot, as the basis for further exploring its use in IM&E of AMD and other IFAD-funded programmes. In Section 2, we first provide some background on the methodological debate in the literature. In Section 3, we give an overview of the added value and rigour SenseMaker® theoretically can bring compared to traditional methods, and subsequently assess to what extent it was able to demonstrate this in the AMD pilot. To support the findings, also a few other relevant experiences with SenseMaker® in international development were consulted. In the last Section, building on the learnings from this and other pilots and IM&E studies, we recommend some ***concrete next steps forward*** to integrate and refine the SenseMaker method and tools within the programme's IM&E system ***for generatively measuring impact*** in terms of adaptive capacity.

2 Background

Development problems and challenges are becoming more complex and politicised nowadays, and appropriate approaches and solutions less straightforward. This is likely to increase with the new Sustainable Development Goals (SDGs) demanding greater sustainability and inclusiveness in addition to development effectiveness. Dominant IM&E practices, however, still largely focus on results presumed to be predictable and traceable with the classic suite of survey-based methods. This works well in controllable and uniform environments, where problems and solutions are known. In more complex contexts, where behaviours change in more irrational ways with less predictable outcomes, traditional IM&E runs the risk of cognitive bias. Change in such contexts does not happen in the desirable straight and upward positive line. IM&E that builds on such premise is likely to draw invalid conclusions (Befani & Mayne, 2014; Woolcock, 2013). A narrow focus on measurable positive results moreover limits the ability to see broader patterns and timely detect signs of emergent positive or negative change essential for timely adapting strategies and plans. Hence changes that are *not* planned for (such as irrational responses to market failure and unstable weather conditions) are often overlooked or ignored (Hummelbrunner, 2010; Jenal, 2014).

A growing number of critiques warn against relying merely on traditional IM&E in complex contexts (Befani et al, 2015; Eyben et al, 2015; Guijt, 2011; Nobuko, 2010). Many argue that if change is contingent upon a high number of unpredictable interactions (also called "high causal density"), then IM&E should take a more open, proactive and explorative approach and collect data on all sorts of interrelated dimensions that plausibly may influence these interactions. An in-depth case study approach using action research is often suggested as a viable alternative (Burns, 2014; Khagram & Thomas, 2009; Woolcock, 2013). Yet in-depth case studies generally don't cover large populations. They serve the purpose of learning about causal mechanisms under particular conditions, but generally are of limited use to draw generalizable conclusions for instance for an entire province or country or region necessary for policy formulation (unless the case itself covers the entire geographic area). Moreover, because they seek to obtain an in-depth and comprehensive understanding of particular cases, they require more time to arrive at conclusions.

For policies and strategies in complex environments at a large scale, real time monitoring methods are needed that can collect quantitative and qualitative data from a large population, and show behavioural patterns and trends in very short time frames with very short feedback loops. Examples are Constituent Voice and SenseMaker®. These methods are often employed in combination with traditional surveys and case studies within a broader systemic IM&E approach (Jacobs et al, 2010; MOFA/GOG, IFAD, & BMGF, 2015; Van Hemelrijck et al, 2011).

3 The potential of using SenseMaker® for IM&E

There are two questions that need to be addressed when reflecting on the potential of using SenseMaker® for IM&E. The first concerns the **added value** (or utility) it offers compared to traditional methods in terms of process and type of knowledge. The second is about **rigour**, which also influences value and uptake by programme management and stakeholders (Casella et al, 2014; Patton, 2012).

3.1 Added Value of using SenseMaker®

3.1.1 Added Value in Theory

SenseMaker® is a methodology with patented software that facilitates mass ethnography and provides a way of nearly real-time mapping of social interactions and individual perceptions and motivations to inform adaptive management and policy formulation. It inquires the evolutionary potential of the present by collecting large amounts of data on people's experiences and how these shape their perceptions of *past* and *future* change in ways that enable us to see emerging patterns influencing actions and decisions. Hence it also fits backward-looking (or summative) and forward-looking (or formative) evaluation, while its strength and added value lays in real-time monitoring of emerging impact. It basically collects fragmented or anecdotal stories (called micro narratives) that are interpreted or signified by the storytellers themselves in relation to a set of interacting dimensions. The software allows us to conduct a statistical analysis of self-signified stories collected at a very large scale.

The method is fundamentally different from traditional methods in its ability to: (a) collect a large amount of *quantified qualitative* data from multiple sources; (b) conduct statistical analysis showing patterns and trends in a fairly short time and at a low cost; and (c) generate evidence that makes it possible to identify and respond to *emergent* opportunities or threats of change close to real time (Deprez et al, 2012). Also Constituent Voice collects and analyses large amounts of self-scored feedback data in a very short time and at an even lower cost. Yet it is different in its focus on performance of specific service delivery mechanisms, as opposed to SenseMaker® that inquires broader patterns and trends of change.

Classic surveys typically produce *estimations* of distributions (or percentages) of values for predetermined variables and characteristics across a large sample population (for instance 900 households). Case studies on the other hand produce *explanations* of changes in the interactions of certain dimensions in a very small sample of cases (for instance 5 villages). SenseMaker® does a bit of both: it generates *patterns* or distributions of interactive values for fragmented stories collected from a very large sample population (e.g. 3000 villagers). In contrast to surveys, it doesn't inquire into specific indicators assessing the intended effects of an intervention. Unlike case studies, it doesn't dig for in-depth explanations of change processes. It investigates patterns of emergent change at a very large scale and seeks to produce evidence of unknown or unexpected influences and effects that conventional methods are not as likely to detect (Jenal, 2014). Thus, it complements other methods and therefore suits mixed-methods IM&E.

3.1.2 Added Value in Practice

Overall, the SenseMaker® pilot in AMD did not reveal anything unique or significantly new. In this instance, it rather confirmed existing challenges and conditions regarding farmers' problems related to climate change, which were already known. This also happened in other SenseMaker® pilots in rural development – such as the pilot in the Triple-S Project of the IRC International Water and Sanitation Centre conducted in 2011-2012 in Ghana and Uganda (Casella et al, 2014).

What it did uncover though allowed for deeper understanding of how climate change as a concept was understood by farmers, and also revealed a different prioritisation of challenges than was at first imagined. This also occurred in other cases, such as for instance a project ran by Global Giving[1] that initially assumed the priority in a refugee settlement in Kenya was food and infrastructure, while the 2000 self-signified stories revealed inhabitants' most pressing need was the repair of social relations for controlling and influencing their access to resources and opportunities (Guijt & Hecklinger, 2010).

There are three plausible explanations for this lack of unique discovery in the AMD pilot:

1. The first is that M&E staff are already closely monitoring field activities and are aware of the problems and conditions in the villages. For instance: the limited access to markets and finance; farmers' dependency on labour and technology; their difficulty with developing viable businesses; as well as the thread of salinity, pests and diseases and rising temperature –all are quite well known and documented elsewhere (e.g. IFAD & BMGF, 2014). In other instances, SenseMaker® collected data on a mix of general scanning and programme-specific variables, which helped monitor both expected/intended and unexpected/unintended changes for reporting and learning. The question for AMD remains: How does, and in which areas, would incorporation of SenseMaker®, in addition to regular IM&E be valuable?
2. Second, one iteration of story collection and analysis is insufficient to reveal new or unexpected patterns and trends. Stories need to be collected at a large enough scale and over a long enough period to be able to do this. Moreover, the study was conducted with little clarity about how to frame "change in adaptive capacity" in the specific contexts of Ben Tre and Tra Vinh, since programme interventions at grassroots level (such as PAR) were yet to begin and take shape.
3. The third plausible explanation is that biases in the sampling and data collection may have influenced the results in a way that what was already known or assumed was confirmed. For instance, the attribution of livelihood improvements to scientists and government intervention suggests such bias. The occurrence of biases is discussed in the next section.

Also in terms of process, this pilot did not incorporate the elements of the full SenseMaker® process that would have made it more innovative. Data was extracted without involving villagers in sense-making of the patterns and validation of the findings and recommendations. This was done in a traditional workshop involving only programme staff and implementing partners. Hence there was no feedback loop back to the villagers that would foster learning. Such a feedback loop can only be established through multiple iterations of story collection and pattern analysis feeding back into the PAR and engaging the PAR actors in sense-making and validation.

[1] Cf. https://www.globalgiving.org/jcr-content/gg/landing-pages/story-tools/files/microsoft-powerpoint---makingsenseofsensemaker.pdf.

3.2 Rigour in using SenseMaker®

3.2.1 Rigour in Theory

In mainstream evaluation practice, rigour is generally understood as the controlled avoidance of bias (Befani et al, 2014; Camfield at al, 2014). Biases occur on the part of the researchers and the researched, and in every phase of inquiry. More broadly, rigour is also defined as the quality of thought put into every step of the design and conduct of an impact inquiry to ensure consistency and responsiveness to the context and knowledge needs at hand. Rigour helps establish confidence in findings and conclusions (Rogers, 2009; Stern et al, 2012; Van Hemelrijck & Guijt, 2016).

SenseMaker® builds on the premise that knowledge of complex change processes in contexts of high causal density can best be obtained by continuously scanning the entire field, detecting signals of potential change, responding to these signals, and continue scanning and detecting the reactions. It is designed in a way that removes intermediary levels of structuring, interpreting and analysing data, thus permitting rapid analysis and response, while avoiding researcher bias at each of these levels. It does so by collecting a large amount of fragmented –thus *unstructured*– experiences or narratives and letting respondents self-signify them, while enabling managers and decision-makers to directly access the raw narratives for making sense of the statistics produced by the SenseMaker® software (Deprez et al, 2012; Snowden, 2002).

The narratives are the kind of fragmented or anecdotal stories people tell around a campfire or in family gatherings[2], reflecting their perceptions of change processes in an unconstructed way. Tools such as *triads*, *dyads* and *stones* enable people to self-signify how their story *proportionally* relate to sets of competing characteristics and variables reflecting plausible trade-offs or choices, which also helps to avoid "gaming" that occur in traditional scoring with scales. By letting people index their own narratives, meta-layers of meaning or interpretation explaining their decisions and behaviour are generated that otherwise could not be obtained if done by the researchers. This process of indexing also permits quantification and statistical analysis. If used on a regular basis, SenseMaker® makes it possible to document patterns and trends of perceptions and behaviours over time across large populations. This enables managers and sponsors to make effective decisions about strategic directions and priorities for investment (http://cognitive-edge.com/SenseMaker®; Jenal, 2014).

[2] Cf. https://www.youtube.com/watch?v=SkRe7Xg7pk4.

3.2.2 Rigour in Practice

SenseMaker® is rigorous in practice in so far as it sufficiently pays attention to the risk of bias in every phase of the inquiry –from the design and sampling to synthesising evidence and formulating recommendations– and is consistent and responsive to context and conditions at any particular time.

Sampling

Through a workshop with the stakeholders from PCU, TVU, DARD, etc., a sample of ten villages (five each) from both Tra Vinh and Ben Tre provinces were identified. These villages were selected to mark different points along the salinity gradient so as to provide greater context for comparison and development of insight specific to the challenges faced by respective villages. They were Hoa An, Long Thoi, Trung Tho, Tan Thanh Binh, Hiep Hoa, Luong Phu, Long Son, An Duc, Long Vinh, and Thanh Phong. A total of 500 stories were collected over a period of one month in April 2015 via pen and paper interviews, and a follow-up with 100 selected farmers was conducted through telephone interviewing in the months of July and August 2015 to ensure that the data would enable better analysis for the development of the participatory action research agenda.

The reflections on the AMD SenseMaker® pilot revealed that:

a) The sampling of the farmers, as well as the introduction of the study to the farmers, was done by programme staff;
b) Researchers introduced themselves as *scientists* who sought to help and support farmers to solve their problems and increase their income; and
c) Farmers were mostly motivated to participate because they hoped to obtain some sort of support or benefit from the programme such as training, services, or finance.

This indicates a relatively high risk of *anchoring* and *confirmation bias* (Copestake, 2013). Indeed, the likelihood that the selection was self-serving programme staff, and that farmers anchored their responses to their expectations and experiences with government programmes, is quite high. Hence findings may just confirm what was already known and expected, due to this bias. Furthermore, as randomisation in terms of sample selection was difficult, impractical, and costly to achieve, this could have resulted in selection bias, which means that the sample was not representative of the entire population of farmers.

In future SenseMaker® studies, careful thought should be put into sampling. The purpose of SenseMaker® is to collect a large amount of widely dispersed and uninfluenced information to enable statistical analysis. The dataset of single stories collected from 500 farmers in the AMD pilot did not provide enough depth and coverage of the complexity and diversity of farmers' day-to-day challenges. Although it is important to move beyond simplistic sampling targets and traditional researcher-led interview capture towards more organic modes of continuous data collection (e.g. by using a farmer diary system), it is equally important that the study maintains sufficient independence and the sampling is unbiased and representative of the population to be inquired. Particularly if using SenseMaker® for IM&E in a mixed-methods approach (thus in combination with other methods), its sampling strategy and frame must be suitable for the entire suite of methods to enable the linking of evidence. The

SenseMaker® mini-pilot conducted in Ghana in the recent impact evaluation of the IFAD-funded "Roots and Tubers Improvement and Marketing Programme" (RTIMP), for instance, demonstrated this principle[3] (IFAD & BMGF, 2015).

A critical question, however, is how to deal with bias when using SenseMaker® for *participatory* impact monitoring –e.g. when participants freely choose to participate, self-capture their experiences in their diaries, and self-signify their stories on an ongoing basis. In such contexts, the "crowd sourcing" of stories should be large enough to outweigh a possible sample bias. Moreover, if this feeds into a PAR in which farmers are not just receivers but also designers of appropriate responses to the problems identified, the risk of self-serving bias is turned into an instrument of empowerment that contributes to greater impact –also called "generative causation" (Pawson, 2013; Van Hemelrijck, 2014). Essential to make this happen, however, is to ensure that the power of the crowd levels out the potential risk of elite dominance (Chambers, 2015). Important in this respect is also to carefully think through how to structure the processes in which farmers are engaged in sense-making and validation of findings.

Data collection and capturing

Experiences with SenseMaker® in agricultural development show that its tools (e.g. triads, dyads and stones) are not easy for farmers to understand. They tend to struggle with grasping the abstract shapes in which they have to situate their story. Illiterate farmers, particularly women, tend to feel intimidated when given a marker to place their dot in abstract figures on a paper with text they can't read. Where the tools are used in focus groups, this may create certain power dynamics between those who can understand the tools and read the text, and those who cannot. Moreover the tools are also challenging for local translators to grasp and translate, which causes great difficulties in contexts where there are many different local languages and dialects[4]. All this may affect the quality of data, and thus the credibility of findings (IFAD & BMGF, 2015).

From the reflections with the researchers in the AMD pilot, it appears that they mastered the tools quite well and so there wasn't a major issue with explaining them to the villagers. They were able to do this quite well by using appropriate examples. This is largely due to the detailed guidance and training that was provided to them. But they did report they needed to be more patient with women to help them understand the tools and overcome their shyness. In future SenseMaker® studies, collecting the data in gender-specific focus groups, and drawing the shapes in the sand/soil while using classic PRA[5] techniques such as proportional piling with pebbles or beans to facilitate the self-signification process, might help overcome inhibitions and power dynamics. This was done for instance in the RTIMP evaluation in Ghana[6] (MOFA/GOG, IFAD, & BMGF, 2015). Technology such as iPads also

---

[3] The report cam be downloaded from: http://www.ifad.org/english/piala/resources/RTIMP_ier_study.pdf. SenseMaker® was employed in combination with four other methods inquiring different aspects of the programme and collecting different types of evidence to permit cross-checking and configurational analysis. All were used in a countrywide random sample of 30 value chain areas across 25 random districts, from which households and beneficiaries were (quasi-)randomly subsampled, representative of the different populations to be inquired.

[4] In Ghana, for instance, there are more than 80 languages, while the researchers spoke maximum 12 languages.

[5] Participatory Rural Appraisal.

[6] The SenseMaker® mini-pilot was conducted in 28 men-specific focus groups involving 211 male farmers and 26 women-specific focus groups counting 189 female farmers.

seem to work well, as the tactile experience of a touchscreen appears to be more intuitive for people with low literacy who are easily intimidated by pen and paper. The technological interface increases curiosity and interactivity, and decreases their fear of "words". The ability to record voice and take direct images also may help overcome this fear.[7] The iPad capture has been used in communities with low literacy in many SenseMaker projects to successful effect (Cordaid, 2015; Girl Hub, 2014; IFRC, 2014). Although iPads were rented to enable voice recording of the narratives in the AMD pilot, these were not fully utilized as the team chose to record the narratives in point form instead. Thus, there was the difficulty of having to working past scientists' biases.

As observed from the pilot, it is crucial that additional training be conducted in the method of data capture as the field team struggled to grasp the concept of narrative research as being verbatim capture. The researchers had interpreted the stories in their own way and captured them as bullet-point lists. These obviously did not reflect people's unrestrained anecdotes and likely generated a bias making the findings more reflective of the researchers' expectations. Thus, the quality of the data was compromised where there was an attempt to rephrase, rewrite and even influence the farmer's narratives.

An additional round of mobile interviews was conducted with a sub-sample of farmers to compensate for this. Given the importance of avoiding bias, either audio recording or closer quality monitoring of data capturing during fieldwork is highly recommended in future SenseMaker® studies. This also, again, emphasises the need to move beyond traditional models of research and data capture to create a truly organic IM&E feedback loop (by for instance using a farmer diary system for data collection and snowballing technique for expanding the samples).

Therefore, it would be useful to implement data capture in a more organic and naturally occurring way, as opposed to using traditional scientist-led interviews. For instance, the journaling method could be tested with a sample of farmers from different villages identified based on certain parameters or characteristics. These farmers would log in daily or weekly entries about their farming techniques, use of fertilisers, and progress on farming, as well as report on any connections with business and commercial forces. In this method of data collection, it would be more important to have richness of data i.e. quality over quantity. Therefore, a sample size of 300-500 journaling over a period of a year would be more than sufficient to garner valuable insights on how to improve farming methods and techniques.

Synthesising evidence and formulating recommendations

The analysis of the 500 self-signified stories was done using the SenseMaker® software. The patterns that emerged from this analysis were interpreted by the researchers. A draft report was produced that synthesised the evidence of the patterns and the findings from researchers' readings of these patterns, with recommendations for the PAR agenda building on these. This synthesis was presented to key stakeholders (incl. programme staff, scientists from participating universities, provincial officials, and technical advisors from IFAD) for validation in a two-day workshop on 14-15 September 2015.

[7] Obviously there are contexts in which recording and pictures might put participants into danger, but research ethics apply to any type of research and any type of tool used for recording.

This process of synthesising the evidence and formulating recommendations seem to have taken place quite rigorously. The patterns show clear knowledge gaps to be tackled through the PAR. But this one-time analysis of 500 stories cannot demonstrate the potential of the method to rigorously track changes in perceptions and behaviours as a result of the PAR activities *over time*. If we want to use the method for impact monitoring, then it needs to not just signal knowledge gaps, or potential threats and opportunities for change, but also document the *pathway of change* that emerges from the interactions between our responses to the signals and the systems we expect to change –in the case of AMD: between the activities planned and implemented through the PAR and the adaptive capacity of communities' farming systems[8].

[8] Incl. infrastructural issues, inputs, technologies, management, finance and market links for agriculture, aquaculture and animal breading.

3.3 Summary of findings

In this first SenseMaker® pilot in AMD, the method did not succeed to produce new insights and demonstrate its added value. But it was only used once to collect and analyze 500 stories. This is largely insufficient to demonstrate value and make a judgment of the potential for IM&E. The method is designed to detect unexpected signals of change by inquiring into trends and patterns among large populations on an ongoing basis. For this to be successful, the method needs to be well-tailored to the specific contexts and needs of the programme. Moreover, staff and managers using the data in both the implementing and the funding agencies need to sufficiently understand what the method *can* and *cannot* do. ***A single iteration is unlikely to make this happen.***

Despite the fact that there were no big surprises, the pilot did show some of the potential for advocacy and action research. The analysis of patterns in farmers' perceptions of changes in their livelihoods and the competing problems and solutions influencing these changes, undoubtedly generated a useful evidence-based overview of essential knowledge gaps and recommendations for shaping the PAR agenda. If the method of capture would be further refined and adjusted to the specific contexts of the two provinces, and used on an ongoing basis and at a larger scale, it will likely produce evidence that is useful for the policy dialogues at the provincial and national levels.

Issues of pests and diseases reflected strongly in the stories shared by the farmers, particularly in plantation farming and livestock rearing. To combat this, farmers resorted to the use of fertilisers. Unfortunately, these were used in excess and the debilitating effects of overuse affected the health of their crops over time. Furthermore, these fertilisers were often purchased from unreliable sources and had long-term negative impacts on the environment and agricultural ecosystem.

Farmers had also expressed difficulty in distinguishing between genuine and fake fertilisers that they had bought in the market. Thus, there should be some form of training and education for farmers to be able to work out the best option when it comes to purchasing and applying fertilisers.

The other issue was interaction with business. A majority of the stories indicated that businesses posed a challenge to them. This could imply that there is a need to promote mutually beneficial relationships between farmers and businesses.

No single method is bias-free. Every inquiry needs to be designed in a thoughtful and careful manner to minimise the risk of bias or dominance of a particular perspective. For SenseMaker® to be rigorous and useful for IM&E to influence policy and planning, serious thought needs to be put into the sampling (or crowd sourcing) frame and strategy. Such a frame and strategy must enable large-scale unbiased story collection for statistical analysis and the linking of evidence from different methods, necessary for probing the programme's hypotheses and making timely adjustments. Moreover, rigorous employment of SenseMaker® also requires that the stories are captured in their original form, and that the challenges related to the use of abstract tools with different (incl. illiterate) groups in different languages are properly addressed. Not paying attention to this may affect the quality of the evidence and thus the credibility of the findings. Last, more thought should be put into the design and facilitation of collective sense-making and validation processes to avoid elite dominance and thus bias, and ensure equal voice in selecting and planning appropriate programme responses.

4 Next steps for integrating SenseMaker® in IM&E of AMD

Expectations in a first pilot are often quite high and generally unrealistic. Sponsors and managers expect magic and tend to forget that innovation takes time and needs sufficient resourcing to mature. They also tend to forget that piloting involves *learning* in order to bring the concept to its mature stage. Thus by expecting that innovation pilots will do magic, one forgets that the only real failure in piloting is the expectation itself causing ***failure to learn.*** We therefore recommend IFAD and AMD to see through the learning process started with this initial small pilot, as to make sure that people's time and the resources invested are generating true learning and not wasted on failure-to-learn.

We recommend the following steps to operationalize SenseMaker® for Impact M&E of AMD:

1. As a first step, we recommend to abstract indicators for IM&E from the findings of this first SenseMaker® study that will help track the influence of AMD, in particular its PAR activities, as well as other influences and their interactions, on adaptive capacity and behavior. These indicators would complement regular programme performance indicators by tracking unexpected or unintended influences alongside those intended and created by the programme. They serve to scan and inquire the systemic interplay of various influences in the broader environment. This can be done through the creation and use of ***narrative landscapes*** (or *probability density maps*) based on the correlational analysis of the various triad and dyad variables. The landscapes show how the different sets of competing variables reflecting the trade-offs or choices in the stories (or in the experiences presented by the stories) interact, thus showing how a complex set of influences affect farmers' adaptive behavior. Analyzing the landscapes helps to identify variables or indicators requiring greater attention for generating the desired changes in adaptive capacity. By doing so, this will also help reframe community engagement in the PAR. A concrete example from the findings of this first SenseMaker® pilot is presented in the Text Box on the next page.
2. Second, the process of identifying indicators also forms the basis for refining the design of the SenseMaker® framework and tools to make it more concise, focused and contextually relevant. Integrating programme-specific performance indicators with the more abstract environment-scanning indicators in the SenseMaker® framework is one of the tasks. Another involves the adaptation of the data collection tools and means. The option of using a simple diary-entry style data collection mechanism that can be disseminated across farming communities could first be tested with participants in the PAR and after 2 or 3 cycles of field testing expanded to the wider communities (by using for instance a snowballing technique) to arrive at sufficiently large sample sizes.
3. Third, as mentioned in the section above on data capture, the initial budget of USD$60,000 was meant to incorporate the use of iPads. However, as a result of the way the data capture was done, additional funds had to be budgeted on top of the initial amount allocated in order to accommodate costs of traditional field capture, data entry, translation, and so on. Therefore, there is a need for more intensive training in data collection techniques among the data collectors/field workers.
4. Finally, if sample representativeness is key, there would be a need to gather large-enough samples that are part of a well thought-through **cluster sample frame**, permitting data integration and linking of findings from SenseMaker with evidence obtained from other IM&E methods, will build coherence in and thus strengthen the IM&E system for AMD. This can be

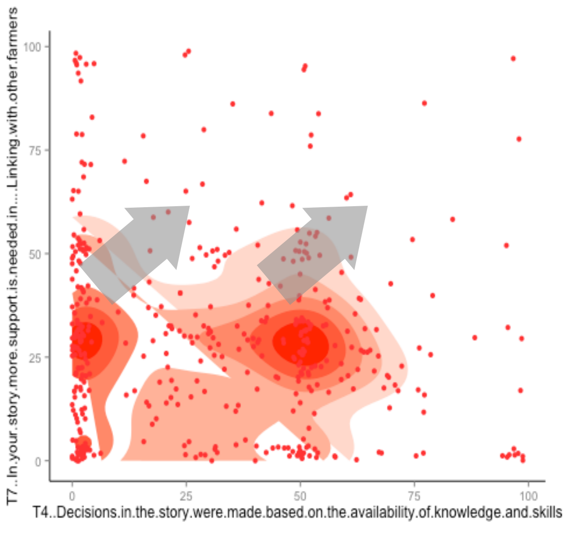
done by defining the **catchment areas of the PARs** as the principle sample unit for measuring changes in adaptive capacity. From these catchment areas, we then can subsample:

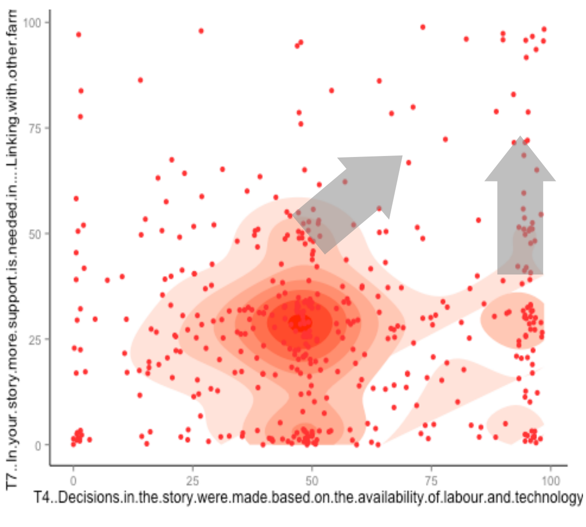
- intended programme beneficiaries for collecting SenseMaker data;
- households for collecting classic survey data;
- service providers for collecting performance data;
- other populations for data collection using other methods.

A **configurational analysis approach** could then be applied for linking all the data from the different methods in each of the catchment areas, and comparing and analyzing different **configurations of PAR treatments, conditions and outcomes** emerging from the evidence across the 2 provinces, in order to arrive at rigorous causal inference and draw conclusions about programme contributions to impact. Such an approach was successfully used in the RTIMP impact evaluation in Ghana (MOFA/GOG et al., 2015).

Example of using landscapes for identifying PAR intervention opportunities and indicators for monitoring

Quantitative indexing of stories provide a statistical basis through which we can map and understand the probabilities of interactions between different conditions and dispositions in the lived and shared experiences of community members, and the way we might be able to influence. The two landscapes presented below, for instance, depict the variables of two triads in the SenseMaker framework developed for AMD: Triad 4 (or T4) which is about the basis for decision-making and Triad 7 (or T7) which is about support needs. The y-axis in both charts represents the need for more support to link with other farmers, which was identified in this study as a desired form of peer-to-peer sharing and support amongst the farmers. The landscapes show how this is mapped against variables of decision-making –on the left *"availability of knowledge and skills"* and on the right "*availability of labour and technology"*. Each individual dot in the landscapes represent a story or lived experience that was collected and self-signified in the AMD pilot.





The red-shaded areas show clusters of stories revealing opportunity for change. The deep red areas represent deep-sunk dispositions that are difficult to influence. For instance, in the right image, the deep-red cluster reveals a relatively low need for support to link up with other farmers and a medium level of dependency on labour and technology for decision-making, which is quite dominant and difficult to change. The smaller cluster on the extreme right, however, is shaded less deep-red, thus revealing an area of *emergent* or *plausible* change that is easier to influence. If we would want to see the dots moving up towards greater famer-to-farmer support (and less dependence on merely market-linking and training services as the other variables in T7 imply), assuming that this would positively influence communities' adaptive capacity, then we might think of conducting small interventions that would move the deep-red cluster of dots *first sideways to the right*, thus into the adjacent possibility of greater decision-making based availability of labour and technology, before trying to move it upwards. We would then closely monitor these interventions on these two variables (decision-making based on availability of labour and technology, and need for farmer-to-farmer linking and support). PAR interventions, thus, can be developed based on the analysis of farmers' stories in clusters that are easier to influence, in order to amplify the trends that are positive or desirable and dampen the ones that are rather negative or undesirable. The use of landscapes helps us shift from a merely **static index-based** to a more **dynamic vector-based approach of impact monitoring**, building on an understanding of impact as change emerging from recursively interacting and constantly evolving conditions and dispositions, which can be attenuated or amplified and directed by crafting small interventions to the opportunities revealed by the evidence for influencing the patterns. Such a vector-based monitoring approach would enable the PAR in the AMD develop a higher level of sensitivity and capacity to respond to the many influences affecting adaptive capacity and behaviour.

References

Befani, B., Barnett, C., & Stern, E. (2014). *Introduction – Rethinking Impact Evaluation for Development.* IDS Bulletin, 45(6). IDS.

Befani, B., & Mayne, J. (2014). *Process Tracing and Contribution Analysis: A Combined Approach to Generative Causal Inference for Impact Evaluation.* IDS Bulletin, 45(6). IDS.

Befani, B., Ramalingam, B., & Stern, E. (2015). *Introduction – Towards Systemic Approaches to Evaluation and Impact.* IDS Bulletin, 46(1). IDS.

Burns, D. (2014). *Assessing Impact in Dynamic and Complex Environments: Systemic Action Research and Participatory Systemic Inquiry.* CDI Practice Paper (8). IDS.

Camfield, L., Duvendack, M., & Palmer-Jones, R. (2014). *Things you Wanted to Know about Bias in Evaluations but Never Dared to Think.* IDS Bulletin, 45(6). IDS.

Casella, D., Magara, P., Kumasi, T. C., Guijt, I., & van Soest, A. (2014). *The Triple-S Project Sensemaker® experience: A method tested and rejected.* (Triple-S Working Paper No. 9). IRC.

Chambers, R. (2015). *Inclusive rigour for complexity.* Journal of Development Effectiveness, 7(3).

Copestake, J. (2013). *Credible impact evaluation in complex contexts: Confirmatory and exploratory approaches (Draft 18 Oct 2013).* Centre for Development Studies, University of Bath.

Cordaid. (2015). *Flourishing community index: Giving local communities a global voice.* (Factsheet). Cordaid.

Deprez, S., Huyghe, C., & Van Gool Maldonado, C. (2012). *The use of Sensemaker for measuring, learning and communicating about smallholder farmer inclusion.* (Case Report). VECO.

Eyben, R., Guijt, I., Roche, C., & Shutt, C. (2015). *The Politics of Evidence and Results in International Development: Playing the Game to Change the Rules?* Practical Action Publishing.

Girl Hub. (2014). *Using SenseMaker® to understand girls' lives: Lessons learned from Girl Hub*. Girl Hub.

Guijt, I. (2011). *Evaluation Revisited: Improving the Quality of Evaluative Practice by Embracing Complexity. Conference Report. (May 20-21, 2010).*

Guijt, I., & Hecklinger, J. (2010, Nov). *Making Sense of Sensemaker®: Evaluating Development Initiatives through Micro-Narrative Capture and Self-Tagging in Kenya.* Presented at the AEA Conference.

Hummelbrunner, R. (2010). *Beyond logframe: Critique, Variations and Alternatives.* In: *Beyond Logframe; Using Systems Concepts in Evaluation*. FASID.

IFAD, & BMGF. (2014). *Impact Assessment of the "Developing Business with the Rural Poor" Programme in Bến Tre, Viet Nam (2008- 2013) in Ben Tre, Vietnam. Report of the Pilot Application of a Participatory Impact Assessment & Learning Approach (PIALA)*. IFAD and BMGF.

IFAD, & BMGF. (2015). *Improved Learning Initiative for the design of a Participatory Impact Assessment & Learning Approach (PIALA): Methodological reflections following the second PIALA pilot in Ghana*. IFAD and BMGF.

IFRC. (2014). *IFRC Southern Africa grassroots consultations: 1000 stories from Southern Africa. Preliminary report.* International Federation of Red Cross and Red Crescent Societies.

Jacobs, A., Barnett, C., & Ponsford, R. (2010). *Three Approaches to Monitoring: Feedback Systems, Participatory Monitoring and Evaluation and Logical Frameworks.* IDS Bulletin, 41(6). IDS.

Jenal, M. (2014). Monitoring in complex systems: from spotlight and indicators to sensor networks and landscapes. Retrieved from https://marcusjenal.wordpress.com/2014/04/16/monitoring-in-complex-systems-from-spotlight-to-sensor-networks.

Khagram, S., & Thomas, C. (2009). *Evidence for development effectiveness.* The Journal of Development Effectiveness, 1(3).

MOFA/GOG, IFAD, & BMGF. (2015). *Final Report on the participatory impact evaluation of the Root & Tuber Improvement & Marketing Program (RTIMP). Pilot Application of a Participatory Impact Assessment & Learning Approach (PIALA)*. IFAD.

Nobuko, F. (2010). *Beyond Logframe; Using Systems Concepts in Evaluation*. FASID.

Patton, M. Q. (2012). *A utilization-focused approach to contribution analysis.* Evaluation, 18(3).

Pawson, R. (2013). *The science of evaluation: a realist manifesto*. SAGE.

Rogers, P. (2009). *Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation.* The Journal of Development Effectiveness, 1(3).

Snowden, D. (2002). *Complex acts of knowing: paradox and descriptive self-awareness.* Journal of Knowledge Management, 6(2).

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). *Broadening the range of designs and methods for impact evaluations.* (Working Paper No. 38). DFID.

Van Hemelrijck, A. (2014). *Understanding "Rigor". Challenges in Impact Evaluation of Transformational Development.* PhD Research Outline Paper (Revised). IDS.

Van Hemelrijck, A., Dirirsa, B., & Ling, A. (2011). *Joint MEL framework for Strengthening the Ethiopian Agricultural Extension System (SEAES). Framework paper approved for funding by the BMGF.* (p. 33). Oxfam America, Sasakawa Africa Association and the MOFA/GOE in collaboration with ALINe and Keystone Accountability.

Van Hemelrijck, A., & Guijt, I. (2016). *Balancing Inclusiveness, Rigour and Feasibility. Insights from Participatory Impact Evaluations in Ghana and Vietnam.* CDI Practice Paper (14). IDS.

Woolcock, M. (2013). *Using case studies to explore the external validity of "complex" development interventions.* Evaluation, 19(3). Sage.