Outcome Evidencing: A Rapid and Complexity-Aware Evaluation Method

Douthwaite, B.[1] and R. Paz-Ybarnegaray[2]

[1] *Independent Learning and Evaluation Researcher and Consultant, Cushalogurt, Kilmeena, Westport, Ireland. bdouthwaite@gmail.com*
[2] *Independent Evaluator*

Abstract: This paper describes the development and use of a rapid evaluation approach to meet program accountability and learning requirements in an on-going research for development program operating in five developing countries. The method identifies clusters of outcomes to which the program has contributed, within program areas of change. In a workshop, change agents describe the causal connections within outcome clusters to identify outcome trajectories for subsequent verification. Comparing verified outcome trajectories with existing program theory allows the program to question its underlying causal premises and adapt accordingly. The method can be used for one-off evaluations that seek to understand whether, how and why program interventions are working. Repeated cycles of Outcome Evidencing can build a case for program contribution over time that can be evaluated as part of any future impact assessment of the program or parts of it.

Key words: Complex systems, outcome trajectories, theory of change, monitoring, evaluation, socio-technical niches, realist evaluation, mechanisms, innovation.

Introduction

Agricultural research for development programs intervene in complex adaptive systems fashioned by people and the agroecologies in which they live. In complex adaptive systems there are rarely ever any magic bullets: no intervention will ever work the same way, everywhere for everyone. In some contexts, some program offerings will work and in others they will not (Pawson, 2013). Evaluation methods therefore need to understand how different aspects of programs work, for whom in different contexts. In other words, they need to unpack the causal black box between program intervention and program outcomes (Astbury and Leeuw, 2010). Most traditional impact evaluation methods do not dig into causality, but rather concentrate on establishing the worth of program intervention, often evaluating it against whether its initial predicted routes to impact have come to pass (Mayne and Stern, 2013; Stern, 2015). Such methods are of little use to staff interested in understanding how their interventions are working so as to improve implementation and the chances of reaching larger numbers of people. Nor are they useful to donors interested in improving their returns on investment by making better investment decisions. Traditional impact evaluation methods risk failing to identify and learn from the parts of the program that are working and have the potential, if supported and scaled, to make a real difference.

The literature that calls for complexity-aware impact evaluation to fill this gap is large and growing (e.g., Patton, 2011; Stame, 2004; van Mierlo et al., 2010; Mayne and Stern, 2013; Rogers, 2008; Douthwaite et al. 2003). The literature has less to say about the experience of developing and using complexity-aware impact evaluation methods and how they work, or not, in programs that are themselves complex and on-going. This paper describes the development of a complexity-aware method called *Outcome Evidencing* within a systems-focused research for development program of the CGIAR. The CGIAR is a worldwide partnership addressing agricultural research for development carried out by 15 research centers through fifteen CGIAR research programs. CGIAR work contributes to the global effort to tackle poverty, hunger and environmental degradation.

Our objectives are two-fold: to describe and critically reflect on an evaluation approach that may be of interest to other programs and to share the practical considerations involved in starting to use complexity-aware evaluation methods.

The need for complexity-aware evaluation in AAS

The goal of the CGIAR Research Program on Aquatic Agricultural Systems (AAS) is to improve the wellbeing of poor people dependent on aquatic agricultural systems by putting in place the capacity for communities to pull themselves out of poverty (AAS, 2011). AAS began in 2011 by establishing programs of work in five geographically defined hubs with an aspirational goal to make positive difference on the livelihoods of 6 million poor and marginalized by 2023 (AAS, 2014). By the end of 2013 AAS was implementing programs of work in the coral triangle of Solomon Islands and the Philippines, the Asia mega deltas of the Mekong and Ganges–Brahmaputra–Meghna river systems (Cambodia and Bangladesh), and the African freshwater systems of the Niger and Zambezi rivers (Zambia), all of which are complex socio-ecological systems, where millions of poor and marginalized small-scale

fishers and farmers make a living. Issues facing these aquatic agricultural systems are often complex because they arise from deep-rooted, complex, interrelated processes that operate across and between different scales from global to local and cannot be understood by separating them out for analysis by single academic disciplines (Halliday and Glaser, 2011).

In the same period the program developed the research in development (RinD) approach as its main vehicle for achieving impact. The RinD approach allows research teams to work as part of a coalition of stakeholders to jointly tack a broad development challenge. The RinD approach creates new and safe dialog and action spaces for stakeholders to engage with one another long enough to build trust, motivation, capacity and insight to do things differently. AAS overarching program theory is based on the premise that agricultural research processes (e.g. multi-partner collaborations) and outputs (i.e. new technologies) work to catalyze and foster processes of rural innovation. It is these innovation processes, that maybe technical, institutional or both, that lead to development outcomes. The RinD approach is a way of building collaborations across institutional and scale boundaries (e.g. between farmers and researchers, or between different government ministries).

The authors, both with responsibility for program evaluation, were aware that the investment being made in AAS was contingent on demonstrating that the RinD approach is working within the first phase of the program scheduled to end in 2016. We expressed the evaluation challenge in terms of two evaluation questions:

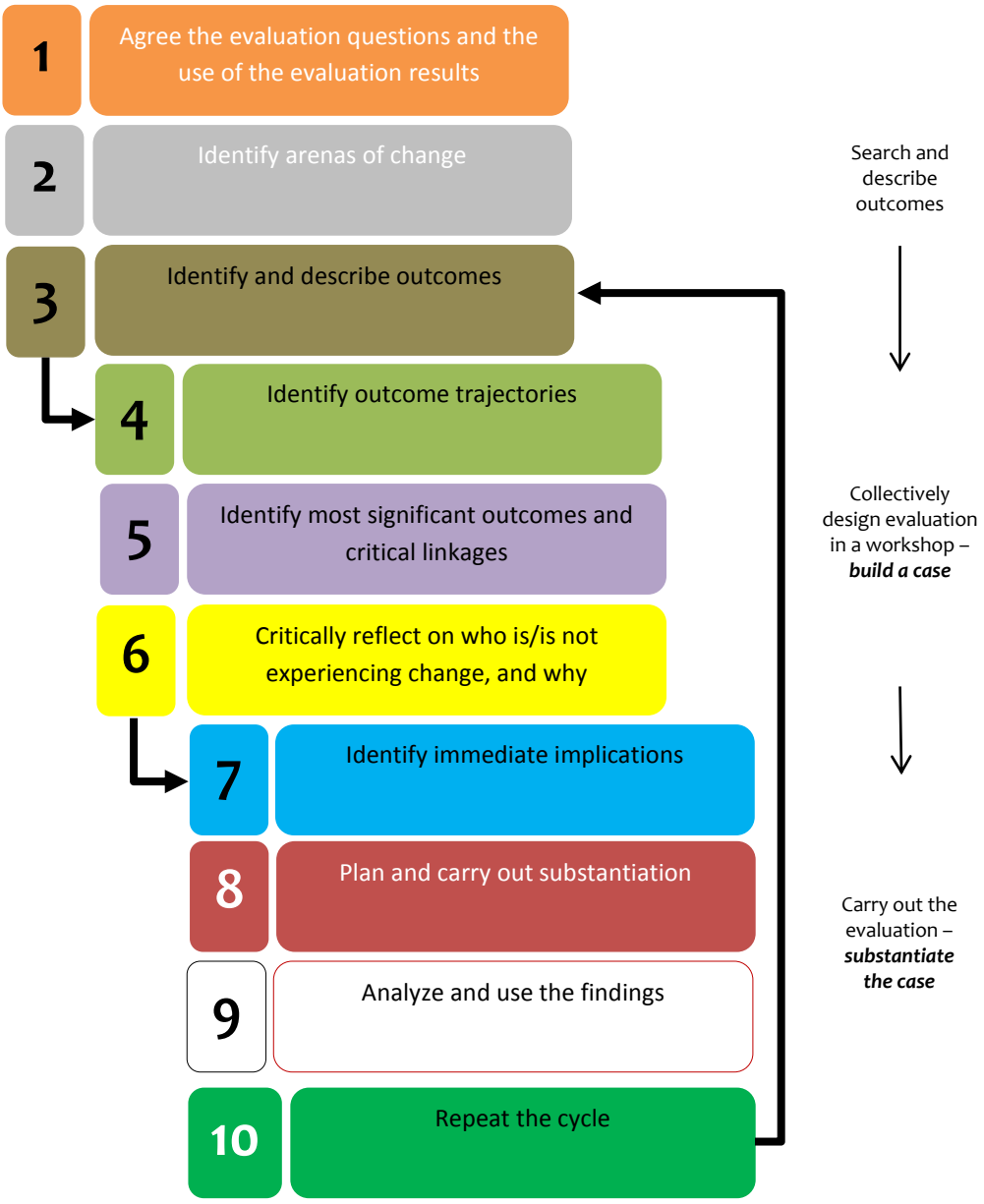
- What types of outcomes is AAS contributing to?
- Do these provide evidence that the overall program theory of change is credible, and how do they help us understand why (or why not)?

The questions were equally motivated by systems thinking (Snowden, 2010) that the way to trigger change in complex systems is to support emerging patterns of positive outcomes resulting from AAS intervention, and at the same time dampen down changes detrimental to the program's beneficiaries. This is similar to Rogers' (2008) idea that program theory can may be used to identify emergent outcomes that have the potential to make a big difference. To work in this way, the program needed a method of quickly identifying emerging outcomes, both expected and unexpected.

The Outcome Evidencing Method

We combined elements of Outcome Harvesting (Wilson-Grau and Britt, 2012) and Scriven's (1976) Modus Operandi methods to meet AAS' evaluation challenge. Outcome Evidencing has ten steps shown in Figure 1 and described below.

Figure 1: Ten steps of an Outcome Evidencing process



Step 1: Agree the evaluation questions and the use of the evaluation results

Step 1 involves program staff agreeing the evaluation questions and how the evaluation results will be used. The AAS question was: "what are the areas of change to which the program is contributing, and how is it doing so?" AAS uses the results to justify its funding and to help achieve impact by early identification of promising areas and early understanding of what the program is doing that is working.

Step 2: Identify key areas of change

In Step 2 knowledgeable program staff identify areas of change to which the program is contributing. These areas of change can be understood as emerging 'socio-technical niches'. Niches are spaces where people experiment with novelty in technology and/or institutions (Klerkx et al. 2012). It is these niches that the program wishes to identify early and support. Niches are a core concept of strategic niche management (Kemp et al., 1998). According to

this theory, when niches are properly constructed and linked they can act as building blocks for broader societal changes towards sustainable development (Schot and Geels, 2008). Hence strategic niche management provides some detail to the AAS' program theory described above, specifically that the program creates, supports and guides socio-technical niches to be building blocks that come together to help achieving the program's goal. Focusing rapid evaluation on if and how program intervention is contributing to niches was a way of answering the second evaluation question relating to the credibility and workings of the AAS program theory. Evaluation findings can guide how the program intervenes in the future to link the niches to bring about broader change.

Step 3: Identify and describe outcomes

Step 3 is to identify and describe outcomes occurring within the identified areas of change. This is done through asking field staff and looking for outcomes recorded in process documentation, particular records kept by field staff. Either way, the outcomes should be described in terms of a single phrase that can be written on card to allow for subsequent clustering in a workshop. Other basic information should also be recorded for each outcome on a simple template. Given that more than 50 outcomes might be identified, filling out any template should not be too onerous.

The next three steps take place in a participatory workshop attended by staff and stakeholders involved with implementing the program in the field. The workshop identifies outcome trajectories by which the program is contributing to areas of change. The outcome trajectories, described as theories of change, identify and explain the causal links connecting program intervention to outcomes contributing to the areas of change. The workshop identifies critical parts of these theories of change for substantiation, and identifies sources of evidence.

Step 4: Identify outcome trajectories

Outcome trajectories are the patterns of change that the program is generating within the areas of change. They are similar to Scriven's Modus Operandi. Scriven argued that interventions, like criminals, have a modus operandi that is recognizable. Just as identifying criminals' modus operandi can help catch them, so identifying programs' modus operandi can help improve them by understanding how the program is or is not working. In Outcome Evidencing, identifying outcome trajectories happens in an annual workshop. Participants first cluster outcomes that they think are related. They then build a causal diagram as a way of collectively agreeing on what those relationships are, and in doing so add in or reject some outcomes. The outcomes and the links between them constitute outcome trajectories. Outcome trajectories are characteristic causal patterns of outcomes, with momentum, contributing to larger or more aggregate impact within and across the identified areas of change.

From a realist evaluation perspective, trajectories of change are mid-level theories of change that take place within a particular context, involve a causal mechanism or mechanisms that produce an outcome or outcomes (Westhorp, 2014). Causal mechanisms are what intervenes between the delivery of program service and the occurrence of outcomes of interest (Weiss, 1997). A mechanism is the response program activities generate in those

involved. The responses happen in peoples' heads and are generally hidden and sensitive to context. Mechanisms have causal power. Identifying outcome trajectories is a way of identifying and describing underlying mechanisms. Box 1 describes causal mechanisms in more detail.

Box 1: Examples of causal mechanisms

The concept of causal mechanisms is fundamental to realist evaluation but is also a cause of misunderstanding (Westhorp, 2014). Gravity is an example of a causal mechanism in the physical world. Gravity is what causes an apple to fall from my hand to the ground. Whether the apple falls or not depends on whether I release my grip. Letting go of the apple is the trigger. Social norms are an example of a mechanism in the social world (Elster, 2007). Social norms suggest a certain way of acting in particular circumstances. For example, whether I act in accordance to the expected behavior of not talking on my mobile in a train carriage will depend on triggers such as a disapproving glance from a fellow passenger or a sign asking passengers to respect others' wish for quiet. The outcome of triggering a mechanism depends on context. If I release an apple at the bottom of a swimming pool it will float because buoyancy replaces gravity as the dominant mechanism. Whether I make a phone call in the railway carriage will depend on the urgency of the situation. Both gravity and social norms are real, but their working is not directly observable. The 'under the surface' nature of mechanisms is a fundamental characteristic.

Step 5: Identify most significant outcomes and critical linkages in the outcome trajectories

The next step is to identify the critical outcomes and linkages within outcome trajectories upon which the program's claim to have made a contribution most depend. Outcome trajectories are theories of change. According to Popper (1992: 94 as quoted by Pawson, 2013: 9) theory is built and verified with the accumulation of explanation, rather than on the bedrock of observational facts.

> "The empirical basis of objective science has thus nothing 'absolute' about it. Science does not rest upon rock-bottom. It is like a building erected on piles. The piles are driven down from above into the swamp, but not down to any natural or 'given' base; and when we cease our attempts to drive our piles into a deeper layer, it is not because we have reached firm ground. We simply stop when we are satisfied that they are firm enough to carry the structure, at least for the time being."

We think Popper's swamp-building analogy helps explain the importance of this step. Some piles in the outcome trajectories are more crucial for understanding and substantiating program impact claims than others: these require greater scrutiny. The scrutiny helps clarify the program's unique modus operandi -- the distinctive set of underlying causal mechanisms that the program is triggering. If the program's claim to contributing to significant outcomes and critical linkages stands scrutiny, if firm enough ground can be reached, the building can continue. If not, the building needs to take on a different shape, and donors informed of the change.

Step 6: Critically reflect on who is experiencing change, and who isn't

AAS uses research to trigger or support processes of innovation. Innovation processes benefit participants more than non-participants (Rogers, 2010). AAS' goal, shared with many other programs, is to benefit the poor and marginalized who are usually by-passed by mainstream development activity. Hence we include a step that involves analyzing outcome trajectories in terms of social and gender equity, inclusion and power. This information helps AAS catalyze, support and modify outcome trajectories to favor poor, vulnerable and marginalized groups, and correct the course of and even curtail potentially harmful ones.

Carrying out this step requires a context-specific understanding of inequalities and gender norms, roles and dynamics. Ideally gender specialists should facilitate and inform this step. Workshop participants in groups analyze and discuss outcome trajectories and the most significant changes along them from a social and gender equity perspective by answering the following questions:

- What vulnerable or marginalized groups are being, or could be, directly or indirectly affected by the change?
- Does the outcome trajectory:
  - Promote equal opportunities for vulnerable and marginalized groups? *Yes/How is that happening? Or No/Why not?*
  - Strengthen positive norms that support social and gender equality and an enabling environment? *Yes/How is that happening? Or No/Why not?*
  - Challenge norms that perpetuate social and gender inequalities. *Yes/How is that happening? Or No/Why not?*

Step 7: Identify immediate implications

The workshop produces learning and insight about which there is sufficient agreement to be acted upon immediately. To make sure this happens a workshop report identifying these measures is written and circulated to relevant people as soon as possible. Another strategy is to hold the Outcome Evidencing workshop immediately before annual planning so that the people involved in both can take the learning from one to the other.

Step 8: Plan and carry out substantiation; analyze the results

The workshop provides sufficient information to plan and carry out the substantiation of the outcome trajectories. Substantiation is carried out by an evaluator, who may be internal or external. Internal, or 'self-evaluation' has been found to be more self-critical and the results more useful to staff than when an external evaluator is used (Douthwaite et al., 2003) whereas external evaluation may carry more weight with an external audience when accountability is more important than learning. Developing and implementing the plan requires a number of decisions to be made as to which key informants to interview, which documentation to check and the evaluation report length and structure.

The substantiation verifies ways in which people are using program resources to generate outcomes. This is then compared to AAS' existing program theory and action taken if required.

Step 9: Analyze and use the findings

The evaluator who has carried out the substantiation and other staff leading the Outcome Evidencing process analyze the findings from the substantiation to complete the evaluation report. Outcome Evidencing was designed to be repeated annually or bi-annually within a program that needed the results to inform its adaptive management. Outcome Evidencing can also be used for one-off evaluations. In either case the authors of the evaluation report have a responsibility to promote the use of the findings, including comparing the findings to existing program theory and making adjustments as necessary.

Step 10: Repeat the Outcome Evidencing cycle

Repeating the Outcome Evidencing cycle annually allows AAS to explore how the outcome trajectories first identified have evolved and grown. This is done in subsequent repetitions of Step 3 by collectively deciding if new outcomes map onto existing outcome trajectories, and if they do whether they add to or challenge the outcome trajectory theory of change. New outcome trajectories may emerge in this process if new outcomes do not map onto existing trajectories. Repeating Outcome Evidencing allows the program to build an increasingly strong case for the changes to which it is contributing. New outcomes and causal explanation can serve to confirm or challenge initial causal claims (Barnett and Munslow, 2014) and program theory. This builds an increasingly sound basis for any future ex-post impact assessment.

Experience using Outcome Evidencing

We piloted Outcome Evidencing first in Bangladesh in 2014. The AAS Country Program Leader identified two main areas of change (Step 2) resulting from community and hub-level engagement respectively. For the first area, ccommunity facilitators in each of the sixteen AAS focal villages produced a list of outcome descriptions gleaned from documentation generated by village-level participatory monitoring and evaluation. The lists were then presented, revised and consolidated in a workshop where community facilitators reviewed, grouped and classified the outcomes. This workshop produced more than 50 outcome statements. Key members of the AAS country team then went through another round of review and consolidation to finally formulate 16 outcome descriptions.

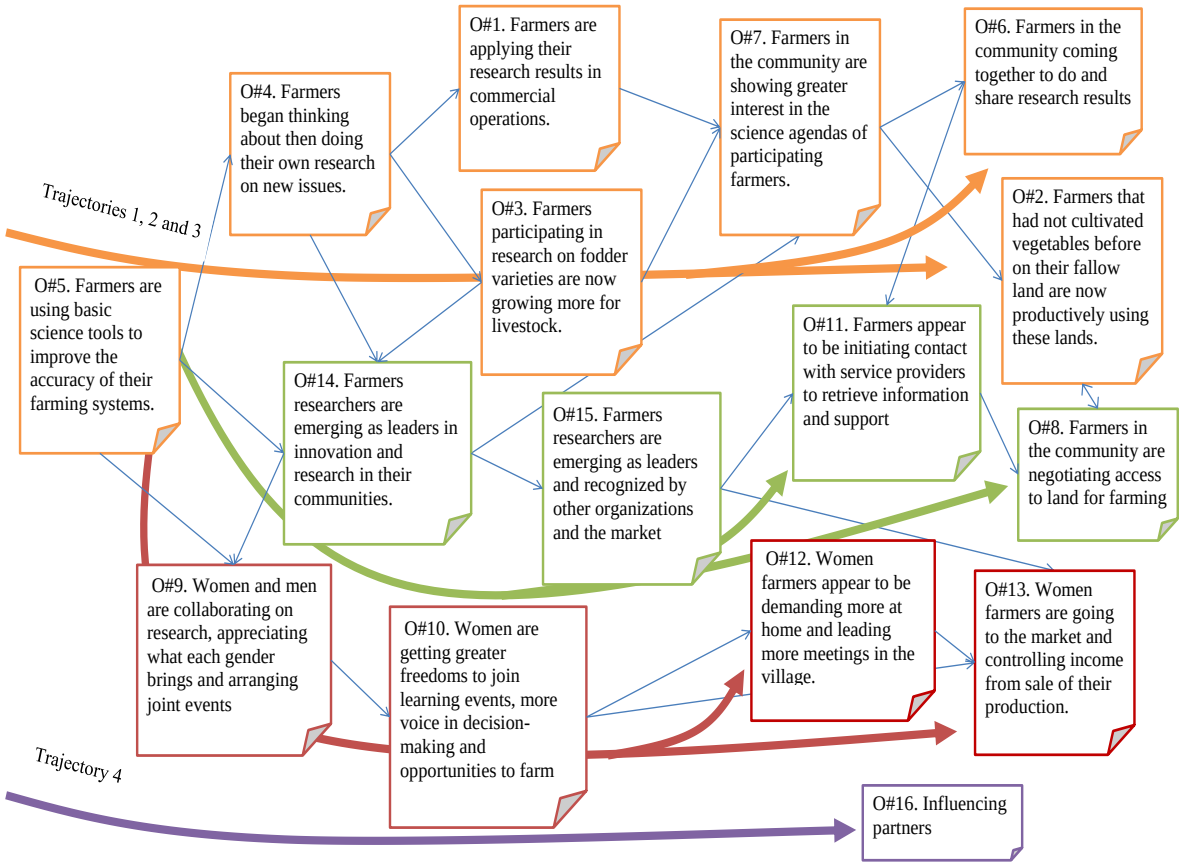
We brought the change agents together in a workshop in May to complete step 3 and carry out step 4 and 5. The change agents were the staff facilitating community engagement, AAS staff working in the hub and key people directly involved with AAS in Bangladesh.

The links participants identified among the 16 outcomes, and the four outcome trajectories they subsequently identified, are shown in Figure 2. The first three trajectories are the result of carrying out participatory action research (PAR) at community level.

1. F*armers doing research*, in particular through engaging with researchers and village facilitators. Outcomes associated with this pathway included changes in knowledge, skills, attitudes and practices in farmer researchers.
2. F*armers becoming self-confident* leading to outcomes such as farmer-researchers taking up leadership roles and becoming recognized by other organizations.

3. *Women and men working together* contributing to outcomes such as women with more say in decision-making and more freedom to join learning events and go the market.
4. I*nfluencing partners* as a result of AAS' engagement with hub-level partners contributing to outcomes such as partners developing greater ownership and understanding of AAS work including adopting elements of the RinD approach.

Figure 2: Outcome trajectories as identified during the workshop in Bangladesh

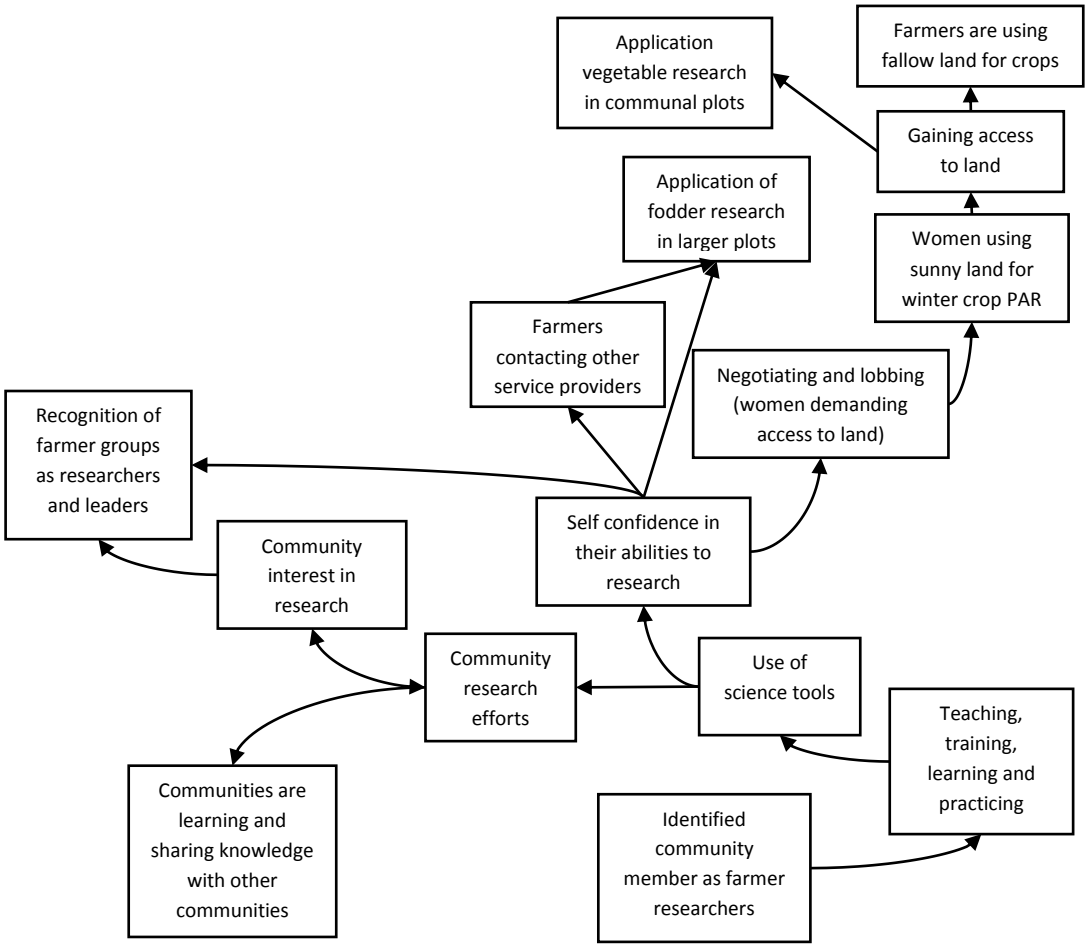


The next step of the workshop was to break participants into groups to develop a theory of change for each of the four identified outcome trajectories. To do this we asked the participants to:

1. Identify what they thought the program had contributed through implementing the RinD approach that had resulted in outcomes;
2. Specify the causal links between the outcomes; and,
3. Predict the likely future direction of the trajectory.

Participants built the respective theories of change by drawing and explaining a causal diagram. Participants in other groups offered challenge and validation during group presentations. Figure 3 shows the diagram produced by the group that worked on the *capacity to do research* outcome trajectory. For simplicity the diagram does not show feedback loops, of which there are several, for example outcomes resulting from increased self-confidence further building confidence, motivation and recognition.

Figure 3: A multi-cause diagram depicting the theory of change for the capacity to do research trajectory



The final part of the workshop was to plan the verification of the outcome trajectories as described in the causal diagrams, and assign responsibilities. We asked participants to identify and further describe the most significant outcomes along the trajectories in terms of actual people and organizations doing things differently, to identify existing documentary evidence of these changes, and key people to interview at community and hub-level for corroboration. Table 1 provides an extract of the output produced by one group.

Table 1: Identifying most significant outcomes and how to verify them for the "influencing partners" trajectory

1. What does the most significant outcomes look like?	2. What evidence do we have in hand?	3. What are the key people to interview?	4. What evidence needs to be collected?

Most significant outcome 1.2: Staff from partner organizations showed interest to participate in different AAS program events	Staff records from: - BRAC - SHUSILON - CIMMYT - Blue Gold - CREL - IWMI - Diversity - CARE	- Community people (Khulna) - Field level staff (Khulna) - National level staff (Dhaka)	Documentation (Relevant meetings, training workshops, etc.) Publications Dissemination (Key AAS outputs)

We did not include explicitly reflect on who is experiencing change, and who isn't (Step 6) in the first workshop. We included this step later on explicitly recognizing the need to critically reflect on 'who' was experiencing change and to be sure AAS was reaching the poor and marginalized.

We hired an external evaluator to carry out the substantiation step for the community engagement area of change. He worked with the AAS team to select significant outcomes per outcome trajectory and the villages where the outcomes were most likely to be present (Table 2). The evaluator visited the majority of the focal villages to build the case for the respective outcome in particular, and other outcomes and the overall trajectory of change in general. The final report included clear implications for the program and recommendations for future action. Box 2 provides an excerpt from the final report of the evaluation of the 'farmers doing research' trajectory of change.

Table 2: Key outcomes and case study villages selected to validate the 'capacity to do research' outcome trajectory at community-level in Bangladesh

Key features of the outcome trajectory	**Selected key outcome**	**Case study village**
Farmers doing research	Farmers are applying their research results in commercial operations	Borea
	Farmers in the community coming together to do and share research results	Habati
Greater farmer self-confidence	Farmers are using basic science tools to improve the accuracy of their farming systems	KDC
	Farmers appear to be initiating contact with service providers to retrieve information and support	Sahos
Women and men collaborating in research	Women and men are collaborating on research, appreciating what each gender brings and arranging joint events	Gangarampur
	Women are getting greater freedoms to join learning events, more voice in decision making and opportunities to farm	Ghonapara

Box 2: Excerpt from final Outcome Evidencing Report for Bangladesh

Trajectory 1: Capacity to do research

There is strong evidence for outcomes on this trajectory. The two case studies (Borea and Habati villages) point to numerous specific instances of farmers mastering components of the Community Life Competency Process - CLCP process [an approach introduced by AAS involving visioning, self-assessment, prioritizing, action planning] and applying their results to other crops and to farmers sharing their results formally and informally. The participatory monitoring and evaluation (PM&E) reports prepared in June, 2014, clearly indicate that farmer researchers have mastered the basics of a scientific approach to testing seed varieties, are applying what they learn to other crops, and are sharing their results within their farmer researcher groups, with neighbors through informal networks, with support agencies via the Research Technical Support teams, and more widely within their own and neighboring communities through highly successful farmer field days. Several examples taken from the PM&E reports are provided below and similar examples can be found for all 16 villages.

The following statements offered as examples were made by farmer researchers in Bengali to Program Officers, who translated them into English.

> ***Akra:*** *Now it is easy to arrange different events like farmer field day, learning session, exposure visit by our leadership.*
> ***Kazla:*** *Other-neighbor-farmers-(non-AAS) communicated with us (AAS farmer) about the research technology like line to line spacing, fertilizer dose, good quality seed, and we (AAS farmers) assisted them in this regards. A good networking developed among all of AAS communities through knowledge fair.*
> ***Tarali:*** *We have been communicated with RTS member, DAE office and other development organization for technical purpose related to PAR [in this context, PAR refers to the technical issues farmer groups are researching].*

Action plans in the PM&E reports for 2014 provide strong indications that farmer researchers will continue to progress along this trajectory by applying the results they obtain from their homestead trial plots to larger plots for commercial purposes and plan to scale out their research to address other topics. For example:

In Borea, the 2014 action plan calls for:

- Carp fish and prawn culture in pond. Linkage with Department of Fisheries and other WorldFish projects for technical support and training.

In KDC the 2014 action plan calls for more focus on fisheries:

- Fisheries. Pond aquaculture. Improved *gher* [dike] aquaculture. Training on pond preparation, stocking, and post stocking management.

Outcome Evidencing in other hubs

We followed a similar Outcome Evidencing process in the other four hubs, with some further adaptations according to local context and capacity. Table 3 summarizes the approach used in each hub to identify and verify outcomes and outcome trajectories.

Table 3: Outcomes identification and classification processes in hubs.

Hub	Method of identifying and clustering outcomes	Method of verification
Southern Bangladesh Polder Zone	– >50 outcomes identified at community level in a workshop – 16 outcome descriptions identified by AAS team. – 4 outcome trajectories identified in Outcome Evidencing workshop	External evaluator
Malaita – Solomon Islands	– 17 outcome descriptions identified through Most Significant Change at community level complemented by other outcomes identified by AAS team. – 5 outcome trajectories identified in Outcome Evidencing workshop	Internal evaluator
Barotse – Zambia	– 70 outcomes identified from learning reports produced by stakeholders and partner organizations. – 6 outcome trajectories identified in first Outcome Evidencing workshop	External evaluator
Tonle Sap - Cambodia	– 12 outcome domains identified from learning reports from focal communities and then revised and verified by AAS team – 3 outcome trajectories identified in Outcome Evidencing workshop	Internal and external evaluators
Visayas and Mindanao - Philippines	– 5 outcome domains identified and described by members of AAS team embedded in communities. – 80 outcomes identified in Outcome Evidencing workshop – 3 outcome trajectories identified in Outcome Evidencing workshop, including 14 sub-trajectories	Internal evaluator

All hubs carried out an Outcome Evidencing workshop to identify outcome trajectories, identify evidence and develop a plan to verify them. All hubs used a mixture of existing documentation and change-agent recall to identify outcomes. Different hubs used different approaches to processing these outcomes prior to the Outcome Evidencing workshop. Zambia and the Philippines clustered relatively large numbers of unprocessed outcomes in the Outcome Evidencing workshop while the other hubs carried out some form of amalgamation, usually by the AAS team, before the workshop. There was also a difference in whether hubs chose to use an external evaluator or use internal resources to verify the outcome trajectories (Table 3).

The Philippines was the last hub to complete their Outcome Evidencing and were able to learn from experience from the other hubs. Their process provides an interesting contrast to that followed in Bangladesh.

The Philippines AAS team spent much more time in their respective focal communities than other hub AAS teams, making them the primary change agents. In other hubs the primary change agents were more junior staff contracted to play the role, or staff of partner organizations. This first-hand experience meant the Philippine AAS team was able to identify five more defined areas of change:

1. Small-scale fisheries management in Barangay Mancilang, Madridejos, Cebu
2. Emerging Community Based Small Scale Fisheries Governance in Balingasag, Misamis Oriental

3. Mango production in Barangay Pinamgo, Bien Unido, Bohol
4. Rehabilitation of Abaca Production three communities in Sogod, Southern Leyte
5. Vegetable home gardening in Barangay Galas, Dipolog City.

With these areas in mind the team organized an Outcome Evidencing workshop to which they invited other respective change agents both from community and hub-levels. Twenty-eight people attended the workshop in which they identified 80 outcomes within and beyond the initial five areas of change. These were clustered according to actor groups involved which were: communities; partners; and, the AAS team. Two groups then worked with the clustered outcomes to build a causal diagram/theory of change for each of the first two actor groupings. This led to the identification of three main outcome trajectories:

1. AAS team is showing ability to influence/develop linkages and partnerships and work in/with communities
2. Partners are recognizing that RinD approach is markedly different from their approaches and starting to adopt aspects of it
3. Communities recognizing their strengths, resources and gaining better linkage with institutions to undertake actions to improve their lives

The causal diagrams also allowed for the identification of key outcomes for verification, existing documentary evidence and key informants to interview. Table 4 shows the key outcomes and sources of documentary evidence identified for the partner outcome trajectory.

Table 4: Outcomes and evidence identified for the partner outcome trajectory in the Philippines

Outcomes identified through drawing a causal diagram	**Specific outcomes selected for validation**	**Evidence that the outcomes have occurred and of AAS contribution to them**
– Stakeholders increasingly committed to tackling hub development challenge – There is emerging buy-in to RinD approach by partners and stakeholders – Partnership and network around AAS program are expanding – Partners using AAS outputs – Partners recognizing the importance of participatory	Endorsement of the AAS Program by the Regional Development Council (RDC)	– RDC endorsements in Regions 7, 8, 10 – Letter of from the Region 10 Director of the Department of Science and Technology (DOST) to the Under-secretary – Minutes of RDC meetings
	Different partners investing in activities that are oriented to tackle the hub development challenge	– Partners' investments in activities to tackle the hub development challenge – Memorandum of understanding, meeting report, and plan of work and budget of WorldFish-PCAARRD Technical Working Group – Memoranda of agreement with Department of Agriculture projects for work on capacity building for AAS, climate change and Tilapia – Memoranda of understanding with DOST Regions 8, 9 & 10
	Stakeholders appreciate	– Focal group discussions and pre-testing

approaches – Key staff of partner organizations are becoming more aware of social inclusion issues and are more conscious of engaging the poor and marginalized particularly in conducting research activities	the RinD process of identifying community needs and use outputs of this process in targeting beneficiaries	of planting material with abaca farmers in Sogod – VisMin Hub Stakeholders' Consultation Workshop (SCW) – SCW for the development of an integrated plan for Abaca rehabilitation in Sogod – Letter of DOST 10 RD Alfonso Alamban to DOST USec Carol Yorobe – Letter of DOST 8 RD Edgardo Esperancilla to Dr. Maripaz Perez of WorldFish
	Transformation of individual commitments to institutional commitments through continuous engagement by the AAS team	– Certificate of services rendered by local community facilitators (LCF) – LCF contracts – MOAs with SUCs and partners – Community immersion team (CIT) reports – Workshop proceedings
	Partners developing a shared vision and acting on a common plan of action thus bringing together fragmented network that AAS facilitated	– Proceedings: SCW for the development of an integrated plan for Abaca rehabilitation in Sogod – Documentation of the training on abaca production – Knowledge, sharing, and learning events

The final evaluation report was written by members of the AAS team. Box 3 presents an excerpt describing and verifying the first two outcomes in the partner trajectory (Table 3). The original report was extensively references with hyperlinks to documentary evidence held on internal site.

Box 3: Edited excerpts from final Outcome Evidencing Report for the Philippines for the partnership outcome trajectory

Endorsement of the AAS Program by the Regional Development Council (RDC)

The RDC is the highest policy-making body and serves as the regional counterpart of the National Economic and Development Authority (NEDA) chaired by the President of the Republic. RDC's primary responsibility is to coordinate and set the direction of all economic and social development efforts in the region. It also serves as a forum where local efforts can be related and integrated with national development activities.

The AAS Program has been endorsed by the RDCs of Region 7, 8, and 10. This was facilitated by our partners who are members of the RDC. Without our partners having sponsored the presentation of AAS in the RDCs' sectoral committees which, in some occasions they head, our entry into the RDCs could have been difficult. The principles we shared with the Regional Offices of DOST facilitated our access into RDCs. In some instances, Department of Science and Technology (DOST) Regional Directors defended the program in full RDC sessions. Table shows the status of endorsement.

Status of endorsement of AAS in the RDC

Region	Status	Resolution	Sponsor
Region 7	AAS endorsed by RDC 7 Economic Development Committee (EDC)	RDC Resolution 1 (s. 2014), "Endorsing to Potential Partner Agencies and Convergence Groups in Central Visayas the Consultative Group of International Agricultural Research (CGIAR)-Research Program on Aquatic Agricultural Systems for Replication in Other Areas in the Region"	Regional Director of DOST 7
Region 8	AAS presented to full council of RDC 8	RDC VIII Resolution No. 21 (s. 2014), "Endorsing the Aquatic Agricultural Research Program to the National Government Agencies and Local Government Units"	Regional Director of DOST 8 and Chair of the RDC 8-Social Development Committee
Region 10	AAS endorsed by RDC 10-EDC	RDC X Resolution No. 33 (s. 2014), "Endorsing the Consultative Group on International Agricultural Research Program on Aquatic Agricultural Systems"	Regional Director of DOST 10

Regional development planning is necessary to address the uneven economic and socio development of the country, and these endorsements open the gates for AAS to engage as active participant in national development.

Different partners investing in activities that are oriented to tackle the hub development challenge

The hub development challenge (HDC) and a strategic framework to tackle it was agreed with stakeholders through a series of regional consultation workshops in 2012 culminating in the stakeholders' consultation workshop (SCW) and Design Workshop in 2013. The collective development of both allowed stakeholder to explore collaboration including the support of the endeavors tackling the HDC. At least USD 390,000 has been invested (both in cash and in kind) by at least nine partners since 2013

The AAS team in the Philippines reflected on the results and came up with important learning and affirmation. For example, from the partnership trajectory they concluded that it is possible within a relatively short period to facilitate research and development organizations to work towards a common goal through a number of initiatives. They realized that what it takes are communities that can organize and express their development requirements and an 'honest broker' able to link communities' visions and dreams and organizational mandates. They concluded that research organizations can play this role because of the neutral space that research provides for people to work together. On the other hand, the Outcome Evidencing exercise helped them realize the resources required carrying out the 'honest broker' role takes resources away from research and a challenge the team faces is getting the balance right.

Like any evaluation method, Outcome Evidencing runs the risk of confirmation bias. This was a particular concern given we were aware that investment being made in AAS was contingent on demonstrating the RinD approach was working. However, part of the RinD approach is that it has in place a monitoring, evaluation and learning system that allows it to learn from what is and **is not** working so as to adjust implementation accordingly. Outcome Evidencing was able to pick up on negative outcomes. For example, a program outcome identified in Zambia was the reduction in the illegal use of mosquito nets for fishing. However, on further reflection in the workshop it emerged there was now an increase in the use of pesticide to poison the fish, a practice that was harder to detect. This led to the realization that better education about the damage done by illegal fishing methods was not working without illegal fishers having some other way of providing for their families.

The other guard against confirmation bias is building staff and key stakeholder capacity to reflect critically as a core program value.

Reflection on novelty of Outcome Evidencing

As already said, Outcome Evidencing is a hybrid of outcome harvesting and the modus operandi methods. Wilson-Grau and Britt (2012) describes Outcome Evidencing as an evaluation approach that starts with emerging outcomes and working back to establish if and how program interventions had contributed by reconstructing and validating causal pathways. The steps in the outcome harvesting method are summarized in Box 1.

Box 1: Outcome harvesting in brief (Adapted from Wilson-Grau and Britt, 2012:4).

1. **Design the Outcome Harvest**: Agree evaluation questions to guide the harvest on what information is to be collected and included in the outcome description.
2. Gather data and draft outcome descriptions: Harvesters glean information about changes that have occurred and how the change agent contributed to these changes. Information about outcomes may be found in documents or collected through interviews, surveys, and other sources. The harvesters write preliminary outcome descriptions.
3. Engage change agents in formulating outcome descriptions: Harvesters engage directly with change agents to review and classify the draft outcome descriptions and identify and formulate new ones.
4. Substantiate: Harvesters obtain the views of independent individuals knowledgeable about the outcomes and how they were achieved; this validates and enhances the credibility of the findings.
5. Analyze and interpret: Harvesters organize outcome descriptions through a database in order to make sense of them, analyze and interpret the data, and provide evidence-based answers to the evaluation questions.
6. Support use of findings: Drawing on the evidence-based, actionable answers to the evaluation questions, harvesters propose points for discussion to harvest users, including how the users might make use of findings. The harvesters also wrap up their contribution by accompanying or facilitating the discussion among harvest users.

The main difference between Outcome Evidencing and outcome harvesting is the focus on identifying and evidencing outcome trajectories, rather than outcomes *per se*. This focus on

patterns of outcomes borrows from the Modus Operandi approach (Scriven, 1976). Outcome Evidencing combines steps 3 and 5 of outcome harvesting – change agent description of the outcomes and their analysis and interpretation – in a workshop involving participatory identification of outcome trajectories. The substantiation step, step 4, had become verification of the theories of change developed to describe the outcome trajectories. Outcome Evidencing does not use "independent but knowledgeable people" to validate outcome claims because when evaluating emerging outcomes, AAS' experience is that the people knowledgeable about them were also likely to be involved with the Program in one way or another, and therefore not independent. Instead Outcome Evidencing uses evaluators for this step.

Outcome Evidencing's claim to novelty is the adaptation of outcome harvesting to include elements of the Modus Operandi approach for the purpose of prospecting for and making sense of emerging outcomes, both expected and unexpected, within a project or program lifespan. Unlike Outcome Harvesting it includes a specific step to look at inclusion and winners and losers.

In this paper we have attempted to give a sense of the practicalities of developing and using a complexity-aware evaluation method in the field. The stepwise method we describe at the beginning is an ideal type constructed from learning from five pilots in five hubs. Outcome Evidencing is still in its formative phase and will no doubt adapt and improve as it is used more. Whether it emerges as a new method in its own right or be seen as an adaptation of Outcome Harvesting remains to be seen. Either way, our hope is that it proves useful.

Conclusions

This paper describes the development of the Outcome Evidencing method to help the AAS program meet learning and accountability requirements as it intervenes in geographic hubs, understood as complex systems. The approach identifies emerging outcomes, both expected and unexpected, happening within program areas of change. It then seeks to understand, describe and verify these outcomes to support learning. The method is centered on a workshop that makes sense of those outcomes in terms of identifying immediate implications. The workshop also identifies outcome trajectories for subsequent substantiation. Comparing substantiated outcome trajectories with program log frames, or equivalent, allows the program to question its underlying causal premises. The method can be used for one-off evaluations that seek to unpack the black box answer evaluation questions relating to what aspects of program intervention worked, for whom, to what extent and why. However, it is likely to be most useful as a central part of program M&E. Repeated cycles of Outcome Evidencing build a case for program contribution over time that can be evaluated as part of any future impact assessment of the program or parts of it. Outcome Evidencing is an adaptation of the outcome harvesting method to include elements of the Modus Operandi Method. The main difference to Outcome Evidencing is it seeks to substantiate program contribution within theories of change rather than program contribution to discrete outcomes.

References

AAS (2011) CGIAR Research Program Aquatic Agricultural Systems: Program Proposal. AAS-2012-07.

AAS (2014) CGIAR Research Program on Aquatic Agricultural Systems: Extension Proposal 2015-2016. Accessed July 2015 from: http://goo.gl/Q4XwIU

Astbury, B., & Leeuw, F. L. (2010). Unpacking black boxes: mechanisms and theory building in evaluation. American Journal of Evaluation, 31(3), 363-381.

Barnett, C. and Munslow, T. (2014) Process Tracing: The Potential and Pitfalls for Impact Evaluation in International Development, Evidence Report 102, Brighton: IDS

Douthwaite, B., Kuby, T., van de Fliert, E., & Schulz, S. (2003). Impact pathway evaluation: an approach for achieving and attributing impact in complex systems. *Agricultural Systems*, *78*(2), 243-265.

Elster J (2007) Explaining social behavior: More nuts and bolts for the social sciences. Cambridge University Press.

Halliday, A. and Glaser, M., (2011). A management perspective on social ecological systems: a generic system model and its application to a case study from Peru. *Human Ecology Review*, *18*(1), pp.1-18.

Kemp, R., Schot, J. and Hoogma, R. (1998). Regime shifts to sustainability through processes of niche formation: the approach of strategic niche management. Technology analysis & strategic management, 10(2), pp.175-198.

Klerkx, L., Van Mierlo, B., & Leeuwis, C. (2012). Evolution of systems approaches to agricultural innovation: concepts, analysis and interventions. In *Farming Systems Research into the 21st century: The new dynamic* (pp. 457-483). Springer, Netherlands.

Mayne, J. and E. Stern. (2013). Impact evaluation of natural resource management research programs: a broader view. ACIAR Impact Assessment Series Report No, 84. Australian Center for International Agricultural Research: Canberra. 79p. http://aciar.gov.au/files/ias84.pdf

Patton M. Q. (2011) *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. Guilford Press.

Pawson, R. (2013). The Science of Evaluation. SAGE Publications. Kindle Edition.

Popper, K. (1992). The logic of scientific discovery. Routledge

Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.

Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, *14*(1), 29-48.

Schot, J., & Geels, F. W. (2008). Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy. *Technology Analysis & Strategic Management*, *20*(5), 537-554.

Scriven, M. (1976). Maximizing the power of causal investigations: The modus operandi method. Evaluation studies review annual, 1, 101-118.

Stame N (2004) Theory-based evaluation and types of complexity. *Evaluation 10*(1): 58-76.

Stern, E., (2015). Impact evaluation: a guide for commissioners and managers. Bond. Accessed in October 2015 from https://www.bond.org.uk/data/files/Impact_Evaluation_Guide_0515.pdf

Snowden, D. (2010). Naturalizing sensemaking. Informed by knowledge: Expert performance in complex situations, 223-234.

van Mierlo, B., Arkesteijn, M., & Leeuwis, C. (2010). Enhancing the reflexivity of system innovation projects with system analyses. *American Journal of Evaluation*, *31*(2), 143-161.

Westhorp, G. (2014). Realist Evaluation: An introduction. Methods Lab-ODI.

Weiss, C. H. (1997) How can theory-based evaluation make greater headway? *Evaluation review 21*(4):501-524.

Wilson-Grau, R., & Britt, H. (2012). Outcome harvesting. Ford Foundation. Accessed, 6, 2012.