Social and Technological Transformation of Farming Systems:

Diverging and Converging Pathways

Proceedings of the 12th European IFSA Symposium
12th - 15th July 2016 at Harper Adams University, United Kingdom

Volume 1

Andrew Wilcox and Karen Mills (Eds.)

Harper Adams University

IFSA INTERNATIONAL FARMING SYSTEMS ASSOCIATION EUROPE GROUP

Workshop 1.2: Monitoring and evaluation for learning and innovation

Convenors: Boru Douthwaite, Ann Waters-Bayer, Bernard Triomphe, Barbara van Mierlo, Cees Leeuwis and Rodrigo Paz.

Rural innovation systems are complex adaptive systems. In such systems the way that interventions, intermediate results and impact are evaluated can positively or negatively affect the capacity of that system to innovate. For example, linear and often exaggerated expectations of impact can lead to over investment in baselines that take energy and funding away from creating, supporting and scaling the innovation processes and capacities required to have impact in the first place. A key ability for researchers and development practitioners wishing to intervene effectively is to do so based on understanding how innovation and change come about in complex systems, not on the basis of linear and usually simplistic logframes and theories of changes that can come with conventional M&E. Evaluation can build this ability by focusing on understanding how project and programme interventions actually work, or not, and - by doing this "in real time" - can, at least in theory, influence on-going implementation by allowing the project or programme to 'learn its way forward'.

The workshop collated and collectively explored a body of empirical evidence and theory as to where and how monitoring and evaluation (M&E) works to support learning, adaptive management and foster innovation. To this end, the workshop called for papers on the following:

- Cases of M&E that have attempted to support learning, adaptive management and/or foster innovation, successfully or unsuccessfully
- The conceptual and theoretical design of M&E approaches that try explicitly to foster learning and innovation
- The use of M&E methodologies to support learning about how innovation and change happens, including the use of impact, pathway, theory of change, realist evaluation, contribution analysis and developmental evaluation
- Understanding and measuring whether and how different types of interventions build capacity to innovate among innovation system actors in a given site and the subsequent manifestations of that capacity.

This workshop explored both theoretical and practical perspectives. The outputs included a set of propositions about how to build and measure system capacity to learn and innovate, and resulting impact, reflecting on experiences and lessons identified during the workshop. The main results will form the basis of a written synthesis and (hopefully) a multi-author journal article. The workshop process consisted of short presentations of selected papers submitted beforehand, followed by group work and facilitated plenary discussion. Interested paper authors were invited to meet to develop an annotated table of contents for a journal article before the end of the Conference.

Applying the Participatory Impact Pathway Analysis (PIPA) approach to enhance co-innovation for sustainability within livestock family farming in Uruguay

Albicette, M.M., Leoni, C., Ruggia, A., Bortagaray, I., Scarlato, M., Scarlato, S., Blumetto, O., Albín, A. and Aguerre, A.

National Agricultural Research Institute (INIA), Uruguay

Abstract: Participatory approaches are needed to ensure learning processes and to incorporate lessons learned during the implementation of a project. This is particularly important when the aim is to improve farm sustainability considering changes in knowledge and skills, natural resources management and networking. This paper describes the Participatory Impact Pathways Analysis (PIPA) implemented within the participatory action research project 'Co-innovating for the sustainable development of livestock family-farming systems in Rocha, Uruguay', which involved stakeholders for planning, monitoring and evaluating of the project's progress. Six workshops were implemented during 2012 - 2015 to enhance the project's actions. Participatory methods were used to adapt PIPA to the Uruguayan culture. During 2013 an interinstitutional network was established, a shared vision of expected project results was defined, as well as impact pathways, goals and activities to achieve them. During the 2014-2015 workshops, reflections and suggestions led in turn to new or modified activities. This process contributed to confidence and commitment building, improving the quality of the established relationships and strengthening networking to enhance the dissemination of the project findings. As a result of the learning process, and inspired by the project's methodological and technological results, one stakeholder organisation established a project for another region. The last workshop focused on a participatory evaluation of the whole project, demonstrating that a successful innovation process took place. This Uruguayan case showed that within the co-innovation framework, the PIPA approach nurtured the creation of a common space for social learning and innovation, providing a useful instrument for rural development.

Keywords: Learning Process, methods, networks, planning, monitoring and evaluation

Introduction

The traditional linear transfer of technology model is a top-down process from research institutions to farmers. This model is still predominant worldwide and has often led to low use of many improved agricultural technologies (Moschitz et al*.,* 2015; Okali et al., 1994). To overcome this, new theoretical perspectives had emerged where the development of network and system approaches and the inclusion of relevant actors to broaden agricultural innovation were incorporated (Klerkx et al., 2012). These new perspectives are most needed when dealing with natural resource management systems to improve farm sustainability (Speelman et al., 2007), where a variety of stakeholders are involved (Schut et al., 2015). However, most of the institutions in charge of fostering innovation are still locked into old approaches and methods of intervention (Moschitz et al., 2015).

Participatory collaboration in knowledge generation, technology development and innovation has proved its ability to add value to science-based technology development (IAASTD, 2009). Working with a network of researchers, extension agents, farmers and local actors, focused on bringing new products or new processes into economic use as well as sharing and exchanging knowledge among them, strengthens innovation (Klerkx et al., 2009). To promote changes in agricultural practices towards more sustainable production systems, a collective learning process among all stakeholders is needed (Dogliotti et al., 2014). Stakeholders are actors interested in addressing a problem and their participation is seen as a critical success factor to solve complex agricultural problems (Schut et al., 2015). Social learning projects should include a reflexive design to encourage and facilitate the learning processes, particularly when outcomes are expected to contribute to sustainable development (Loeber et al., 2007).

Project monitoring and evaluation (M&E) can be used to enhance learning during its implementation and not only for accountability issues (Douthwaite et al., 2003; Rossing et al., 2010). M&E is increasingly seen as crucial to the success of rural research and development projects because it supports a real-time feedback (Douthwaite et al., 2007a). Furthermore, stakeholders should periodically reflect on the validity of the impact hypotheses and the entire process should be facilitated (Moschitz et al., 2015) and documented so as to better understand the mechanisms through which socio-technical changes are fostered

Participatory Impact Pathway Analysis (PIPA) draws from program theory evaluation (Rockwell and Bennett, 2004), social network analysis and ongoing research for development to understand and foster innovation and is related to designing strategies, as well as a basis to set out a monitoring and evaluation framework (Alvarez et al., 2010). The PIPA method was successfully used by Douthwaite et al. (2007b) to enhance the developmental impact of projects through better impact assessment, to provide a M&E framework, to allow stakeholders to learn for future initiatives and to provide information that can be used for public policies.

Between 2000 and 2011 in Uruguay a 21% reduction occurred in the number of farms (most of which were family farms). At the livestock farming level, there are more than 26,000 farms in 11.7 million hectares, 60% of which are family farms (Tommasino et al., 2014). There are opportunities for improving family farms sustainability by re-designing those systems through an adequate selection and orientation of production activities and the use of appropriate technologies and farming management skills, through a participatory intervention to promote learning and innovation (Albicette et al., 2016).

Between 2012 and 2015 a group of researchers at INIA (Spanish acronym for National Agricultural Research Institute) implemented the project ‘Co-Innovating for the sustainable development of family-farming systems in Rocha-Uruguay’. The project presupposed an innovation paradigm shift through participatory research, aiming to contribute to the improvement of livestock family farms’ (LFF) sustainability and rural development. Three interconnected and simultaneous participatory processes took place: at farm level; at research team level and at Rocha regional level, with specific methods for each one (Albicette et al.,

2016). In this article we focus on the regional level where the PIPA method was adapted to plan, monitor and evaluate the co-innovation process throughout the three years of the project, engaging regional stakeholders in a participatory learning process. We describe the method used, the M&E activities, the results obtained and the lessons learned.

Methodology

The co-innovation approach is considered as a participative and interactive approach to foster effective innovation across stakeholders (Coutts et al., 2014), combining farming systems theory, social learning and dynamic project monitoring and evaluation (Dogliotti et al., 2012; 2014; Rossing et al., 2010). In this project the approach was implemented between 2012-2015, considering three interconnected simultaneous processes: (i) at farm level, seven representative LFF based on native grasslands (project farms) (Albicette et al., 2016) were selected as case studies to asses sustainability using the MESMIS framework (Spanish acronym for Evaluation of Natural Resource Management Systems Incorporating Sustainability Indicators (Masera et al., 2000)); (ii) at research team level, a participatory action research (PAR) methodology was used to implement the project; and (iii) at regional level, the PIPA method (Alvarez et al., 2010) was adapted to involve local actors to monitor and evaluate the project.

PIPA participatory methods, techniques and tools

Specifically at regional level, the PIPA method was adapted and implemented to plan activities, to M&E the project, and to include the lessons learned during the process in real time. The PIPA method was implemented through workshops held at the local offices of farmers' organisations. Thus, six half-day PIPA workshops (PW1 to PW6) took place between July 2012 and August 2015 with two major objectives: (i) to share and discuss results at the farm level as a monitoring and evaluation process, promoting a learning process among participants; and (ii) to jointly develop activities to share the results.

In order to promote a constructive atmosphere to monitoring and evaluating the project advances and generate a learning process during the six project workshops (PW), six key points were considered: (i) each PW was carefully planned using a script with roles and responsibilities, specifying timetable and methodological tools, and the expected outputs were written down and distributed among the research team (Schut et al., 2015); (ii) the agenda for each PW was written on a flip chart to share it with the participants; (iii) the date for each PW was coordinated among project farmers, research team and other stakeholders, who were invited by e-mail (with the agenda and the minutes of the previous workshop attached) to be used as a kick-off point for the PW; (iv) moderation cards and visualisation charts (Schut et al., 2015) using different participatory techniques selected from a toolkit (Knowledge Sharing Toolkit, 2009) were applied during the PW, leading to a collaborative knowledge and reflection process; (v) a facilitator (member of the research team) oriented each PW introducing the methodology, guiding plenary sessions, monitoring group sessions and facilitating workshop sessions (Home and Rump, 2015); (vi) all materials and results were photographed and presentations were recorded in order to document the information (Akpo et al., 2015).

PIPA steps

Originally the PIPA process starts with the definition of a problem tree to understand the problems that the project addresses and what needs to change (Alvarez et al., 2010). In our case, a Rapid Rural Appraisal of the region (Capra et al., 2009) was used, with which the main constraints of LFF systems were identified. We therefore started the process by inviting farmers, researchers and local actors to build a regional interinstitutional network (IN) (Table 1) to plan and M&E the project 'Co-Innovating for the sustainable development of family-farming systems in Rocha-Uruguay'.

Table 1. Stakeholders of the Interinstitutional Network (IN)

Stakeholder groups[1]	IN Stakeholder
Research and Training	INIA - Research Team (Research Institute)
	Facultad de Agronomía – FAGRO (University - Research and Education)
	Centro Universitario Regional Este - CURE ((University - Research and Education)
	Instituto Plan Agropecuario -IPA (Extension and Training Institute)
Government	Intendencia Municipal de Rocha - IMR (Local Government)
	Sistema Nacional de Áreas Protegidas - SNAP (Environmental Ministry)
	Ministerio de Ganadería Agricultura y Pesca - MGAP (Livestock Ministry)
Non-governmental organisations (NGO) and civil society organisations	Comisión Nacional de Fomento Rural - CNFR (National Farmers' Union)
	Delegates from SFR 109 and SFR-C (Local Farmers' Organisation)
Farmers	Project farmers
	Farmers of the region

[1]*Based on Schut et al. (2015)*

Based on the project objectives, in PW1 the participants expressed what their goals concerning the project were and what they expected at the end of it (the shared vision). The starting question to elaborate that vision was: "What will be happening in 2015 with farmers, professionals, organisations and their relationships after the successful ending of the project?". Based on the vision, the IN generated the outcome model considering the following questions: "What changes do we intend to undergo?", "Which are the actors expected to change?", "What is needed to achieve the expected changes?", "Through which activities?" and "Who will do/implement them?" (Table 2).

As a way to implement the activities on the outcome model, the IN suggested the development of a Communication Plan (CP). For this, a committee of 4 IN stakeholders designed a strategy which was presented as a draft at the PW2, in order to discuss and formalise a CP. Finally the CP was defined and the activities were planned annually at PW3 and PW5, considering three target groups: farmers, professionals and organisations related to rural development.

During PW2 to PW6, M&E was done by IN stakeholders through the discussion of partial results presented by the research team and the analysis and reflexion of the outcome model and CP. These cycles of M&E led to a continuous process of knowledge acquisition, where strengths and weaknesses, suggestions to improve the project implementation to achieve its goals and IN vision were identified. Thus, several learning cycles occurred.

A final participatory evaluation of the project methodology and results was carried out with a survey of 17 questions and an open space for comments, answered by 18 IN stakeholders (INIA researchers were not included). The survey (see Appendix) was designed following Bennett´s hierarchy criteria (Bennett, 1975; Rockwell & Bennett, 2004). The evaluation was processed (i) during the PW6 where participants copied the answers onto a pin board where the questions were written, and the results were discussed and analysed in a plenary session (Figure 1) and (ii) after PW6 where it was processed in order to collect and make sense of quantitative and qualitative results. Quantitative results were determined calculating the average score for each response, which had been previously scored on a 5 points scale (++ = 5, + =4, 0 =3, - =2, -- =1).

Figure 1. Plenary discussion during final evaluation

Results and Discussion

Significant changes were obtained at the three levels where the project was implemented to enhance co-innovation in order to improve LFF sustainability. At farm level, the farming systems were re-designed by adjusting the stocking rate and sheep:cattle ratio, allocating

pasture according to biomass height and using low cost breeding practices. These changes in turn resulted in a 23% meat production increase and a 56% increase in net income, while maintaining natural resources untouched. Furthermore a 25% reduction was observed in the estimated workload on animals and pasture management, a 97% implementation of 11 of the proposed technologies and the incorporation of mid-term planning. The mentioned changes revealed changes in farmers´ knowledge and skills related to their LFF system (Albicette et al., 2016).

At the research level, by applying PAR methodology it was possible to consolidate a 'research team', with a mutual understanding of how to address the problem and the methodological approach to adopt to face it. From a group of researchers with a varied range of backgrounds and expertise - agronomy, environmental and social sciences - (Albicette at al., 2016), transdisciplinarity emerged as a new property where disciplinary scientific knowledge (scientific evidence) and knowledge from other sources (field experiences) were combined (Moschitz et al., 2015; Wiesmann et al., 2008).

At regional level PIPA method was adapted for planning and reviewing the project's progress towards its objectives, becoming more impact oriented (Alvarez et al., 2010). The key results of this level are described below.

A network perspective

Six PWs (PW1 - PW6) were organised during the project´s implementation with an average number of 32 participants (from 29 to 39). All IN stakeholders that were invited to PW1 in 2012 participated throughout the six PWs. This outstanding level of participation demonstrated that stakeholders were supporting what was taking place during the PWs and were highly involved in the process (Home & Rump, 2015).

During the PW1 IN stakeholders developed a shared vision of the project. In their own words: (i) "*There is a considerable improvement of farms sustainability, using suitable technologies that resulted in higher income, preservation of natural resources and life quality improvement";* (ii) *"farmers adopt an interactive working style";* (iii)*" farmers and professionals acquire knowledge and develop skills for specific techniques and resource management";* (iv) *"regional organisations are involved in the improvement of LFF, working as a network"* and (v) *"appropriate knowledge is being shared through presentations and field days and through mass media, making an efficient use of communication tools".* This vision was a clear expression of the stakeholders' dreams about the project's impact and became a strong motivating force for them to design a clear strategy and activities to be implemented (Douthwaite et al. 2007b).

Based on the IN vision, stakeholders defined impact pathways using an outcome model (Alvarez et al., 2010), describing what is expected from the project, the ways in which stakeholders can adjust their behaviours and the interactions needed to achieve their project vision (Table 2). The outcome model was agreed during the PW2 and was used as a basis for the project M&E (Alvarez et al., 2010). This strategy presupposes a paradigm shift during the

research process (design and methodologies) in order to achieve development impacts, where end-users are proactive actors in socio-technical changes (Akpo et al. 2015).

Table 2. Outcome Model elaborated by the Interinstitutional Network

What changes do we intend to undergo?	Which are the actors expected to change?	What is needed to achieve the expected changes?	Through which activities?	Who will do/implement them?
Sustainability of the PF[1] enhanced	PF and RT[2]	Interaction between field agronomist, RT and PF. Discussions and agreements to generate learning and change.	On farm Work	Mainly PF and RT. Strategically, IN[4] professional
		Commitment of RT and PF		
Farmers in the region are aware of technologies promoted in the project	PF	PF interacting with their groups	Strategic group meetings	PF, professionals and IN organisations
	Organisations' farm members	Members of the organisations interacting with PF	Yearly meetings for presentation and exchange of ideas	PT [3]
	Farmers in the region	Farmers interacting with PF	Various activities: face to face, mass media, web.	PT
	Professionals working in organisations linked to the project	Professionals interacting with PF	Strategic visits to PF	PT
		Interaction of professional teams in the region	Professional teams meetings	
	Professionals working on other organisations	Inform and raise awareness about the results and ways of working in the region	Various activities	
	IN representatives	Encourage networking	Lead by example	PT
		Channelling issues to corresponding organisations	Acts as an emissary of the new ideas	Delegate
		Encourage organisations to adequate approach for working with LFF	Workshops with organisations/ policy makers for awareness	PT

[1]PF: Project Farmers, [2]RT: Research Team, [3] PT: Project Team = project farmers, research team and professionals of the IN, [4] IN: Interinstitutional Network

High stakeholder involvement was achieved as they were asked to monitor the process. Constraints and interests of different stakeholder groups were considered, allowing the triangulation and validation of the products generated by the IN (Schut et al., 2015). Some

conflicts emerged in relation to high stakeholder expectations of the project considering the resources available to implement it. Negotiating was therefore necessary to balance the demands into a mutually acceptable solution (Leeuwis, 2000). As stakeholders' suggestions were incorporated, the original high motivation level remained. Farmers and other stakeholders were seen as relevant actors in the process (Leeuwis & Van der Ban, 2004), rather than perceived as technology consumers (Moschitz et al., 2015).

Two annual communication plans (CP) (2013 and 2014-2015) were elaborated by the IN stakeholders focused on activities to share project methodologies and results related to the LFF. These aimed to promote learning among different target groups i.e. farmers, professionals and organisations. On-farm meetings and local activities were included to enhance the interactive learning process among farmers and professionals in the region. Dissemination activities were important to respond to local farmers´ demands. Field days were key activities for sharing results and interacting with a broader audience. Promotion of the projects progress using mass media was important to reach politicians and people from outside the region. Project strategy and results were also presented at national and international academic events and at activities related to LFF policy makers (Table 3).

Table 3. Summary of the communication plans elaborated by the Interinstitutional Network (IN)

Activities	Objective	Who were invited	Channels used for invitation	N° of Activities and Total Participants
On-farm meetings	Share and discuss implementation and results of the project at farm level	PF[1] groups, their agronomist and neighbours	Personal invitation SMS[2]	9 meetings 100 participants
Local activities	Exchange information on specific technological topics (e.g. cattle body condition scoring, pastures management, ovarian diagnosis activity)	IN[3] stakeholders	SMS	3 on-farm activities 90 participants
Dissemination activities demanded by local farmers organisations	Exchange information on technological topics (bull and cow management for mating, heard management under drought conditions)	Farmers in the region	SMS	2 meetings 70 participants

PW1[4] to PW6	Planning, M&E	IN stakeholders	Personal e-mails and SMS	6 PW 180 participants
Seminars for professionals related to rural development	Technical discussion (Social and Environmental issues)	Professionals	E-mail	2 Seminars 90 participants
Field days at PFs	Share project results and processes	Open invitation	Personal Invitations, Newsletter INIA, Twitter, SMS, advertising in mass media, Web, flyers	5 Field days 600 participants
Participation in mass media	Disseminate of project activities and results	PT[5]		9 articles in rural magazines. Radio and TV interviews.
Participation in national and international academic events	Disseminate of project methodologies and results	RT[6]		20 activities with 1900 participants, 13 conference papers
Participation in interinstitutional meetings	Discuss co-innovation related to LFF policy makers	PT		5 activities 110 participants

[1]PF: Project Farmers, [2] SMS: Short Message Service, [3] IN: Interinstitutional Network, [4] PW1: Workshop 1 to PW6: Workshop 6, [5] PT: Project Team = project farmers, research team and professionals of the IN, [6]RT: Research Team

On December 8th 2015, 200 people attended a field day where project final results were presented. The evaluation sheet for the activity was responded to by 98 participants: 93% considered the activity excellent or very good; 83% indicated that the proposed technologies presented during the field day were feasible to implement in their own farms. Furthermore, at the end of that day, seven national authorities stressed the importance of the project approach and results in relation to: (i) fostering farm sustainability through an intensification process and adaptation to climate change, while maintaining farming families on their land; (ii) enhancing farmers´ knowledge and skills; (iii) promoting regional networking; (iv) generating scientific data to support family farm policies. Finally, the process was highlighted as a methodological innovation for INIA, as "*a way of working which thinks on what (...) and how things are done*". A summary (in Spanish) of the field days of 2014[1] and 2015[2] is available at INIA's website.

[1] 2014 Field day: http://www.inia.uy/estaciones-experimentales/direcciones-regionales/inia-treinta-y-tres/jornada-de-producci%C3%B3n-familiar

[2] 2015 Field day: http://www.inia.uy/estaciones-experimentales/direcciones-regionales/inia-treinta-y-tres/hacia-una-ganader%C3%ADa-familiar-sustentable-jornada-final-del-proyecto-co-innovando-en-rocha-2012-%E2%80%93-2015

Monitoring and evaluating the project´s progress as a learning process

Throughout PW2 to PW6 research team members and farmers presented and shared the project activities and farm results, so that anyone could follow and monitor the project advances. Special attention was paid to reporting all activities; a key factor in the co-production of knowledge in a multi-stakeholder processes (Akpo et al. 2015). The participatory M&E process enhanced stakeholders' learning through a regular reflection on the project progress and results, using a different perspective of impact assessment (Douthwaite et al., 2003; Rossing et al., 2010). M&E is traditionally used for accounting project achievements whereas we used it for analysing the process and emphasising the importance of real time feedback, thus promoting learning (Douthwaite et al., 2003; 2007a).

Due to active M&E stakeholders suggested improvements to the outcome model (Table 2) and to the CP (Table 3). This process contributed to confidence and commitment building, improving the quality of the relationships and strengthening networking. As posited by Schut et al. (2015) different stakeholders enhanced insights into the different dimensions of a problem and could look for different types of solutions. As an example of the depth of the M&E process, some reflections of the 30 participants during the mid-term workshop (PW3, September 2013) are presented in Table 4. Participants analysed project achievements and difficulties as well as elaborated suggestions for enhanced project implementation. Most project achievements were related to the co-innovation approach used. At farm level this was reflected by the 'good farm results', at the research team level by an 'efficient methodology to work with at farm level' and at the regional level by the consolidation of an Interinstitutional Network (Table 4). Difficulties at the farm level were associated with the decision making process and animal health management, whereas at the region level difficulties were due to interinstitutional coordination and the scope of the process. Suggestions for improvement were considered and most of them included in the CP (Table 3). Incorporating a specialist in the area of animal health was not possible.

Table 4. Stakeholder's perception of project achievements, difficulties and suggestions for improvement analysed during mid-term PIPA workshop (PW3, September 2013)

	Stakeholders' reflections
Project achievements	*"There are already good farm results".** • Unexpected productive performance (positive results) • High increase in meat production • Farm planning • Learning about production technologies • Continuous technical support • Be aware of a different way of working (*"we have changed our minds"*) "*There is an efficient methodology to work with at farm level*". • Interaction farmers – research team • Assess natural resources management in relation to production activities Consolidation of Interinstitutional Network (IN)

	• Knowledge acquisition by IN stakeholders • More linkage between regional organisations
Project difficulties	At the project farms • Difficulties with implementing changes on the farm. • Some technical issues uncovered ("*we need animal health assistance"*). At IN • *"Project coordination with other regional organisations was difficult at the beginning"* • *"We have difficulties following the process"* • *"How project results could reach other actors is not clear"*
Suggestions for project improvement	At the project farms • Include animal health plan in the re-design of the LFF • Include on-farm meetings with neighbours to learn about the LFF re-design process At the research team • Incorporate a veterinarian At IN • Exchange information among IN stakeholders on specific technologies being used at the farms • Efforts to reach more farmers: use mass media to enhance dissemination of project results • More coordination with other organisations, especially with MGAP to disseminate on-farm approach.

** Statements in italics are stakeholders statements recorded during PW3*

The last workshop (PW6) focused on a participatory evaluation of the whole project and tackled different topics: global project assessment, goal achievement, project performance and personal changes in knowledge and practices. The methodology used allowed the participants to immediately visualise the results of the survey and reflect on the process (Knowledge Sharing Toolkit, 2009), and was aligned with the whole participatory process (Home & Rump, 2015). The collective reflection on the individual responses constituted the global perception of project results and was later reinforced by the quantitative analysis of the survey.

The quantitative analysis of the survey valued the overall project performance positively: all topics were rated above 3 in a 5 point scale (Table 5). Global project assessment in particular was rated highly with a mean value of 4.22 out of 5, whereas the achievements reached at the farm level (questions 2a, 2b, 3 and 8) were among the highest values. The weakest points were related to knowledge of project results (question 5) and future impact of project results in the region (question 16), with mean values of 3.61 and 3.44 respectively. Several open

questions allowed participants to express their own ideas and perceptions of the process. Some comments of the participants were: "*The project improved over time"; "It is a very valuable experience"; "The methodology of co-innovation stands out"*.

Table 5. Final Project Evaluation results

	Question N° [1]	Question topic	Mean Value [2]
Global Project	1	Global project´s assessment	4.22
Goals Achievement	2 a	Changes in the farms	4.17
	2 b	Relevance of the farm changes	4.28
	3	Methodology used to work with project farmers	4.44
	4	Methodology used to work with local actors	3.67
	5	Knowledge of project's results	3.61
	6	Knowledge of technological information to be promoted by public policies	4.06
Project Performance	7	General project´s implementation performance	4.44
	8	Work at farm level	4.44
	9	Communication plan	3.94
	10	Interinstitutional coordination	4.17
	11	Incorporation of suggestions during the project	4.17
Other Topics	12	Fulfilling of personal expectations	4.06
	13	Knowledge related to LFF	3.94
	14	"New ways to do things"	3.72
	15	Personal feeling related to participation	4.39
Future	16	Impact of project´s results in the region	3.44

[1] See Appendix for detailed questions. [2] Mean values were calculated as the average of the responses. Each response was rated on a 5 point scale: ++ = 5, + =4, 0 = 3, - = 2, -- =1

PWs were the key-elements for interaction among stakeholders (Home & Rump, 2015). All project outcomes were possible because participation of IN stakeholders during PWs was adequately organised and facilitated. Considerable time and resources were allocated to this process (Klerkx & Leeuwis, 2009; Home & Rump, 2015). The social learning process could be seen as different stakeholders interacting to solve a problem, while simultaneously acquiring new skills (both technical and social), producing knowledge, as well as developing relationships (Schut et al., 2015). This participatory process continued as an interactive experimental learning cycle (Douthwaite et al., 2002).

As a result of the learning process, and inspired by the project's methodological and technological results, one stakeholder organisation established a project for another region. The CNFR Farmer Union had highly valued this way of working and presented a project to a competitive fund which, if approved, will allow them to obtain funds to work with other farmers using the co-innovation approach.

Lessons learned

Stakeholders successfully worked together in the IN over a three year period to support a participatory and collaborative process to generate innovation in the seven LFF, which

contributed to the success of the whole project and to the dissemination of its results. The stakeholders involved in the IN were strongly committed from the beginning of the project, their continuous engagement was essential in the building of network reliability (Akpo et al., 2015). Furthermore, some process characteristics were particularly relevant: (i) clear objectives; (ii) negotiation and facilitation; (iii) systematisation and keeping up with a certain continuity and coherence between PW; (iv) consideration of local culture to define when and where to set the PW and (v) a clear agenda for PW.

The shared vision of the project´s expected results along with the required activities to achieve them contributed to a clear and common understanding of the desired project outcomes and was also a source of motivation. As the impact pathways to achieve the project vision were validated and made explicit throughout activities, it was easier to M&E the advances and final results. However, specific indicators would be needed in the future for M&E. In our case, during PWs stakeholders reflected on the obtained results and shared ideas for project improvement in real time, identifying whether or not interventions successfully contributed to achieving the vision.

The co-elaboration of CPs by the IN with specific activities was a strong tool that generated interaction and promoted coordination among local organisations. From the activities organised for farmers and technicians to share results and to exchange ideas on new technologies, we would highlight on-farm field days. An evolution of project farmers' role was noticeable: after three years they directly explained to others the changes and associated results introduced in their LFF, reflecting the undergoing learning process.

The spaces generated with the PWs and the implementation of the CPs could be seen as platforms for social learning and innovation for farmers, researchers and local actors where the 'real world' actors are involved in the process. These spaces demanded time and resources to reach a common understanding of how the project would achieve the desired impact. Within this platform, the M&E allowed visualising of the changes and their relevance at farm level in a particular context, while making learning cycles explicit throughout the process.

Final reflections

This Uruguayan case demonstrated that applying the PIPA method enhances co-innovation at regional level and nurtures the creation of a common space for networking, participatory planning, M&E and social learning. Researchers, farmers and organisations were capable to plan, monitor and evaluate the project focusing on sustainable LFF production systems. Several learning cycles took place to adapt and adjust the project in real time, while strengthening the impact-oriented vision of the project. The PIPA method provided a good framework for innovation towards sustainable LFF providing a useful instrument to contribute to rural development. As stakeholders understood the benefits of the approach, they effectively used the new knowledge for their own organisations to identify key issues for future initiatives and to provide information for agricultural family farming policy makers.

Acknowledgements
The authors are grateful to the families who shared their experiences with us and to the local actors who participated in the PWs and in the elaboration and implementation of the CP. The authors also thank the whole project research team and Sophie Alvarez for their suggestions during the writing process, Fiorella Cazzuli for her support on language editing and INIA for its financial support.

References

Akpo, E., Crane, T. A., Vissoh, P. V., & Tossou, R. C. (2015). Co-production of knowledge in multi-stakeholder processes: Analysing joint experimentation as social learning. Journal of Agricultural Education and Extension 21(4): 369-388.

Albicette, M.M., Leoni, C., Ruggia, A., Scarlato, S., Albín, A., & Aguerre, V. (2016). A co-innovation approach in family-farming livestock systems in Rocha-Uruguay: a three-year learning process. IFSA 2016. Shropshire, UK.

Alvarez, S., Douthwaite, B., Thiele, G., Mackay, R., Córdoba, D., & Tehele, K. (2010). Participatory impact pathways analysis: a practical method for project planning and evaluation. Development in Practice 20(8): 946-958.

Bennett, C. (1975). Up the hierarchy. Journal of Extension 13(2): 7-12.

Capra, G., Albin, A., Ruggia, A., García, F., Aguerre, M., Vilaró, F., Gilsanz, J., Arboleya, J., & Rachetti, M. (2009). Tecnología para la producción familiar del departamento de Rocha: avances del acuerdo de trabajo INIA-CNFR. Revista INIA. No 18. pp 40-44. Uruguay.

Coutts, J., Botha, N., & Turner, J. (2014). Evaluating a Co-innovation Policy Initiative in New Zealand. In: Proceedings of Farming Systems Facing Global Challenges. IFSA 2014. Berlin.

Dogliotti, S., Abedala, C., V. Aguerre, V., Albín, A., Alliaume, F. Alvarez, J., Bacigalupe, G. F, Barreto, M., Chiappe, M., Corral, J., Dieste, J. P., García de Souza, M. C, Guerra, S., Leoni, C., Malán, I., Mancassola, V., Pedemonte, A., Peluffo, S., Pombo, C. Salvo, G., & Scarlato, M. (2012). Desarrollo sostenible de sistemas de producción hortícolas y hortícola-ganaderos familiares: una experiencia de co-innovación. INIA, Uruguay. (Serie FPTA; 33).

Dogliotti, S., García, M. C., Peluffo, S., Dieste, J. P., Pedemonte, A. J., Bacigalupe, G. F., Scarlato, M., Alliaume, F., Alvarez, J., Chiappe, M.,& Rossing, W. A. H. (2014). Co-innovation of family farm systems: a systems approach to sustainable agriculture. Agricultural Systems 126: 76-86.

Douthwaite, B., Keatinge, J. D. H., & Park, J. R. (2002). Learning selection: an evolutionary model for understanding, implementing and evaluating participatory technology development. Agricultural Systems 72: 109-131.

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: 243-265.

Douthwaite, B., Schulz, S., Olanrewaju, A., & Ellis-Jones, J. (2007a). Impact pathway evaluation of an integrated *Striga hermonthica* control project in Northern Nigeria. Agricultural Systems 92: 201-222.

Douthwaite, B., Alvarez, S., Cook, S., Davies, R., George, P., Howell, J., Mackay, R., & Rubiano, J. (2007b). Participatory impact pathways analysis: a practical application of program theory in research-for-development. The Canadian Journal of Program Evaluation 22 (2): 127-159.

Home, R., & Rump, N. (2015). Evaluation of a multi-case participatory action research project: the case of SOLINSA. The Journal of Agricultural Education and Extension 21 (1): 73-89.

IAASTD. International assessment of agricultural knowledge, science and technology for development. (2009). Executive summary of the synthesis report. pp. 1-23. Washington: Island Press

Klerkx, L., & Leeuwis, C. (2009). Operationalising demand-driven agricultural research: institutional influences in a public and private system of research planning in The Netherlands. Journal of Agricultural Education and Extension 15(2): 161-175.

Klerkx, L., Hall, A., & Leeuwis, C. (2009). Strengthening agricultural innovation capacity: are innovation brokers the answer? International Journal of Agricultural Resources, Governance and Ecology. 8(5/6) 409-438.

Klerkx, L., van Mierlo, B., & Leeuwis, C. (2012). Evolution of systems approaches to agricultural innovation: concepts, analysis and interventions. In I. Darnhofer, D. Gibbon and B. Didieu (Eds) Farming Systems Research into the 21st century: The New Dynamic pp 457-483. Dordrecht:Springer.

Knowledge Sharing Toolkit. (2009). FAO, CGIAR. Retrieved January 13th, 2016. http://www.kstoolkit.org/

Leeuwis, C. (2000). Reconceptualising participation for sustainable rural development: towards a negotiation approach. Development and Change 31: 931-959.

Leeuwis, C., & Van der Ban, A. (2004). Communication for rural innovation: rethinking agricultural extension. Oxford: Blackwell Science.

Loeber, A., Mierlo, B. v., Grin, J., & Leeuwis, C. (2007). The practical value of theory: conceptualising learning in the pursuit of a sustainable development. In A. E. J. Wals (Ed.) Social Learning towards a Sustainable World. Principles, Perspectives and Praxis pp. 83-97. Wageningen Academic Press

Masera, O., Astier, M., & López-Ridaura, S. (2000). Sustainability and natural resource management. The MESMIS evaluation framework. Working Document D36. Grupo Interdisciplinario de Tecnología Rural Apropiada A.C. México.

Moschitz, H., Roep, D., Brunori, G., & Tisenkopfs, T. (2015). Learning and innovation networks for sustainable agriculture: processes of co-evolution, joint reflection and facilitation. Journal of Agricultural Education and Extension 21(1): 1-11.

Okali, C., Sumberg, J., & Farrington, J. (1994). Farmer Participatory Research. Rhetoric and Reality. London: ODI.

Rockwell, K., & Bennett, C. (2004). Targeting outcomes of programmes: a hierarchy for targeting outcomes and evaluating their achievement. Faculty Publications: Agricultural Leadership, Education & Communication Department, University of Nebraska. Lincoln. Paper 48.

Rossing, W.A.H., Dogliotti, S., Bacigalupe, G.F., Cittadini, E., Mundet, C., Mariscal Aguayo, V., Douthwaite, B., & Alvarez, S. (2010). Project design and management based on a co-

innovation framework. In Building Sustainable Rural Futures: The Added Value of Systems Approaches in Times of Change and Uncertainty pp. 402-412. IFSA 2010. Austria.

Schut, M., Klerkx, L., Rodenburg, J., Kayeke, J., Hinnou, L.C., Raboanarielina, C.M., Adegbola, P.Y., van Ast, A., & Bastiaans, L. (2015). RAAIS: Rapid Appraisal of Agricultural Innovation Systems (Part I). A diagnostic tool for integrated analysis of complex problems and innovation capacity. Agricultural Systems 132:1-11.

Speelman, E.N., López-Ridaura, S., Colomer, N.A., Astier, M., & Masera, O.R. (2007). Ten years of sustainability evaluation using the MESMIS framework: lessons learned from its application in 28 Latin American case studies. International Journal of Sustainable Development & World Ecology 14: 345-361.

Tommasino, H., Cortelezzi, A., Mondelli, M., Bervejillo, J., & Carrazzone, S. (2014). Tipología de productores agropecuarios: caracterización a partir del Censo Agropecuario 2011. In Anuario 2014 pp 491-508. Uruguay: OPYPA.

Wiesmann, U. Hirsch-Hadorn, G., Hoffman-Riem, H., Bibber-Klemm, S., Grossenbacher, W., Joye, D., Pohl, C., & Zemp, E. (2008). Enhancing transdisciplinary research: a synthesis in fifteen propositions. In Hirsch Hadorn, G., Hoffmann-Riem, H., Biber-Klemm, S., Grossenbacher-Mansuy, W., Joye, D., Pohl, C., Wiesmann, U. and Zemp, E (Eds) Handbook of transdisciplinary research. Chapter 27, pp. 433-444, SpringerLink.

Appendix

FINAL PROJECT EVALUATION

Co-innovating for sustainable development of family farm production systems in Rocha-Uruguay

To answer the following questions there are five alternatives:
(++) (+) (0) (-) (--)
Please indicate with (X) which best reflects your opinion.

OVERALL

1) How do you assess the project globally?
Excellent ( ) ( ) ( ) ( ) ( ) Very bad
Comments:

_____

_____

GOAL ACHIEVEMENT:

Considering the objectives of the project and the shared vision generated by the interinstitutional network (IN), assess the following:

2) Did the project allow positive changes in the 7 farms considering their sustainability?
a) Many changes ( ) ( ) ( ) ( ) ( ) No changes
b) Very relevant ( ) ( ) ( ) ( ) ( ) No relevant
Please indicate the most significant changes for you:

_____

3) Was an appropriate methodology used in working with project farmers?
Very appropriate ( ) ( ) ( ) ( ) ( ) Inadequate
Comments:

_____

4) Was adequate methodology used to promote networking and thus contribute to regional development?
Very appropriate ( ) ( ) ( ) ( ) ( ) Inadequate
Comments:

_____

5) Do farmers, professionals and local organisations know the results of the project?
Know much ( ) ( ) ( ) ( ) ( ) Unknown
Comments:

_____

6) Do farmers, professionals and organisations know high-impact information technology to be promoted through public policies for family farming?
To a high degree ( ) ( ) ( ) ( ) ( ) to a low degree
Comments:

_____

_____

PROJECT IMPLEMENTATION

7) Generally speaking, how do you think the project has worked during this time?
Appropriately ( ) ( ) ( ) ( ) ( ) Unsuitably
Please indicate the most significant aspects for you:

_____

8) To what extent was it appropriate at farm level?
To a high degree ( ) ( ) ( ) ( ) ( ) to a low degree
Comments:

_____

9) To what extent were the activities foreseen in the communication plan drafted by the IN appropriate?
To a high degree ( ) ( ) ( ) ( ) ( ) to a low degree
Comments:

_____

10) To what extent was the institutional coordination of the activities from the project adequate?
To a high degree ( ) ( ) ( ) ( ) ( ) to a low degree
Comments:

_____

11) To what extent you believe the project was "permeable" to suggestions for improvement made by yourself during the process?
To a high degree ( ) ( ) ( ) ( ) ( ) to a low degree
Comments:

_____

_____

OTHER TOPICS

12) How far were your expectations regarding the project fulfilled?
To a high degree ( ) ( ) ( ) ( ) ( ) to a low degree
Comments:

_____

13) To what extent have you improved your knowledge on technological strategies for family livestock?
To a high degree ( ) ( ) ( ) ( ) ( ) to a low degree
Please indicate the most significant aspects for you:

_____

14) To what extent do you have "new ways of doing things" in relation to your work with livestock farming?
To a high degree ( ) ( ) ( ) ( ) ( ) to a low degree
Please indicate the most significant aspects for you:

_____

15) How did you feel about your participation in this process?
Comfortable ( ) ( ) ( ) ( ) ( ) Uncomfortable
Comments:

_____

_____

TOWARDS THE FUTURE

16) To what extent do you believe the project results may impact in the region?
To a high degree ( ) ( ) ( ) ( ) ( ) to a low degree
Please indicate three aspects that facilitate this process:

_____

Please indicate three aspects that limit this process:

_____

17) Based on the experience and learning generated within the framework of this project, which suggestions would you make to policy makers when defining public policies for family farming?

Name three aspects or policies considered critical to support family farming in Uruguay:

_____

I would also like to add:

_____

INDIVIDUAL CHARACTERISTICS

Are you?

Mark with an X the option/s that correspond

(I) Project farmer? ----------

(II) Farmer representative of a SFR? ----------

(III) If you represent a SFR, please give us its name ----------

(IV) A professional ----------

(V) If you are a professional, which organisation do you belong to ……

THANK YOU

What is capacity to innovate and how can it be assessed? A review of the literature

Allebone-Webb, S.[1], Douthwaite, B.[2], Hoffecker, E.[3], Mathé, S.[1,4] and Triomphe, B.[1]

[1] *CIRAD, UMR Innovation*
[2] *Consultant, Enabling and Evaluating Innovation*
[3] *Massachusetts Institute of Technology*
[4] *International Institute of Tropical Agriculture, Cameroon*

Abstract: 'Capacity to innovate' is an emerging concept, especially in agriculture and rural development. There is no universally agreed definition for this concept, but many authors agree that it refers generally to the ability of actors to continuously identify constraints and opportunities and to mobilise capabilities and resources in response – i.e. to produce and sustain innovation processes in a dynamic systems environment. Increasingly, capacity to innovate (C2I) is recognised as playing a critical role in successfully responding to a changing external environment. Facilitating and building this capacity is therefore crucial for adaptable farming systems and for improving the resilience and livelihoods of poor farmers and other rural actors. This paper summarises the findings of a targeted literature review aiming to unpack the concept of C2I, exploring its meaning across all research sectors and ways to assess it in agricultural communities. We propose that the various dimensions of C2I identified through the literature review can be a starting point for developing an assessment framework to measure changes in C2I. Specifically, we identify four core capacities that make up C2I: (1) to envision and create new ways of doing things; (2) to connect with others to access and understand new information and resources; (3) to experiment, test, assess and adapt; and (4) to work with others to achieve action and change. We review previously described indicators to measure these concepts and accordingly propose an initial set of metrics for use in agricultural communities. We conclude that the C2I concept puts a spotlight on process-driven approaches to innovation that have previously been undervalued.

Keywords: Innovation, capacity, agricultural innovation systems, capacity to innovate, metric, capabilities

Introduction

Innovation plays a fundamental role in economic development and is considered a key factor in determining the 'success' of societies, sectors and firms (Bell & Albu, 1999; Freeman, 1987; Mytelka, 2000). Defined as "the new use of existing or new ideas or the combination of ideas that have social or economic significance" (Mbabu & Hall, 2012), innovation is increasingly seen as critical to achieving economic, social and environmental goals in a rapidly changing world (Jones, 2004). While not a panacea nor an end in itself (sometimes resisting change may be what is needed), agricultural innovation may be particularly vital for feeding a growing global population in a sustainable manner (FAO, 2014; Jones, 2004), and is especially important in developing countries where agriculture plays a critical role in the local and national economy (Thomas & Slater, 2006; World Bank, 2008).

Conventional approaches towards agricultural innovation involve the creation of new technologies by research and development organisations, and then 'pushing' them to farmers and other end-users. This assumes that the lack of (adequate) technology is the primary obstacle to agricultural innovation and development. However, the limitations of this technology-led approach have been increasingly recognised (Clark, 2005; Hall et al., 2007; Johnson & Segura-Bonilla, 2001). Many scholars and practitioners acknowledge that the constraints to agricultural innovation and development are not only the ability to produce new knowledge or technologies, but also the ability of stakeholders to put relevant knowledge and technological inventions into use. This includes adapting inventions and practices to rapidly changing conditions and locally-specific contexts and often requires changes to social, economic, institutional and technological systems (Chataway et al., 2005; Hall et al., 2007; Schut et al., 2015).

The understanding of the importance and nature of innovation led to the development of the innovation systems concept, defined as the complex networks of interacting actors (individuals, organisations and enterprises) involved in developing and putting an innovation into use, together with the institutions and policies that support this (World Bank, 2007). Innovation systems thinking is now commonly applied to agriculture, as Agricultural Innovation Systems (AIS) (Assefa et al., 2009; Klerkx et al., 2012; Pant & Hambly Odame, 2009). Along with the emergence of innovation systems analysis, and reflecting the importance of actors' capacities to engage in innovation processes, over the past twenty years a related concept has emerged, that of 'capacity to innovate' or C2I.

As a concept, C2I is significant not only in the agricultural sector (Schut et al., 2015) but also in business (Hult et al., 2004), medicine (Caccia-Bava et al., 2006), engineering (The Royal Academy of Engineering, 2012), education (Grogger & Hanson, 2011) and in relation to national innovation systems (Wonglimpiyarat, 2010). It is closely related to the concepts of adaptive capacity and capacities for social learning and has been increasingly seen as playing a key role in helping local system actors respond effectively to rapidly changing external contexts, including climate change (Berkes, 2007; Lybbert & Sumner, 2012). Despite this, the lack of a universally accepted definition for C2I reflects a certain 'fuzziness' about what it means (Chuluunbaatar & LeGrand, 2015; Hall, 2005; Hall et al., 2007).

With the greater focus on C2I has come an increasing concern over how to evaluate it (Furman et al., 2002; OECD, 2012). Measuring C2I is important to evaluate the efficacy of interventions and to assess C2I changes over time, and for this more robust monitoring and evaluation (M&E) tools are needed than those currently available.

This paper proposes an approach to developing metrics for assessing C2I. By taking a broad look at the growing literature on C2I and related terms, we review how this concept has been defined and identify its key conceptual components. From this, we identify the dimensions of the concept which are particularly relevant to agricultural innovation systems and propose a framework for understanding and for measuring C2I. This framework then serves as the starting point for developing a proposed set of metrics and indicators for assessing C2I within the context of rural communities.

Searching the literature: methods and bibliometric data

Bibliographic searches

We searched for references to C2I across a range of peer-reviewed and practitioner publications, including in the Scopus, Web of Science and AGRIS databases, and by searching donor, implementer and research institution websites. Our search focused on keywords used in the literature, covering all terms related to "capacity to innovate", including "capacity for innovation", "innovation capacity", and "innovation capability". The resulting documents were screened for relevance if they made reference to:

- C2I concept and/or component capacities;
- interventions aiming (explicitly or implicitly) to improve C2I; *or*
- indicators or methods of assessment or evaluation of capacity to innovate

While reading these papers, relevant cited references were also added to the database, as were references suggested directly by a handful of knowledgeable resource persons.

Bibliographic results

From a total 2254 documents retrieved through the above searches, 748 passed title and abstract screening and 149 passed full text screening as referring to the C2I definition, concept/component capacities, interventions or indicators. As expected for an emerging topic, the number of documents retrieved by year of publication has increased sharply since 2000 (Figure 1). Of these, more papers used "innovation capacity" than "capacity to innovate".

More literature was published[1] in China than any other country (**Error! Reference source not found.**), probably due to a strong national emphasis on it becoming an "innovation-oriented nation" by 2020 (Zhang & Wu, 2012). Other major sources of references include the USA and then Western European countries (UK, Spain, Germany, France). By subject, C2I occurs most in business and management literature (21%), followed by social science (18%) and engineering (14%). However, this varied by country, with over half of all publications from China being from business and management sectors. Agricultural and biological sciences are only 7th on the list, illustrating that the C2I concept is used in a number of contexts beyond that of agricultural innovation systems.

[1] The country/territory detailed in Scopus is determined by the location of the publisher, see Scopus coverage guide https://www.elsevier.com/__data/assets/pdf_file/0007/69451/scopus_content_coverage_guide.pdf

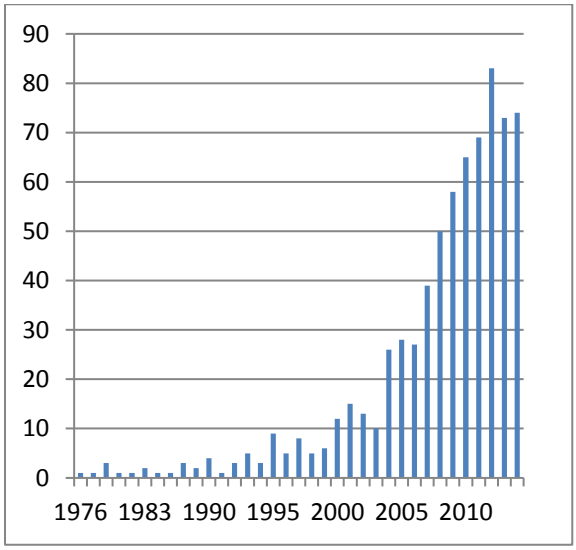

Figure 1. Number of publications related to C2I by year (N = 749 papers passing title and abstract screen)

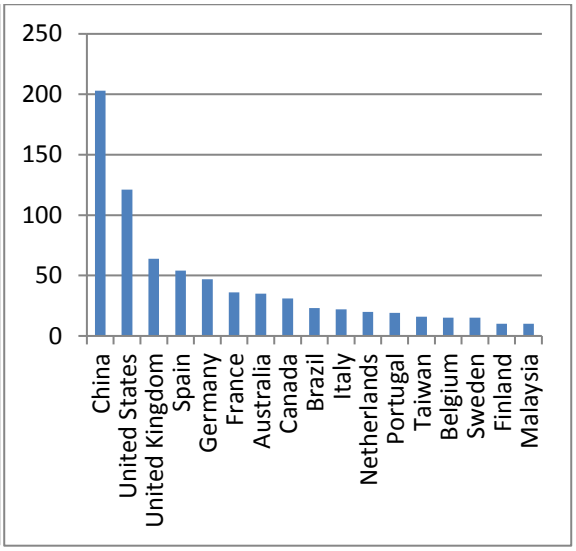

Figure 2. Number of publications by country from countries with 10 or more publications (N = 741)

Capacity to innovate: definitions and component capacities from the literature

Definitions of capacity to innovate

Coined by Burns and Stalker (1961)[2], the term *capacity to innovate* (C2I) has changed over time and across sectors of use. Early use of the term described it as the capacities "to successfully adopt and implement innovations", seen as distinct from the capacities needed to "initiate and be receptive to innovations" which were termed *innovativeness* (Hurley & Hult, 1998; Hurley et al., 2005). Modern use of the C2I term includes these two capacities and additional sub-capacities which are seen as integral to the ability to produce and/or to use innovation (Hall et al., 2009; Leeuwis et al., 2014; Mayne & Douthwaite, 2015).

For some, C2I is defined simply (and redundantly) as the increased capacity to be able to innovate. Others have chosen not to give a one-phrase definition, going straight into detailing what component capacities are encompassed by the C2I concept. The few non-redundant, one-sentence definitions describe the capacities to access new innovations and apply them over time, e.g. "the continuing ability to combine and put into use different types of knowledge" (Chuluunbaatar & LeGrand, 2015). See Table 1 for a list of distinct, relevant definitions identified.

[2] *As reported in (Hurley & Hult, 1998)*

Table 1. Definitions of C2I and IC (in chronological order)

Author	Definition or description of C2I or related terms	CI term[3]	Research sector
Cohen & Levinthal (1990)	"the ability of a firm to recognise the value of new external information, assimilate it and apply it to commercial ends"	Absorptive capacity; Innovative capabilities	Business
Hurley & Hult (1998)	"The *capacity to innovate* […] is the ability of the organization to adopt or implement new ideas, processes, or products successfully" "*Innovativeness* is the notion of openness to new ideas as an aspect of the firm's culture [it] is a measure of the organisation's orientation toward innovation." "Innovative capacity relates to […] absorptive capacity".	C2I; Innovativeness Innovative capacity; Absorptive capacity	Business
Neely & Hii (2001)	"Innovative capacity is the internal potential of a firm to generate new ideas, identify new market and technological opportunities, and implement innovations by leveraging resources and capabilities. In short, innovative capacity determines a firm's ability to innovate."	Innovative capacity	Business & management SME
Hult et al. (2004)	"Innovativeness is defined here as the capacity to introduce some new process, product, or idea in the organisation"	Innovativeness	Business
Caccia-Bava et al. (2006)	"the organisation's capacity to innovate (absorptive capacity), [is] the organisation's ability to recognise the value of new information, assimilate it, and apply it to productive ends…"	C2I Absorptive capacity	Health
Skiltere & Jesilevska (2013)	"the ability to generate new knowledge, new technology and new artefacts and to apply these novelties in a useful way. The concept of innovative capacity evaluates not only the current capabilities to innovate but also the innovative potentials that may affect innovativeness in the longer period of time."	Innovative capacity; IC; Innovative potential; Innovativeness	Business & economics
Nair et al. (2014)	"Innovation capacity is the collective ability of a firm to look into future through the eyes of customers and reengineer products and services accordingly"	IC	Business
Chuluunbaatar & LeGrand (2015)	"the continuing ability to combine and put into use different types of knowledge"	C2I IC	Agriculture

[3] The principal term (C2I, IC, etc) used for the definition is listed first, with other terms used interchangeably in the same text listed afterwards

Mayne & Douthwaite (2015)	"Capacity to innovate is then the ability to combine some or all of hardware, software and orgware to bring about innovation"	C2I	Agriculture
Turner et al. (2015)	"Innovation capacity is the capability of actors to continuously identify and prioritise constraints, and in response mobilise new and existing capabilities and resources, i.e. adapt to realise opportunities in a dynamic systems context". "to mobilise, combine and create resources and capabilities to successfully innovate"	IC C2I	Agriculture

Key terms related to the concept of C2I

As Table 1 illustrates, the terms *innovation capacity* (IC) and C2I have often been used interchangeably (e.g. Chuluunbaatar & LeGrand, 2015; Turner et al., 2015). However, particularly in business and management research, IC may be used to refer to how many innovations an organisation can produce and implement successfully, rather than the capacities needed to do so. These definitions of IC may include 'structural properties' of an organisation (e.g. measures of organisation size, finance and machinery) that most definitions of C2I do not include, as well as 'human qualities' such as communication, tolerance for risk-taking, power sharing, learning, collaboration, and participative decision-making that do align with C2I capacities (Aiken et al., 1980; Hurley & Hult, 1998). Thus, understanding of IC may vary with sector from being comparable to C2I, to being something much broader. We have not seen C2I used in this broader sense, and so consider it as the less ambiguous and therefore preferable term.

Another area where terms may be ambiguous relate to the concept of *capacity* itself. In the literature, *competence*, *capability*, and *capacity* are used in relation to innovation and yet the distinction is not always clear. Some authors use these in a nuanced way, differentiating between *capacity* as an overall ability of individuals, groups or systems to do something and *capabilities* as specific sets of skills (Mauerhofer, 2010; Pant, 2012; Sen, 1993; Skiltere & Jesilevska, 2013), while others use these interchangeably (Neely & Hii, 2001). While we appreciate the distinction, given that these related terms and concepts are not always used consistently, we have decided to pull all capacities, capabilities and competencies together in Figure 4.

The ***process*** by which individuals, organisations and societies obtain, strengthen, adapt and maintain capacity to set and achieve objectives over time has been called capacity development (CD) (UNDP, 2009), capacity strengthening (Hartwich et al., 2007) or capacity building (DFID, 2008), each having slightly different meanings (Hambly & Sarapura, 2009; Pant, 2012). For the purposes of this paper, we will treat them as overlapping aspects within the same overall process, which we will call CD. Thus, CD for innovation can be thought of as the process by which C2I may be achieved. Some authors indeed specifically equate CD for AIS as C2I (TAP, 2016) or IC (Pound & Essegby, 2008).

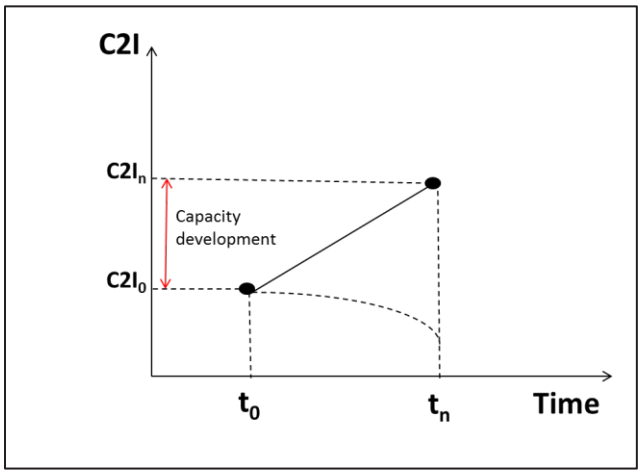

Figure 2. Conceptual relationship between C2I (a state which changes dynamically over time) and CD (a process and set of activities which contributes to increasing C2I). The dashed line indicates that while the goal is to increase C2I, it may go up or down over time.

What specific capacities are needed to innovate?

The literature on C2I points to certain capacities and capabilities required by individuals, organisations and/or institutions which when combined create the capacity needed to innovate and sustain innovation processes over time. We sorted these into groups of capacities that were most alike or linked, giving four broad capacity groups each of which may occur at, or are supported by, individuals, organisations and the enabling environment, and which can be further divided into several sub-capacities (Figure 4):

(1) To envision, create and be open to new ways of doing things - to individually and/or jointly envision something new and improved;
(2) To connect with others to access and understand new information and resources – to form new connections and to use both new and existing relationships with diverse actors (individuals and entities) to obtain, share and understand information and resources;
(3) To iteratively experiment, test, assess and adapt – to conduct experimentation involving iterative learning and improved processes and results over time; and
(4) To work with others to achieve action and change - to work together formally and informally in order to take effective collaborative action and achieve common objectives.

The recognition of capacities to envision, generate and welcome new ideas **(1)**, as separate from capacities to adapt **(3)** apply **(4)** those innovations, reflects earlier definitions of *innovativeness* (Hurley & Hult, 1998; Hurley et al., 2005), described as essentially a cultural trait (Woodside, 2005). Turner et al. (2015) named a similar grouping of capacities *innovation capabilities*, describing them as "processes for exploring and exploiting opportunities to innovate", and encompass the capacities we describe in **(1)** while also overlapping with some of the capacities described under **(2)** and **(3)**.

The capacities to connect with others to access and understand new information **(2)** are most closely aligned with definitions of *potential absorptive capacity*, which is the capacity to acquire and assimilate knowledge (Zahra & George, 2002). Some authors have bundled the concepts of acquiring and assimilating knowledge together with the capacities to use and apply that information. Both are reliant on networks and encompassed by definitions of *absorptive capacity* (Cohen & Levinthal, 1990) or *absorptive capability* (Turner et al., 2015). In this paper we have put the capacities to use and apply knowledge into a separate grouping - the capacity for collaborative action **(4)**. We consider that these two sets of capacities are inherently different as 'to understand and know' does not automatically translate into 'being able to do'.

Figure 3. Capacities to innovate

1. To envisage, create and be open to new ways of doing things

1.1 To generate new ideas and foster creativity: •Capacity to generate new ideas, products, processes for action[1,2,3,4,5,6]; •To foster creativity[5,7,8,9,10]; •Entrepreneurial spirit[8]

1.2 To be open to new ideas and actions (individuals, leaders and organisations)[1,3,4,8]

1.3 To identify and prioritise problems and opportunities and adapt/explore them accordingly[6,11,12,13,14,15,16]

2. To connect with others to access and understand new information & resources:

2.1 To link with others/network: •Develop, maintain and use effective networks[5,9,11,12,14,17,18,19,20,21,22,23,24,25];
•To intermediate/facilitate/broker for linkages, interactions and networks[12,22,26,27,28,29,30,31,32,33,34,25,6];
•Institutions support networks and collaboration, policy supports the development of networks[35]

2.2 To access, share and process information:
•To have processes for acquiring, assimilating and transforming external knowledge[12,13,36,15];
•Capacity to link with others to access, share and process information[37,17,38,31,4,32,34,14,5,25]; •Institutions support knowledge sharing and interactive learning[27,23,4,6]

2.3 To understand and learn to process information:
•Understand new knowledge (ideas, things, resources) and put to (productive) use[1,17,11,39,9,25]; •Capacity for reflection and learning[8, 39,12,26,22,4,5,25,40,41]; •Organisations and Institutions support learning[1,27,23,4,5,6]

3. To iteratively experiment, test, take risks, analyse, assess

3.1 To test, experiment and assess
•To experiment and assess arising trade-offs[17,42,3,31,14]; •Institutions support social and technical experimentation [12,42,43,30,31,23,33];
•Capacity to assess and take risks and a culture that supports that[38,9,42,31,14,6]

3.2 To adapt to change, be flexible:
•Ability to change approach and partnerships/ networks/ interactions in response to change[44,11,12,13,45,30,22,5,15];
•Embed innovation and research activity in ongoing process of change[22,14,6];
•Leadership, institutions and culture support & embrace change and allow for rapid response / adaptive management[44,45,42,3,31,23,46,14,5,6];
•Flexible solutions to allow for revision[12,43]

4. To work with others to achieve action and change

4.1 To be motivated and to motivate others: •Individuals motivated to participate[11,9,30,14,40];
•Project champions[47,26,30,31]

4.2 To work with others effectively to achieve action: •Collaborate and work with others to achieve action[38,18,32,14,25,16]; •Capacity to mobilise resources and form support coalitions around promising options[22,23,14]; •Share risks and benefits/Diversify risks and share uncertainties[38,22,3]; •Institutions for sharing risks and benefits[12,31,23]; •Build a shared vision/goal and realise shared values[5,25]

4.3 To mediate and facilitate: •Actively manage interdependent and unpredictable interactions among network partners[12,28,22,33]; •Leaders and facilitators orchestrate and facilitate to enable action, can understand how change happens and how to intervene effectively[14]; •Mediate diverse groups with different skills[5], mediate power-imbalances[43]; •Allow all members of the group to influence decisions[1]; •Leadership able to balance individual and collective interests to meet individual and collective needs[12,43,3,46]

[1](Hurley & Hult, 1998); [2](Johnson & Segura-Bonilla, 2001); [3](Rufat-Latre et al., 2010); [4](Pant, 2012); [5](Nair et al., 2014); [6](Hueske et al., 2015); [7](King & Anderson, 1990); [8](Hult et al., 2004), [9](Howard & Gillies, 2009); [10](Yang & Konrad, 2011); [11](Hall, 2005); [12](Smart et al., 2007); [13](Wang & Ahmed, 2007); [14](Leeuwis et al., 2014); [15](Turner et al., 2015); [16](TAP, 2016); [17](Dalohoun, 2005); [18](Hartwich et al., 2007); [19](Carlsson & Sandström, 2008); [20](Rohrbeck et al., 2009); [21](Chatenier et al., 2010); [22](Klerkx et al., 2010); [23](Musiolik et al., 2012); [24](Lambrecht et al., 2014); [25](Chuluunbaatar & LeGrand, 2015); [26](Douthwaite et al., 2009); [27](Klerkx et al., 2009); [28](Adner & Kapoor, 2010); [29](Dijkman, 2010); [30](Douthwaite & Gummert, 2010); [31](Traitler et al., 2011); [32](World Bank, 2012); [33](Brusoni & Prencipe, 2013); [34](Hermans et al., 2013); [35](Ugbe, 2010); [36](Boly et al., 2014); [37](Hosmer, 1995); [38](Adner, 2006); [39](Caccia-Bava et al., 2006); [40](Mayne & Douthwaite, 2015); [41](TAP, 2016); [42](Röling, 2009); [43](Chatenier et al., 2010); [44](Senge, 1990); [45](Fernandez-Gimenez et al., 2008); [46](Nettle et al., 2013); [47](Douthwaite, 2002).

Capacities are interlinked and multi-dimensional

The broad groups of capacities presented in the previous section and Figure 4 can be organised in a multitude of ways: the proposed grouping is not definitive and these capacities are interlinked, overlapping and exist on different dimensions. For example, while we have considered that 'to learn and understand' **(2.3)** most closely aligns with our C2I group of capacities 'to connect with others to access and understand new information and resources' **(2)**, learning and understanding is also necessary for iterative experimentation.

In trying to group capacities, it is useful to think of 'higher-level' capacities needed to produce or sustain innovation compared to 'building block' capacities or sub-capacities that underpin or precede them. For example, the 'higher-level' capacity to envisage, generate and access new ideas assumes the presence of 'building-block' capacities to form and access networks (from where to find new ideas), to learn and understand those ideas, and to identify opportunities (see Figure 4 which shows some of the main links between these capacities).

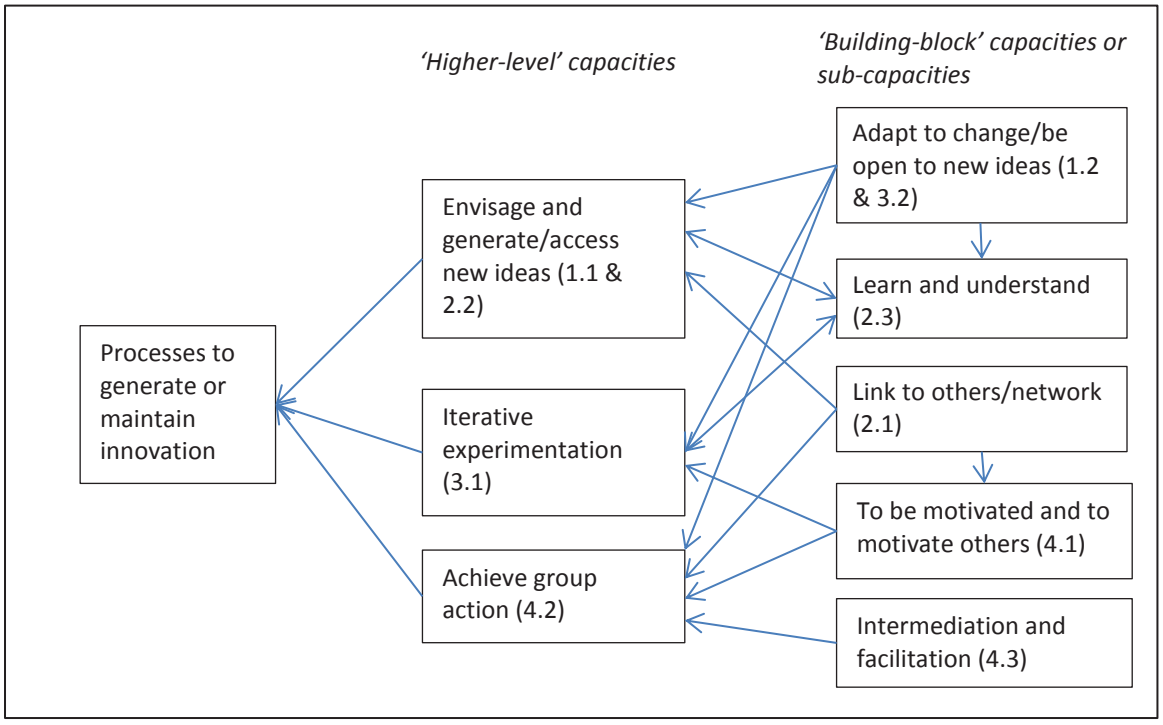

Figure 4. Main links between 'building-block' and 'higher-level' capacities

How to assess C2I

Evaluating C2I has become a concern for those working to strengthen AIS (Furman et al., 2002), as well as within firms, and at regional and national levels (OECD, 2012). In addition to measuring the efficacy of interventions for reporting purposes, it is also important to assess C2I in order to inform organisational learning and the ability of programs and interventions to adapt (Mayne & Douthwaite, 2015). Measuring C2I is difficult as many of the desired results refer to processes and have no clear completion mark (Daane et al., 2009). In addition, it is important that measures be useful and accurate, particularly given that the way in which something is defined and assessed often affects how it is managed (Chuluunbaatar & LeGrand, 2015).

Indicators proposed in the literature

The literature reviewed outlines two main types of indicators for C2I. Most common are measures to assess C2I as a whole, using outcome indicators reflecting the presence or change in this overall capacity. Also proposed are measures aiming to assess the capacity directly using indicators linked to essential elements of C2I.

Measurements linked to innovation outputs and outcomes

Given that C2I refers to the ability to produce innovation, overall capacity can be assessed by looking at innovation outcomes. Thus, if C2I has increased over time, we should expect to see evidence of more or improved quality innovations, more effective innovation processes and/or innovation activity spread more broadly.

This is essentially what most business, national and regional measures of C2I assess. Common measures include indicators of new product output or the number of patents and patent citations (proxy measures of new product development) (Skiltere & Jesilevska, 2013; Song et al., 2014).

Outcome measures of C2I proposed in agricultural contexts include measures of up-scaling and out-scaling such as a) interlinked technical and social-institutional innovations, b) innovations being tested outside the initial intervention area, c) growing coalitions for change and d) lessons learned/principles/methods/strategies adopted elsewhere (Leeuwis et al., 2014; Mayne & Douthwaite, 2015). Others use higher-level measures of development and well-being such as job creation and income (Dalohoun, 2005).

Measurements linked directly to C2I

Within firms, there has been a move away from innovation output-only metrics towards the evaluation of multiple factors, including indicators of C2I itself (Boly et al., 2014). Similarly, most proposed measures of C2I within agricultural systems include indicators which seek to directly measure the component capacities of C2I. 'Opening the black box' allows us to see the extent to which C2I may have changed, even if innovation processes are still mid-course, which may be useful given that innovation processes typically take time, so there may be a considerable lag between the time when C2I is developed and the time when it manifests through specific measurable innovation outputs. It also facilitates understanding of which aspects of C2I have changed, which may be important for research, as well as for programs that seek to strengthen specific dimensions of C2I.

Indicators of C2I can therefore be categorised according to the type of capacities they measure, using the same general groupings proposed above. Thus indicators related to **(1) envisioning, generating or being open to new ideas** include measures of a change in mind-set, attitude, confidence or conducive modes of thinking (Leeuwis et al., 2014; Mayne & Douthwaite, 2015; Van Veldhuizen & Water-Bayer, 1997); or responsiveness of organisations to innovation opportunities (Spielman & Kelemework, 2009).

Indicators related to **(2) connecting to others to access and understand new information and resources** focus on assessing a) the scale of networks (e.g. # networks and initiatives involved in social enquiry/learning, or the diversity of those networks) (FARA, 2014; Leeuwis et al., 2014; Spielman et al., 2011; Spielman & Kelemework, 2009; Temel, 2004; Van

Veldhuizen & Water-Bayer, 1997); b) the use of those networks to access information (Clark, 2006; Dalohoun, 2005; FARA, 2014; Jang et al., 2002; Leeuwis et al., 2014; Michailova & Husted, 2003; Van Veldhuizen & Water-Bayer, 1997); or c) learning and development or changes in learning processes (Boly et al., 2014; Dalohoun, 2005; Hurley & Hult, 1998).

Indicators of **(3) testing, experimenting and analysing** focus on the number of technical and social experiments done, which may include the number of novelties identified, tested, or discarded, and changes in the way that selection decisions are made (Leeuwis et al., 2014; Mayne & Douthwaite, 2015; Van Veldhuizen & Water-Bayer, 1997).

Finally, indicators to assess **(4) achieving action as a group** look at a) the number or scale of new ideas or practices adopted (Dalohoun, 2005; Hurley & Hult, 1998; Mayne & Douthwaite, 2015); b) organisational development (Van Veldhuizen & Water-Bayer, 1997) or coalition formation around promising initiatives (Leeuwis et al., 2014); c) measures of power-equity and participatory decision-making (Hurley & Hult, 1998); d) measures of leadership (Van Veldhuizen & Water-Bayer, 1997) and e) resource mobilisation (Van Veldhuizen & Water-Bayer, 1997).

Proposed indicators to assess C2I

The conceptual outline of C2I and its component capacities and indicators developed in this paper are based on the review of the literature on C2I across all sectors. In this section, we apply this to the agricultural context, proposing an exploratory framework and set of indicators for assessing C2I directly in the context of agricultural communities.

Approaching the assessment of C2I: what makes a 'good indicator'?

When developing a method of assessment that can be used by researchers as well as project implementers seeking to measure C2I in a community-based setting, we propose that the following principles be applied:

1. Where possible, use validated indicators for a specific capacity before creating new ones (e.g. tested methods of assessing individual and collective efficacy already exist);
2. When choosing among indicators, prioritise those which can be readily measured;
3. To avoid wasting resources on unnecessary measurement, use as few indicators as possible – 'bellwether indicators' rather than complex sets of interacting factors.

Regarding the latter, to ascertain which indicators may be most suitable as 'bellwether indicators', it is useful to return to the idea of 'higher-level' and 'building-block' capacities described above. For example, if we find that a group of farmers has engaged in a series of experiments resulting in improved practices or prototypes over time (demonstrating iterative learning), we can assume that at least some of the precursor or underpinning sub-capacities (e.g. to identify opportunities for learning, to devise experiments and test different approaches, to analyse results of experiments and trials, and to reflect and learn from results) are present. If, however, we assess this group of farmers at the level of various sub-capacities—the capacity to devise experiments or analyse trade-offs emerging from experiment results, say—we may or may not find that this results in the higher-level capacity to conduct iterative experimentation. In complex, adaptive systems such as AIS, higher-level capacities—including the capacity to innovate itself—are emergent properties of systems dynamics and do not reliably or predictably emerge when only some lower-level system conditions are

present. In developing metrics to assess C2I, we therefore propose an approach that focuses on defining indicators for the highest-level capabilities that are needed in order to produce and sustain innovation.

That said, we note that while bellwether indicators may be helpful to track aspects of C2I over time or space, in some cases it may be necessary to further unpack the C2I 'black box'. For example where there is a lack of innovation and C2I at these higher-level capacities, understanding the development of some of the building block capacities may be necessary to ascertain obstacles and change practices accordingly.

In developing indicators of C2I, we should also be mindful of the final objectives of development projects such as those aiming to develop C2I. Collecting data on, or including indicators of, innovation and of development outcomes is important in order to understand whether innovation has actually taken place, and whether that innovation has been accompanied by improvements (or not) in well-being. Assessing C2I allows us to understand how and by what processes these interventions have worked (and to continually develop and adjust interventions), while assessing development outcomes allows us to understand (eventually) if these interventions have worked. It may also be useful for local stakeholders themselves to reflect on their own C2I.

Defining components of C2I

The first task in developing metrics to assess capacity to innovate is therefore to define the core, high-level capacities within the C2I concept. We propose the following, which we consider as vital to C2I at the local community and local system levels:

1. **Creative drive and innovativeness**
2. **Networking and leveraging of linkages to access resources**
3. **Iterative experimentation**
4. **Collaborative action**

Selecting indicators

We suggest the following indicators in order to assess C2I. They reflect elements of the four core capacities as represented by, and to be measured at, the individual, community and local system levels.

At the individual level:

1. **Confidence** in ability to develop new and useful solutions to challenges and/or to experiment with and create new ways of doing things.
2. **Increased skills and abilities** associated with C2I at the individual level.
3. **Quantity and quality of experimentation,** e.g. more/better experiments or more/more diverse experimenters.

At the group/community level:

4. **Quality and effectiveness of stakeholders**: e.g. number/diversity of stakeholders, quality and effectiveness of their engagement
5. **Existence or growth in numbers of groups** or other organisations/community institutions with an innovation-related role

6. **Increased collective efficacy:** increases in strength and performance of these groups, such as a) increased confidence in the ability of groups to achieve objectives; b) improved inter-group dynamics; c) number of successful collective actions achieved by the group
7. **Quantity and quality of innovation output**[4]: more/better/more widely used innovations produced by groups or networks of people and individuals
8. **Increased and strengthened linkages:** network size, strength, effectiveness in sharing resources/support.

At the local system level:

9. **Strengthened enabling environment:** e.g. elements of the enabling environment are strengthened or added to; local people are better at sustaining the various supportive elements of the local enabling environment.
10. **Changes in norms, attitudes, policies, rules, funding/resource availability** that reduced barriers to innovation and/or facilitated the ability of local people to advance innovation processes.

These indicators reflect a selection of what we consider to be the essential components of C2I, but we stress that these are not yet field tested. Some of these indicators specifically address 'higher-level' capacities (e.g. 3, 4, 6) while others are indicators of 'building-block' capacities that we consider to be critical and measurable capacities (1, 2, 5, 8, 9, 10). Finally, we also include measures of innovation outcomes (7). In addition to these, it may be important to add indicators directly measuring the **processes** which contribute to bringing about innovation and supporting ongoing C2I, as well as the presence, strength and effectiveness of certain **structures and conditions** that enable innovation to take place.

Discussion

In this paper we have used a focused literature review to develop a list of categories representing core elements of C2I. We suggest that these elements are a first step to developing monitoring and evaluation tools for measuring C2I, and then make steps towards describing indicators. We acknowledge that, as with many literature reviews based on the use of keyword searches, there are limitations to this approach. While we aimed to use a range of terms around this concept, and followed up on references through a snowball search, we are assuming that the finite number of terms searched are linked to the complex phenomena of C2I. In reality, these concepts are also discussed using different language, particularly in the wider social sciences and psychology literature. Thus, the approach taken here is likely to be a fair representation of the C2I concept amongst those using C2I terminology, particularly amongst those in the agricultural development field, but may not fully explore other interpretations of the concept, and so we should be wary of applying these conclusions and tools to other fields. Another limitation of this approach is that have restricted our search to literature which is online and relatively accessible. While we tried to search for practitioner publications as well as peer-reviewed publications, the former are not always as well-archived and bibliographic databases may be skewed towards the latter.

[4] *Importantly this occurs and could be measured at the level of the individual and the group. i.e. innovations may have been spearheaded by an individual 'innovator' or by group activity*

Having accepted the limitations to our approach, we can still make some valid conclusions. The importance of catalysing and strengthening innovation for development outcomes is now accepted (FAO, 2014). The dominant approach to this has been to focus on the development of new technologies through agricultural research, the adoption of which is assumed to generate outcomes and impact. The value of the C2I concept is that it puts emphasis on the causal power of the process component of innovation (which is often over-looked), rather than on the artefact (technology) component. Research processes can build all of the capacities to innovate, and yet some don't, or are not explicit about wanting to. C2I has the potential to increase the impact of agricultural research, in particular for more marginalised people for whom connectivity and capacity is more of an issue than available technology.

While some CD projects are moving towards developing capacities for innovation (e.g. the CD AIS project (TAP, 2016)), we argue that using the C2I concept bundles together a group of important capacities that conventional CD projects may overlook. Focus on the technology component of innovation has meant that conventional CD projects concentrate on scientific or technical capacities (TAP, 2016). These, while integral to innovation, are alone not sufficient to drive it (Dijkman, 2010). Instead C2I approaches emphasise transferable skills such as those needed to learn and access knowledge (Pant, 2012) and to combine research-based knowledge with context-specific knowledge (often 'tacit knowledge'[5] that may not be written down) (Chuluunbaatar & LeGrand, 2015) facilitating the adaptation of innovations to local settings (Hall et al., 2009).

Projects focusing on, or recognising, C2I may also put more emphasis on cross-dimensional interventions, including improving capacities at different scales (TAP 2016) which ensures a more cohesive approach. In contrast, traditional CD interventions may fail to capture the full complexity of innovation processes (Aerni, 2013; TAP, 2016). Similarly, C2I approaches stress the importance of networking and participation (TAP, 2016), while traditional approaches may fail to strengthen inter-relational capacities (Gottret & Córdoba, 2004).

Key to making process outcomes of agricultural research more visible is rigorously showing that they exist and demonstrating their value. This requires measurement, and this paper has made steps towards developing a measurement system. The indicators developed here are targeted, measurable indicators linking to specific capacities to innovate. The next step will be testing and refining these indicators in the field, in order to develop a robust monitoring and evaluation tool, a work that is currently being done by the authors in a number of case studies.

Conclusion and Perspectives

The literature shows that although the emerging C2I concept is not concretely defined, most scholars now agree that it involves the continuing ability to access or generate innovations and to successfully apply them. There are the beginnings of a consensus over what component capacities the term encompasses, with most focusing on the capacities to generate or access innovations through networks, test and adapt innovations, and work with others to apply and adopt them. Accurate indicators linked to C2I capacities, rather than innovation outcomes, will allow us to assess the efficacy of different intervention types for different capacities. Testing the indicators proposed here may allow us to improve interventions for greater C2I and thus improve development outcomes.

[5] *Knowledge based on experience in a specific situation and which is less likely to be codified*

References

Adner, R. (2006). Match your innovation strategy to your innovation ecosystem. Harvard Business Review 84(4): 98.

Adner, R., & Kapoor, R. (2010). Value creation in innovation ecosystems: how the structure of technological interdependence affects firm performance in new technology generations. Strategic Management Journal 31(3): 306-333.

Aerni, P. (2013). Assessment of current capacities and needs for capacity development in agricultural innovation systems in low income tropical countries. Synthesis report. . Rome, Italy: Tropical Agriculture Platform (TAP).

Aiken, M., Bacharach, S.B., & French, J.L. (1980). Organisational structure, work processes and proposal making in administrative bureaucracies. Academy of Management Journa, 23: 631-652.

Assefa, A., Waters-Bayer, A., Fincham, R., & Mudahara, M. (2009). Comparison of frameworks for studying grassroots innovation: agricultural innovation systems (AIS) and agricultural knowledge and innovation systems (AKIS). In P. C. Sanginga, A. Waters-Bayer, S. Kaaria, J. Njuki and C. Wettasinha (Eds.) Innovation in Africa: Enriching Farmers and Livelilhoods. UK: Routledge.

Bell, M., & Albu, M. (1999). Knowledge systems and technical dynamism in industrialised clusters in developing countries. Industrial and Corporate Change 2(2): 157-210.

Berkes, F. (2007). Understanding uncertainty and reducing vulnerability: lessons from resilience thinking. Natural Hazards 41(2): 283-295.

Boly, V., Morel, L., Assielou, N.G., & Camargo, M. (2014). Evaluating innovative processes in French firms: methodological proposition for firm innovation capacity evaluation. Research Policy 43(3): 608-622.

Brusoni, S., & Prencipe, A. (2013). The organisation of innovation in ecosystems: problem framing, problem solving and patterns of coupling. In R. Adner, J. E. Oxley and B. S. Silverman (Eds.) Collaboration and Competition in Business Ecosystems (Vol. 30, pp. 167-194).

Burns, T., & Stalker, G.M. (1961). The Management of Innovation. London, UK: Tavistock Publishing.

Caccia-Bava, M.D.C., Guimaraes, T., & Harrington, S.J. (2006). Hospital organisation culture, capacity to innovate and success in technology adoption. Journal of Health, Organisation and Management 20(3): 194-217.

Carlsson, L., & Sandström, A. (2008). Network governance of the commons. International Journal of the Commons 2(1): 33-54.

Chataway, J., Smith, J., & Wield, D. (2005). Partnerships for building science and technology capacity in Africa: Canadian and UK experience paper prepared for the Africa-Canada-UK

Exploration: Building Science and Technology Capacity with African Partners, 30th January - 1st February 2005. Canada House, London: UK.

Chatenier, E.d., Verstegen, J.A.A.M., Biemans, H.J.A., Mulder, M., & Omta, O.S.W. (2010). Identification of competencies for professionals in open innovation teams. R&D Management 40(3): 271-280.

Chuluunbaatar, D., & LeGrand, S. (2015). Enabling the capacity to innovate with a system-wide assessment process. Occasional Papers on Innovation in Family Farming. Rome, Italy: Food and Agriculture Organisation of the United Nations.

Clark, L. (2006). Building farmers' capacity for networking (part ii): strengthening agricultural supply chains in Bolivia using network analysis. KM4D Journal 2(2): 19-32.

Clark, N. (2005). Science policy and agricultural research in Africa: a capacity building needs assessment: NEPAD.

Cohen, W.M., & Levinthal, D.A. (1990). Absorptive capacity: a new perspective on learning and innovation. Administrative Science Quarterly 35: 128-152.

Daane, J., Francis, J., Oliveros, O., & Bolo, M. (2009). Performance indicators for agricultural innovation systems in the acp region. Paper presented at the International Expert Consultation Workshop. Synthesis Report, CTA Headquarters, Wageningen, The Netherlands. 15-17 July 2008.

Dalohoun, D.N. (2005). Learning to innovate. Agricultural innovation and entrepreneurship: the case of songhaï farmers in Benin. United Nations University, Maastricht, The Netherlands.

DFID. (2008). Capacity building DFID Research Strategy 2008-2013. Working Paper Series: Capacity Building: Research for Development. Department for International Development, UK.

Dijkman, J. (2010). Africa matters. Emergent policy-relevant lessons from the RIU Africa country programmes Discussion Paper 9. London, UK: Research Into Use (RIU) Programme, Department for International Development (DFID).

Douthwaite, B. (2002). How to enable innovation. Paper presented at the Agricultural Engineering International the CIGR. Journal of Scientific Research and Development. Invited Overview Paper.

Douthwaite, B., Beaulieu, N., Lundy, M., & Peters, D. (2009). Understanding how participatory approaches foster innovation. International Journal of Agricultural Sustainability 7(1): 42-60.

Douthwaite, B., & Gummert, M. (2010). Learning selection revisited: how can agricultural researchers make a difference? Agricultural Systems 103(5): 245-255.

FAO. (2014). State of food and agriculture 2014. Innovation in family farming. Rome, Italy: Food and Agricultural Organization of the United Nations.

FARA. (2014). FARA's 2014-2018 strategic plan: Enhancing African innovation capacity for agricultural transformation. Accra, Ghana: Forum for Agricultural Research in Africa

Fernandez-Gimenez, M.E., Ballard, H.L., & Sturtevant, V.E. (2008). Adapative management and social learning in collaborative and community-based monitoring: A study of five community-based forestry organisations in the western USA. Ecology and Society 13(2): 4.

Freeman, C. (1987). Technology and economic performance: lessons from Japan. London, UK: Pinter.

Furman, J.L., Porter, M.E., & Stern, S. (2002). The determinants of national innovative capacity. Research Policy 31: 899-933.

Gottret, M.V., & Córdoba, D. (2004). Políticas y procesos de innovación tecnológica con productores de pequeña escala en honduras y nicaragua. Cali, Colombia: Centro Internacional de Agricultural Tropical.

Grogger, J., & Hanson, G.H. (2011). Attracting talent: location choices of foreign-born PhDs in the US (IGC working paper) (pp. 39). London, UK: International Growth Centre (IGC).

Hall, A. (2005). Capacity development for agricultural biotechnology in developing countries: an innovation systems view of what it is and how to develop it. Journal of International Development 17(5): 611-630.

Hall, A., Sulaiman, R., Beshah, T., Madzudzo, E., & Puskur, R. (2009). Agricultural innovation system capacity development: tools, principles or policies? Capacity.org, Issue 37: 16-17.

Hall, A., Sulaiman, R., & Bezkorowajny, J.P. (2007). Reframing technical change: livestock fodder scarcity revisited as innovation capacity scarcity. A conceptual framework. Nairobi, Kenya: ILRI and UNU/MERIT

Hambly, H., & Sarapura, S. (2009). Ensuring gender equality in capacity development - opportunities for rural employment and sustainable development. Paper presented at the FAO-IFAD-ILO Workshop: Gaps, Trends and Current Research in Gender Dimensions of Agricultural and Rural Employment: Differentiated Pathways out of Poverty, Rome, Italy. http://www.fao-ilo.org/fileadmin/user_upload/fao_ilo/pdf/Papers/20_March/Hambly_Sarapura-paper.pdf

Hartwich, F., Gottret, V., Babu, S., & Tola, J. (2007). Building public-private partnerships for agricultural innovation in latin america. Lessons from capacity strengthening. IFPRI Discussion Paper. Washington DC: International Food Policy Research Institute.

Hermans, F., Stuiver, M., Beers, P.J., & Kok, K. (2013). The distribution of roles and functions for upscaling and outscaling innovations in agricultural innovation systems. Agricultural Systems 115: 117-128.

Hosmer, L.T. (1995). Trust: the connecting link between organisational theory and philosophical ethics. The Academy of Management Review 20(2): 379-403.

Howard, J., & Gillies, A. (2009). Knowledge to innovate: developing a tool to assess and assist the development of the capacity to innovate in small and medium sized enterprises. 4th European conference on entrepreneurship and innovation: ECEI2009

Hueske, A.-K., Endrikat, J., & Guenther, E. (2015). External environment, the innovating organisation, and its individuals: a multilevel model for identifying innovation barriers accounting for social uncertainties. Journal of Engineering and Technology Management,35: 45-70.

Hult, G.T.M., Hurley, R.F., & Knight, G.A. (2004). Innovativeness: its antecedents and impact on business performance. Industrial Marketing Management 33(2004): 429-438.

Hurley, R.F., & Hult, G.T.M. (1998). Innovation, market orientation, and organisational learning: an integration and empirical examination. Journal of Marketing 62: 42-54.

Hurley, R.F., Hult, G.T.M., & Knight, G.A. (2005). Innovativeness and capacity to innovate in a complexity of firm-level relationships: a response to Woodside (2004). Industrial Marketing Management 34(3): 281-283.

Jang, S., Hong, K., Bock, G.W., & Kim, I. (2002). Knowledge management and process innovation: the knowledge transformation path in Samsung SDI. Journal of Knowledge Management 6: 479-485.

Johnson, B., & Segura-Bonilla, O. (2001). Innovation systems and developing countries: experiences from the Sudesca project. DRUID Working Paper No. 01-12: Danish Research Unit for Industrial Dynamics (DRUID).

Jones, M. (2004). Strengthening agricultural research in Africa. 2020 Africa Conference Brief 9. Washington DC: International Food Policy Research Institute.

King, N., & Anderson, N. (1990). Innovation in working groups. In M. A. West and J. L. Farr (Eds.) Innovation and Creativity at Work. UK: Wiley.

Klerkx, L., Aarts, N., & Leeuwis, C. (2010). Adaptive management in agricultural innovation systems: the interactions between innovation networks and their environment. Agricultural Systems 103(6): 390-400.

Klerkx, L., Hall, A., & Leeuwis, C. (2009). Strengthening agricultural innovation capacity: are innovation brokers the answer? International Journal of Agricultural Resources, Governance and Ecology 8(5-6): 409-438.

Klerkx, L., van Mierlo, B., & Leeuwis, C. (2012). Evolution of systems approaches to agricultural innovation: concepts, analysis and interventions. In I. Darnhofer, D. Gibbon and B. Dedieu (Eds.) Farming Systems Research into the 21st Century: The New Dynamic. Dordrecht: Springer Science & Business Media.

Lambrecht, E., Kühne, B., & Gellynck, X. (2014). How do innovation partners differ with respect to innovation type and stage in the innovation journey of farmers? International Journal of Entrepreneurship and Innovation 15(3): 191-203.

Leeuwis, C., Schut, M., Waters-Bayer, A., Mur, R., Atta-Krah, K., Douthwaite, B., . . . others. (2014). Capacity to innovate from a system CGIAR research program perspective. Penang, Malaysia: CGIAR Research Program on Aquatic Agricultural Systems (AAS) Program Brief AAS-2014-29.

Lybbert, T.J., & Sumner, D.A. (2012). Agricultural technologies for climate change in developing countries: policy options for innovation and technology diffusion. Food Policy 37(1): 114-123.

Mauerhofer, V. (2010). Social capital, capacity and carrying capacity: exploring basics of 'socially sustainable economic degrowth'. Paper presented at the 2nd conference on Economic Degrowth for Ecological Sustainability and Social Equity, Barcelona, Spain.

Mayne, J., & Douthwaite, B. (2015). Beyond problem solving: building and evaluating capacity to innovate. AAS Working Paper.

Mbabu, A.N., & Hall, A. (2012). Capacity building for agricultural research for development. United Nations University - Maastricht Economic and Social Research and Training centre on Innovation and Technology (UNU-MERIT). Maastricht: The Netherlands.

Michailova, S., & Husted, K. (2003). Knowledge-sharing hostility in russian firms. California Management Review 45; 59-77.

Musiolik, J., Markard, J., & Hekkert, M. (2012). Networks and network resources in technological innovation systems: towards a conceptual framework for system building. Technological Forecasting and Social Change 79(6): 1032-1048.

Mytelka, L.K. (2000). Local systems of innovation in a globalised world economy. Industry and Innovation 7(1): 15-32.

Nair, H.A.P., Kumar, D., & Ramalu, S.S. (2014). Assessment framework for innovation capacity. Man in India 94(4): 775-795.

Neely, A., & Hii, J. (2001). Innovative capacity of firms in east of England: actions and performance: Centre for Business Performance Cranfield School of Management.

Nettle, R., Brightling, P., & Hope, A. (2013). How programme teams progress agricultural innovation in the Australian dairy industry. Journal of Agricultural Education and Extension 19(3): 271-290.

OECD. (2012). Innovation for development. A discussion of the issues and an overview of work of the OECD directorate for science, technology and industry: Organisation for Economic Co-operation and Development (OECD).

Pant, L.P. (2012). Learning and innovation competence in agricultural and rural development. Journal of Agricultural Education and Extension 18(3): 205-230.
Pant, L.P., & Hambly Odame, H. (2009). Innovation systems in renewable natural resource management and sustainable agriculture: a literature review. African Journal of Science Technology Innovation & Development 1: 103-135.

Pound, B., & Essegby, G. (2008). Agricultural innovation systems SCARDA Briefing Papers (Vol. 3, pp. 46-53). Forum for Agricultural Research in Africa (FARA). Accra: Ghana.

Rohrbeck, R., Hölzle, K., & Gemünden, H.G. (2009). Opening up for competitive advantage - how deutsche telekom creates an open innovation ecosystem. R and D Management 39(4): 420-430.

Röling, N.G. (2009). Pathways for impact: scientists' different perspectives on agricultural innovation. International Journal of Agricultural Sustainability 7(2): 83-94.

Rufat-Latre, J., Muller, A., & Jones, D. (2010). Delivering on the promise of open innovation. Strategy & Leadership 38(6): 23-28.

Schut, M., Klerkx, L., Rodenburg, J., Kayeke, J., Hinnou, L.C., Raboanarielina, C.M., . . . Bastiaans, L. (2015). RAAIS: Rapid Appraisal of Agricultural Innovation Systems (part i). A diagnostic tool for integrated analysis of complex problems and innovation capacity. Agricultural Systems 132: 1-11.

Sen, A.K. (1993). Capability and well-being. In M. C. Nussbaum and A. K. Sen (Eds.) The Quality of Life (pp. 30-53). Oxford, UK: Clarendon Press.

Senge, P.M. (1990). The fifth discipline. The art and practice of the learning organisation. New York, USA: Doubleday.

Skiltere, D., & Jesilevska, S. (2013). Building the system of innovation capability indicators: case of Latvia. Eurasian Journal of Business and Economics 6(12): 113-128.

Smart, P., Bessant, J., & Gupta, A. (2007). Towards technological rules for designing innovation networks: a dynamic capabilities view. International Journal of Operations & Production Management 27(9-10): 1069-1092.

Song, Y., Cheng, Y.X., & Meng, L.S. (2014). DEA evaluation of strategic and emerging industries in Hubei province of China based on endogenous competitiveness.

Spielman, D., Davis, K., Negash, M., & Ayele, G. (2011). Rural innovation systems and networks: findings from a study of Ethiopian smallholders. Agriculture and Human Values 28: 195-212.

Spielman, D.J., & Kelemework, D. (2009). Measuring agricultural innovation system properties and performance: illustrations from Ethiopia and Vietnam. IFPRI DIscussion Paper No. 851. Washington DC, USA: International Food Policy Research Institute.

TAP. (2016). Common framework on capacity development for agricultural innovation systems. Guidance note on operationalisation. Draft. Rome, Italy: Tropical Agriculture Platform (TAP) project "Capacity Development for Agricultural Innovation Systems (CDAIS), jointly implemented by Agrinatura-EEIG and FAO.
Temel, T. (2004). Mapping organisational linkages in the agricultural innovation system of Azerbaijan. International Journal of Agricultural Resources, Governance and Ecology 3: 134-152.

The Royal Academy of Engineering. (2012). Educating Engineers to Drive the Innovation Economy. London, UK.

Thomas, G., & Slater, R. (2006). Innovation, agricultural growth and poverty reduction. International Journal of Technology and Globalisation 2(3/4): 279-287.

Traitler, H., Watzke, H.J., & Saguy, I.S. (2011). Reinventing R&D in an open innovation ecosystem. Journal of Food Science 76(2): R62-R68.

Turner, J.A., Klerkx, L., White, T., Payne, T., Everett-Hincks, J., Mackay, A., & Botha, N. (2015). Unpacking systemic capacity to innovate: how projects coordinate capabilities across agricultural innovation system levels.

Ugbe, U.P. (2010). It May Take a Little While... Insights on Agricultural Research for Innovation and Development in Nigeria. pdf dfid.gov.uk

UNDP. (2009). Capacity development: a UNDP primer. New York, USA: United Nations Development Programme, Bureau for Development Policy. Capacity Development Group.

Van Veldhuizen, L., & Water-Bayer, A. (1997). Developing technology with farmers: a trainers' guide for participatory learning. New York, USA: ZED Books Ltd.

Wang, C.L., & Ahmed, P.K. (2007). Dynamic capabilities: a review and research agenda. [Review]. International Journal of Management Reviews 9(1): 31-51.

Wonglimpiyarat, J. (2010). Innovation index and the innovative capacity of nations. Futures 42: 247-253.

Woodside, A.G. (2005). Firm orientations, innovativeness and business performance: advancing a system dynamics view following a comment on Hult, Hurley, and Knight's 2004 study. Industrial Marketing Management 34(2005): 275-279.

World Bank. (2007). Enhancing agricultural innovation: how to go beyond the strengthening of research systems. Washington DC.

World Bank. (2008). World development report 2008: agriculture for development. In World Bank (Ed.) Washington DC, USA.

World Bank. (2012). Agricultural Innovation Systems: an Investment Source Book.

Yang, Y., & Konrad, A.M. (2011). Diversity and organisational innovation: the role of employee involvement. Journal of Organisational Behavior 32: 1062-1083.

Zahra, S.A., & George, G. (2002). Absorptive capacity: a review, reconceptualisation and extension. Academy of Management Review 27(2): 185-203.
Zhang, F., & Wu, F. (2012). "Fostering indigenous innovation capacities": the development of biotechnology in Shanghai's Zhangjiang high-tech park. Urban Geography 33(5): 728-755.

Evaluating for learning and accountability in system innovation: incorporating reflexivity in a logical framework

Botha, N.[1], Coutts, J.[2], Turner, J.A.[1], White, T.[1] and Williams, T.[3]

[1] *Ruakura Research Centre, New Zealand*
[2] *Coutts J&R Queensland, Australia*
[3] *The New Zealand Institute for Plant & Food Research Limited*

Abstract: Approaches to accelerate innovation such as Agricultural Innovation Systems and co-innovation (Brunori et al., 2008; Knickel et al., 2009; Fischer et al., 2012) have become more integrated and systemic over time. Primary Innovation is a New Zealand co-innovation programme in which innovation is conceived as being 'co-produced' by stakeholders who contribute their unique knowledge to solving a problem or realising an opportunity. In co-innovation, cyclical processes of planning, doing, observing and reflecting enable innovation to emerge from interactive learning among stakeholders (Botha et al., 2014). The value of applying logic models (logical frameworks, programme theories or theories of change) and concurrently evaluating the effects of co-innovation practices (particularly reflexive processes) in order to understand the extent of learning in and impact from systemic projects have been questioned and debated (Klerkx et al., 2012; Regeer et al., 2016). In this paper we argue that, when flexibly applied and adapted to capture dynamics typical in systems innovation projects, the Log Frame Approach (LFA) (Gaspar, 1999; AusAid, 2005; Kaplan, 2015) and logical frameworks (Kaplan, 2015) have considerable utility to support evaluation for both learning and accountability, and for identifying and addressing institutional logics which leads to system innovation. We demonstrate this for the case of Primary Innovation, and compare our experiences with the limitations and solutions suggested by Regeer et al. (2016) when applying logic models, logical frameworks, programme theories or theories of change as part of an `adapted accountability framework'.

Keywords: Log frame approach, reflexivity, co-innovation, system innovation, impact, accountability, learning

Introduction

Expectations of the impact from public and private investments in innovation are increasing as commissioners seek reassurance that innovation programmes are successfully addressing the complex problems challenging global primary sectors (Coutts et al*.,* 2014). As a consequence, a more systemic and integrated perspective of innovation is emerging in the form of Agricultural Innovation Systems theory and practice.

Agricultural Innovation Systems are complex adaptive systems characterised by a "large number of actors, diverse interactions and relationships and constantly changing influences emerging from technological, market, policy, cultural and other socioeconomic factors'' (Spielman et al*.,* 2009). From these complex adaptive systems, innovations emerge through a co-evolutionary process that combines technological, social, economic and institutional change (Klerkx et al., 2012). While such co-evolutionary processes generally take place

autonomously (see Ekboir, 2003), the principles behind these processes can also inform approaches designed to accelerate innovation (see for example Nederlof et al., 2011 and Nettle et al., 2012). This implies a move from linear, technology transfer (technologies invented by science and disseminated by extension) to a wider system innovation (Fischer et al., 2012) involving interactive learning among all relevant actors in the agricultural sector, including farmers, growers, consultants, banks, agri-businesses, Government, NGOs and entrepreneurs. The key characteristics of system innovation projects are (Klerkx et al., 2010; Van Mierlo et al., 2010a):

1. Numerous stakeholders with multiple and often conflicting goals;
2. A focus on reflection, learning and action;
3. Co-evolution of technical, social, market and institutional changes; and
4. Emergent outcomes that are modified in response to changes in system understanding and external system changes.

In innovation projects these systemic aspects of system innovation create new challenges for traditional monitoring and evaluation methods and tools such as logical frameworks, which were historically used to meet the accountability needs of commissioners.

Evaluation for learning and accountability in a system innovation programme: Primary Innovation

Primary Innovation seeks to explore ways actors in the New Zealand primary industries can work together to jointly learn and co-develop innovations to complex industry challenges, such as water and land management (Botha et al., 2014). The programme also has wider system innovation ambitions to stimulate changes in the New Zealand Agricultural Innovation System so that adequate and complementary innovation policies, funding frameworks and organisational cultures enable the optimal performance of innovation networks (Turner et al., 2016).

Achieving system innovation requires reflexivity, learning and action among multiple actors in the Agricultural Innovation System (Van Mierlo et al., 2010a; Van Mierlo et al., 2010b) in order to tackle the key institutional logics that hinder the formation and functioning of innovation networks (Turner et al., 2016). Institutional logics are historically built-up and persistent structures and institutional arrangements that lock systems into current arrangements (Fuenfschilling & Truffer, 2014). Learning needs to stimulate structural (or system) change by participants in the system. This kind of learning is facilitated by a reflexive perspective, enabling things that are usually taken for granted to be challenged (Loeber et al., 2007; Van Mierlo et al., 2010a; Van Mierlo et al., 2010b). The outcome of reflexivity and learning in system innovation projects is practical action in the wider system and subsequent double loop learning from this action (Van Mierlo et al., 2010b).

In system innovation projects such as Primary Innovation, monitoring and evaluation are not separate, but become an integral part of reflexivity and learning (Van Mierlo et al., 2010b). To be effective in this highly adaptive setting evaluation must be flexible to respond to (i) learning from action and (ii) external system changes, especially to seize ‘windows of opportunity’ (Van Mierlo et al., 2010a; Beers et al., 2014). This places evaluation in a constructivist perspective that recognises the importance of multiple perspectives and values in shaping system innovation (Arkesteijn et al., 2015). Evaluation also needs to deal with ambiguity and disagreement and their influence on lock-in in systems due to existing institutional logics (Van Mierlo et al., 2010a).

At the same time, system innovation projects are also accountable to commissioners, who often have in place formal procedures for evaluating the efficiency and efficacy of project funding (Botha et al., 2014; Gosling & Edwards 2003b; Roberts & Coutts 2011). A common approach for achieving this is the use of classical project planning and evaluation such as the Logical Framework Approach (LFA) which has been characterised as a tool (Gaspar, 1999), methodology (AusAid, 2005) or approach (Kaplan, 2015) for designing, executing and assessing projects and programmemes (AusAid, 2005). This approach was originally developed from an instrumental perspective, with the aim of meeting the accountability needs of project funders – upwards accountability (Gosling & Edwards 2003; Regeer et al., 2016). In this paper Logical Framework Matrix (LFM) (AusAid, 2005) or logical framework (Kaplan, 2015) refers to a matrix with columns and a number of rows that show inputs, outputs, short-, medium and long term outcomes. Its purpose is to translate a Log Frame Approach into action, and as a document it forms the basis of an actionable work plan that guides implementation through the programme lifecycle (Kaplan, 2015).

Applying logical frameworks - as tools for planning - to achieve upwards accountability has been based on rational planning and problem solving in which causes and effects are assumed to be predictable ex-ante and uncertainty reduced and managed (Arkesteijn et al., 2015). As has previously been observed, the use of such classical project evaluation creates tensions with the focus on learning that characterises system innovation projects (Botha et al., 2014; Regeer et al., 2016). This has led to five criticisms regarding the use of the Logical Framework Approach or log frame (Arkesteijn et al., 2015) in system innovation projects in which reflexivity, learning and action are essential:

1. It fails to recognise complexity and inherent uncertainty in system change
2. It focuses on what is agreed in formulating the project and does not make transparent the points of disagreement or conflict in system innovation
3. It assumes that society is amenable to rational design
4. Accountability outcomes tend to take priority over learning outcomes
5. Accountability to other stakeholders, besides funders, is ignored.

These criticisms raise several issues regarding the practical application of logical frameworks to support reflexivity and learning in system innovation projects while also meeting the upwards accountability needs of funders (Regeer et al., 2016).

Regeer et al., (2016) outline eight practical limitations to the application of logic models, logical frameworks, programme theories or theories of change in system innovation projects:

1. They cannot capture the diverse dimensions of accountability and learning because they are typically separated and not dealt with simultaneously in innovation (niche) projects
2. They are developed ex-ante to describe the expected causal relations between inputs, activities and desired programme outcomes
3. They are often used ex-post to assess whether, and to what extent, programme goals and objectives have been achieved – upwards accountability
4. They typically see accountability in terms of predefined goals and relationships with interventions that “presuppose a relatively stable programme, whose activities, goals and intended effects can be univocally described.” (Regeer et al., 2016)

5. They adopt a goal-oriented evaluation that does “not sufficiently take into account the emergent nature of complex projects and their multifaceted environment” (Regeer et al., 2016)
6. They fail to capture the effects of external and internal pressures on projects, which can affect the ambition for change, and do not reflect back these pressures as potential for structural change
7. They typically are not used during intervention processes nor include all stakeholders
8. They are commonly implemented by external evaluators and, depending on the commissioner, results may be publicly available. As a result, participants in the evaluation may be more inclined to defend the outcomes generated by their actions and decisions rather than be willing to internalise the findings and learn from them.

Recent work (Botha et al., 2014; Regeer et al., 2016) has called for the need to explore ways to reconcile these practical issues in meeting accountability needs from evaluation, with the reflexivity, learning and action needs of system innovation projects. This has stimulated modifications of logical frameworks to increase their flexibility (e.g. MERI (Monitoring, Evaluation, Reporting and Improvement)) (Dart, 2007) and, to some extent, the Theory of Change (Funnell & Rogers 2011). Regeer et al., (2016) argue for a reframing of accountability to not only funders, but all stakeholders in system innovation, with the implication that evaluation is undertaken to address these multiple accountabilities. Botha et al. (2014) call for ‘learning by doing’ to operationalise the use of logical frameworks for accountability, reflexivity, learning and action in a system innovation project. In this paper we present the insights gained from this ‘learning by doing’ in the Primary Innovation programme. Botha et al. (2015a) discussed the challenges when using a logical framework (Kaplan, 2015) in co-innovation projects and concluded that it worked well in Primary Innovation, was particularly useful in providing a monitoring and evaluation framework and guiding change, and matched a co-innovation approach to fostering change.

We argue that the practical limitations of the use of a logical framework to support both learning and accountability in system innovation projects, such as Primary Innovation, can be overcome by the way it is implemented during a project. We demonstrate this for the case of Primary Innovation and compare our experiences with the solutions suggested by Regeer et al. (2016) to the eight limitations listed above.

The paper is organised as follows: the next section provides a description of how a logical framework was implemented in Primary Innovation. The following section then describes how the programme addressed each of the eight practical limitations to using a logical framework in system innovation (Regeer et al., 2016). We conclude the paper with a discussion of the main insights on how to use a logical framework to simultaneously meet accountability, reflexivity, learning and action needs in system innovation projects.

Methods used to implement a logical framework in Primary Innovation

At the start of Primary Innovation, while there was no requirement to develop or use a logical framework, there was a recognition that it was a complex programme that required effective evaluation. To this end, the planned steps, process and aspirational outcomes needed to be defined and understood by the programme team (researchers and social scientists across participating research organisations) and endorsed by its 23 stakeholders (Botha et al., 2014). The large number and diversity of Primary Innovation stakeholders makes it impractical for everyone to be involved in the logical framework and monitoring and evaluation throughout

the life of the investment. While all Primary Innovation's stakeholders are familiar with the logical framework, a small team consisting of individuals from AgResearch and Plant and Food Research, as well as Coutts J&R (a contracted evaluator), uses it to review the programme's logic, expected outputs and theory of action, and evaluate impacts and processes rather than checking whether a predetermined path is unquestioningly pursuing contract milestones. The small team shares and discusses their findings and activities with the other 23 stakeholders through a newsletter, quarterly reports and other communication methods like phone calls, meetings, webinars and workshops.

The outcomes from Primary Innovation are emergent because system innovations have multiple socio-technical components that cannot be defined in advance (Wiskerke & van der Ploeg, 2004). However, to secure funding Primary Innovation was required to identify expected outcomes in the funding application and contract. These were used to help develop the first logical framework.

The levels used in the logical framework (Coutts et al., 2014) were : Longer term Outcomes – towards which the programme is intended to contribute along with other complementary initiatives; Key Result Areas – specific measurable short term impacts or achievements to which the programme is planning to deliver on in its life (including unintended benefits or consequences); Uptake Strategies – approaches used to reflect, communicate, influence, assist and/or encourage appropriate people or groups to effectively engage; Underpinning Activities – Research, Development & Planning Activities and Outputs needed or used (from other sources) to provide the science, tools, information or materials to support systems change processes; Supporting Structures – resources, staff, management processes, Steering Groups and other structures to oversee and undertake programme activities; and Context – political, economic, climatic and other factors that can affect the success or otherwise of the programme and process.

The first logical framework was developed by the programme team with the assistance of Coutts J&R who were embedded in the monitoring and ex-durante evaluation of Primary Innovation. This first draft was circulated for feedback to the programme team and the content was developed in a workshop where they created a flow diagram to develop a greater understanding and sense of ownership. This provided a focus for reflecting, discussing, understanding and refining the programme and its change ambition, and highlighted the key elements for evaluation in order to generate accountability and enable learning. The logical framework was revisited annually over the next two years as a basis for reviewing the programme and highlighting gaps in monitoring and evaluation data.

Three workshops were held with a monitoring and evaluation focus – all based around the logical framework. For example, the flow diagram developed in the first workshop was used at subsequent workshops to focus discussion, reflect on progress (including the ambition for change), analyse monitoring and evaluation data captured and look for gaps. Re-visiting the logical framework also provided an opportunity for ongoing reflection and discussion of what co-innovation meant in practice and how best to evaluate it. Following the workshops, participants were provided with structured feedback sheets and asked to reflect on their learning about and understanding of monitoring and evaluation as a result of participating in the development of the logical framework.

A team member was appointed to support the evaluation process and collect data to support the utility of the logical framework. Amongst other responsibilities, a reflexive monitor also kept track of how team members were responding to changes in the wider Agricultural Innovation System that were affecting Primary Innovation, such as the establishment of new platforms for innovation (e.g. the National Science Challenges), as an opportunity for furthering programme ambitions (Hekkert et al., 2007; Bussels et al., 2013; Rijswijk et al., 2015). A reflexive monitor's role is to help programme participants reflect on process, action and progress towards agreed research goals (Van Mierlo et al., 2010a). It is a mechanism that the Primary Innovation team is using to help identify opportunities where co-innovation can enhance impact in the design or management of a series of innovation projects (case studies). Reflexive monitoring is also being used by the research team and stakeholders to remind them of their ambitions for system innovation, such as by challenging and changing presumptions, current systemic practices and underlying institutions (Botha, 2013). This is different from the role of a facilitator, in that reflexive monitors challenge, as well as support, participants to reflect on and address how the way they work together enhances or hampers progress towards their shared ambition for change (Botha et al., 2014).

Narrative reflective processes during the workshops helped participants to reflect upon their own and others' learning (Schwind et al., 2012). These were recorded in a Dynamic Learning Agenda (van Veen et al., 2014).

The logical framework was also used for a mid-project evaluation during which available monitoring and evaluation data were analysed against the accountability performance measures at every level of the logical framework. This highlighted data gaps and issues that needed to be considered by the programme team and stakeholders, providing another opportunity to reflect on the extent of learning underway in the programme as well as increase accountability around programme performance.

Results

This section provides a description of ways the logical framework was applied in the Primary Innovation programme that may address each of the eight practical limitations, identified by Regeer et al., (2016), to using such a framework in system innovation. Lessons number two and three have been combined as they are closely related.

Logical frameworks do not deal with accountability and learning simultaneously

By using the logical framework in Primary Innovation in a flexible and reflective way, tensions that surfaced around learning across the programme were identified quickly and addressed. For example, tensions arose during a programme workshop about the speed of progress in the programme between the management team, researchers and practitioners when discussing and reflecting upon processes, outputs and learning. Learning about and understanding co-innovation was more important to researchers than practitioners who preferred learning-by-doing, while programme delivery and impact was important for upward (funder) accountability of the programme management team. These tensions were dealt with through regular telephone and face-to-face conversations between programme participants and a reflexive monitor for the programme who helped discussions to focus on the shared ambition for change. This has resulted in increased involvement of leaders of Primary Innovation case studies in reflexive sessions as well as increasing their involvement in writing up the results of the research.

Logical frameworks are developed ex-ante and applied ex-post

In Primary Innovation the logical framework was for practical reasons constructed and reflectively used ex-durante in order to re-visit and discuss the underpinning 'theory of action' of the programme and causalities assumed in the original project proposal. The ex-post evaluation is yet to be completed as the programme is still underway. However, a mid-term review was undertaken using the logical framework. This review included consideration of changes in the logical framework itself.

Logical frameworks typically see accountability in terms of predefined goals and relationships

The use of monitoring and evaluation with a logical framework has, in the case of Primary Innovation, been used to identify where changes are needed. One example is the membership and functioning of the Community of Practice – a mechanism through which Primary Innovation intends to 'scale up' co-innovation across New Zealand (that is, to influence and stimulate system innovation at the Agricultural Innovation System level) (Botha et al., 2014). Over time, through reflection, considering member feedback, an openness towards adaptation and by being pragmatic and flexible the role and membership of the Community of Practice has evolved. As well, the programme has created a Community for Change that consists of three integrated but smaller self-selected groups, with sufficient homogeneity and focus as well as ambition for change, to achieve impact around three distinct opportunities to enhance New Zealand's Agricultural Innovation System. Further, two social scientists have joined the research team, bringing expertise in change management and using Communities of Practice to support institutional change, demonstrating how the logical framework supported identification of the need for new relationships. Another example of how the logical framework enabled accountability measures to be redefined is that one of the key measures of successful programme impact, 'rate of adoption', was changed to 'rate of innovation' to better reflect the ambition for change in the programme.

Logical frameworks fail to capture the effects of external and internal pressures on the ambition for change

Reflexive narratives by programme participants provide evidence of the programme's contribution towards system changes and how current interventions, like the Community of Practice and case studies, could be adapted to suit external changes in the primary industry. For example, severe market volatility in dairying provided an opportunity for systemic change as well as adapting programme delivery.

Although this took resources away from the dairy innovation project within Primary Innovation, it created opportunities to discuss with industry partners different modes of intervention, like co-innovation and its benefits and costs (Botha et al., 2015b). These discussions challenge current modes of operation and institutions and provide opportunities to explore structural and institutional changes at the innovation project level as well as at the Agricultural Innovation System level.

Logical frameworks do not take into account the emergent nature of complex projects

The logical framework enabled a holistic view of Primary Innovation that included processes and assumptions like how programme goals, in the form of outputs and outcomes, would be achieved, as well as the anticipated role and impacts of the Community of Practice. The focus was on the *type* of impacts that might be observed from the research programme rather than specific pre-determined actions and impacts. This high-level perspective was critical to deep

reflexivity because it raised and encouraged in-depth discussions amongst the team that went well beyond merely achieving pre-set goals. Avoiding a solely goal-oriented evaluation approach helped the programme management team to 'see' the need for changes, like how the Community of Practice was being used. Combining reflexive narratives with the logical framework also helped to articulate and 'picture' programme outputs and impacts, and inform how anticipated interventions could be adapted to fit, and more strongly influence, the changing programme environment.

Logical frameworks are typically not applied during intervention processes nor include all stakeholders

As mentioned already the logical framework was used on an ongoing basis to guide evaluation and reflection, i.e. ex-durante, and most stakeholders were involved in these processes. For example, feedback was sought from all participants (including farmer and Māori representatives), directly after Community of Practice workshops and again through interviews. Evaluation was ongoing and involved the full project team in reviewing information provided from interviews. All of the project team had access to evaluation data and reports as they were collected and collated during the project.

Logical frameworks are commonly performed by external evaluators

In Primary Innovation the external evaluator has been contracted for the duration of the programme to support monitoring and evaluation, provide ongoing advice and build monitoring and evaluation capacity. However, rather than hindering the open exchange of views and learnings amongst the research team the external evaluator's involvement is treated as an essential component and a valuable, well-utilised programme resource. For example, in the initial monitoring and evaluation workshop in 2013 (which used the logical framework), 19 participants rated the extent to which the workshop process clarified the objectives and performance indicators of the programme as good (7.5/10; range of 5-9), the extent to which they were comfortable with the monitoring and evaluation approach as 7.5/10 and their understanding of their role in the monitoring and evaluation process as 6.8/10. Seventeen of the 19 participants also indicated that they would consider the evaluation approach for other projects in their organisations. By the second monitoring and evaluation workshop (July 2014) there were a number of comments that indicated some participants already had a good understanding of monitoring and evaluation (from the first workshop) and the value of reinforcement and review. One noted that "*the ability to physically 'spread out' the monitoring and evaluation plan* (i.e., the logical framework) *on the wall was a huge bonus - I could see it. We were able to locate and augment the framework so it became a tool to enhance project planning and management. I can now see how monitoring and evaluation works!"* Another pointed out the value in being able to pick up the gaps.

Through the involvement of the external evaluator during the research programme the team's understanding grew over the three workshops and there was an increasing consensus and commitment built around the programme aims and process as well as their role in its evaluation and reporting. As noted earlier, a team member was appointed to facilitate the evaluation process through the project with the external evaluator providing an ongoing source of mentoring and support that has created a high-level of trust and confidence in using the logical framework.

Discussion

This section of the paper discusses the main insights on how to simultaneously meet accountability, reflexivity, learning and action needs in practice in system innovation projects. We argue that the ways in which logical frameworks are utilised determine their suitability to achieve system innovation while simultaneously meeting the accountability needs of funders.

System innovation is challenging and involves long-term change (Geels, 2004). We are, however, realistic about what can be achieved through Primary Innovation and acknowledge – with Van Mierlo et al., (2010a: 144) – that, by definition, significant system change is complex with long time horizons. While this may limit what is achievable by a programme like Primary Innovation, contributing to system innovation by stimulating learning in the sense of a change of thinking and acting is, at this stage, the Primary Innovation team's shared `ambition for change', towards which the logical framework is enabling steady progress to be made.

Co-innovation occurs when different stakeholders collaborate and, over time, share their unique knowledge to solve a problem or realise an opportunity. The mid-term review has shown that the logical framework has enabled participants, through joint learning, to reach consensus on the overall plan for this complex programme by bringing the goals (impacts), objectives, outputs and activities together. This is important because shared goals, and a common understanding of the approach being taken to reach these, maximise the likelihood of research having impact (Campbell et al., 2015). Using the logical framework provided an agreed, consistent and coherent summary of the programme to all the stakeholders, including the funders. It supported culture change and helped discussions to focus on the collective goals, or 'ambition for change' (see e.g. Van Mierlo et al., 2010c) by creating a shared visual record of where things are heading.

The logical framework was useful to guide responses to changes or issues because it was used in a flexible way. A flexible approach to monitoring and evaluating progress is very important in allowing for ongoing adaptation, particularly when sustainable solutions are pursued (Rijswijk et al., 2015). Consideration of any unexpected, or even expected, negative outcomes is provided for every time the programme team meets to minimise the risk of them becoming fatal flaws and enable the programme to adjust in order to manage them. The mid-term review has confirmed that the way in which the logical framework was used in Primary Innovation provided sufficient flexibility, allowed adaptation, and thereby effectively guided co-innovation, while also providing accountability measurements and responses to the funders.

By explicitly including the process and principles of co-innovation in the way the logical framework was designed and applied, the team has been able to continuously evaluate in real time how well the process and principles are being applied across the programme work streams while learning together about the value of their application. It also facilitated accountability by measuring progress being made towards the delivery of programme impact.

Conclusion

Our use of a logical framework provided a point of convergence that stimulated project team members to spell out their assumptions about the relations between project activities and long-term goals as well as their own viewpoints and actions, and subject them to scrutiny and evaluation (see also Arkesteijn, van Mierlo, & Potter, 2007). This has helped participants to understand and, through reflexive narratives, learn about each other's viewpoints and actions.

In our view this is an essential component of identifying and addressing institutional logics, which support system innovation.

When logical frameworks are employed adaptively they can, like in Primary Innovation, enhance project functioning in ways that create change while staying accountable. In our experience, when used adaptively, logical frameworks can greatly enhance the possibility of achieving impact by enabling reflection upon, and responding to, important contextual changes. We acknowledge that, when logical frameworks are used to control resource allocation and/or to manage contracts based on accountability, they can become a fixed and prescriptive mould that locks a project team into pre-determined actions, activities and outcomes while ignoring contextual changes.

Acknowledgements
We would like to acknowledge the Ministry of Business, Innovation and Employment for funding the Primary Innovation programme (CONT-30071-BITR-AGR) through a Biological Industries Targeted Research grant as well as funding by New Zealand dairy farmers through DairyNZ (RD1429). We would also like to acknowledge the substantial co-funding supplied by a range of New Zealand agricultural and horticultural sectors to this programme as well as in-kind support from a number of research providers. This paper was greatly improved by discussions with all participants, particularly Laurens Klerkx, at the two-day Primary Innovation write shop in February 2016 facilitated by Lora Hageman. Finally thank you to Helen Percy, Simon Fielke and Barbara van Mierlo for their critical review of the paper.

References

Arkesteijn, M., van Mierlo, B., & Leeuwis, C. (2015). The need for reflexive evaluation approaches in development cooperation. Evaluation 21(1): 99-115.

AusAid. (2005, October 2005). 3.3 The Logical Framework Approach. AusGuideline. Activity design: Commonwealth of Australia. Retrieved 23 May 2016, 2016, from http://www.sswm.info/sites/default/files/reference_attachments/AUSAID%202005%20The%20Logical%20Framework%20Approach.pdf.

Beers, P. J., Hermans, F., Veldkamp, T., & Hinssen, J. (2014). Social learning inside and outside transition projects: playing free jazz for a heavy metal audience. NJAS - Wageningen Journal of Life Sciences 69: 5-13.

Botha, N. (2013). Co-innovation and Innovation Systems: Glossary of Terms. N. Botha. Hamilton, New Zealand: AgResearch.

Botha, N., Coutts, J., & Turner, J. A. (2015a). Using a Logical Framework Approach to evaluate and guide co-innovation: a New Zealand case study. 22nd European Seminar on Extension and Education (ESEE 2015). Wageningen, The Netherlands.

Botha, N., Klerkx, L., Small, B., & Turner, J. A. (2014). Lessons on transdisciplinary research in a co-innovation programme in the New Zealand agricultural sector. Outlook on Agriculture 43(3): 219-223.

Botha, N., Sinclair, S., Turner, J., Blackett, P., Brazendale, R., Dirks, S., Lambert, G., & Kaine, G. (2015b). Estimating the economic impact of a co-innovation approach: the case of dairy heifer rearing in New Zealand. Impacts of Agricultural Research - Towards an Approach of Societal Values. Paris.

Brunori, G., Rand, S., & Proost, J. (2008). Towards a conceptual framework for agricultural and rural innovation policies. WP1 Synthesis Report, IN-SIGHT EU project.

Bussels, M., Happaerts, S., & Bruyninckx, H. (2013). Evaluating and monitoring transition initiatives. Lessons from a field scan. Leuven, Policy Research Centre TRADO.

Campbell, C. A., Lefroy, E. C., Caddy-Retalic, S., Bax, N., Doherty, P. J., Douglas, M. M., Johnson, D., Possingham, H. P., Specht, A., Tarte, D., & West, J. (2015). Designing environmental research for impact. Science of The Total Environment 534: 4-13.

Coutts, J., Botha, N., & Turner, J. A. (2014). Evaluating a co-innovation policy initiative in New Zealand. Farming systems facing global challenges, Berlin, 11th European International Farming Systems Symposium.

Dart, J. (2007). Training in Monitoring, Evaluation, Reporting and Improvement (MERI). Melbourne: Clear Horizon.

Ekboir, J. (2003). Why impact analysis should not be used for research evaluation and what the alternatives are. Agricultural Systems 78(2): 166-184.

Fischer, A. H., Beers, P., van Latesteijn, H., Andeweg, K., Jacobsen, E., Mommaas, H., van Trijp, H. M., & Veldkamp, A. (2012). Transforum system innovation towards sustainable food. A review. Agronomy for Sustainable Development 32(2): 595-608.

Fuenfschilling, L., & Truffer, B. (2014). The structuration of socio-technical regimes - conceptual foundations from institutional theory. Research Policy 43(4): 772-791.

Funnell, S. C., & Rogers, P. J. (2011). Purposeful programme theory: effective use of theories of change and logic models. San Francisco: Jossey-Bass/Wiley.

Gaspar, D. (1999). Evaluating the "logical Framework Approach" - towards learning-oriented development evaluation. Working Paper Series. I. o. S. Studies. The Hague.

Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems:insights about dynamics and change from sociology and institutional theory. Research Policy 33(6–7): 897-920.

Gosling, L., & Edwards, M. (2003). Toolkits: a practical guide to planning, monitoring, evaluation and impact assessment. Save the Children UK.

Hekkert, M. P., Suurs, R. A. A., Negro, S. O., Kuhlmann, S., & Smits, R. E. H. M. (2007). Functions of innovation systems: a new approach for analysing technological change. Technological Forecasting and Social Change 74(4): 413-432.

Kaplan, J. (2015, 3rd February). Logframe. Retrieved 23 May 2016, from http://betterevaluation.org/evaluation-options/logframe

Klerkx, L., Aarts, N., & Leeuwis, C. (2010). Adaptive management in agricultural innovation systems: the interactions between innovation networks and their environment. Agricultural Systems 103(6): 390-400.

Klerkx, L., van Mierlo, B. & Leeuwis, C. (2012). Evolution of systems approaches to agricultural innovation: concepts, analysis and interventions. In I. Darnhofer, D. Gibbon and B. Dedieu (Eds) Farming Systems Research into the 21st Century: The New Dynamic, pp. 457-483, The Netherlands: Springer.

Knickel, K., Brunori, G., Rand, S., & Proost, J. (2009). Towards a better conceptual framework for innovation processes in agriculture and rural development: from linear models to systemic approaches. The Journal of Agricultural Education and Extension 15(2): 131-146.

Loeber, A., Van Mierlo, B., Grin, J., & Leeuwis, C. (2007). The practical value of theory: conceptualising learning in the pursuit of a sustainable development. In (Eds) Social Learning Towards a Sustainable World: Principles, Perspectives, and Praxis, pp. 83-97.

Nederlof, S., Wongtschowski, M., & van der Lee, F. (2011). Putting heads together: agricultural innovation platforms in practice. The Netherlands: KIT publishers.

Nettle, R., Brightling, P., & Hope, A. (2012). How programme teams progress agricultural innovation in the Australian dairy industry. The Journal of Agricultural Education and Extension 19(3): 271-290.

Regeer, B. J., de Wildt-Liesveld, R., van Mierlo, B., & Bunders, J. F. G. (2016). Exploring ways to reconcile accountability and learning in the evaluation of niche experiments. Evaluation 22(1): 6-28.

Rijswijk, K., Bewsell, D., Small, B., & Blackett, P. (2015). Reflexive monitoring in New Zealand: evaluation lessons in supporting transformative change. Evaluation Journal of Australia 15(4): 38-43.

Roberts, K., & Coutts, J. (2011). Monitoring, evaluation and reporting of extension. In J. Jennings, R. Packham and D. Woodside (Eds) Shaping change: Natural Resource

Management, Agriculture and the Role of Extension. Australia: Australasia Pacific Extension Network (APEN).

Schwind, J. K., Cameron, D., Franks, J., Graham, C., & Robinson, T. (2012). Engaging in narrative reflective process to fine tune self-as-instrument-of-care. Reflective Practice 13(2): 223-235.

Spielman, D. J., Ekboir, J., & Davis, K. (2009). The art and science of innovation systems inquiry: applications to Sub-Saharan African agriculture. Technology in Society 31(4): 399-405.

Turner, J. A., Klerkx, L., Rijswijk, K., & Barnard, T. (2016). Systemic problems affecting co-innovation in the New Zealand Agricultural Innovation System: identification of blocking mechanisms and underlying institutional logics. NJAS - Wageningen Journal of Life Sciences 76: 99-112.

Van Mierlo, B., Arkesteijn, M., & Leeuwis, C. (2010a). Enhancing the reflexivity of System Innovation Projects with System Analyses. American Journal of Evaluation 31(2): 143-161.

Van Mierlo, B., Regeer, B., van Amstel, M., Arkesteijn, M., & Beekman, V. (2010b). Reflexive Monitoring in Action. A Guide for Monitoring System Innovation Projects. Wageningen/Amsterdam: Communication and Innovation Studies, WUR.

Van Mierlo, B., Regeer, B., Van Amstel, M., Arkesteijn, M. C. M., Beekman, V., Bunders, J. F. G., De Cock Buning, T., Elzen, B., Hoes, A. C., & Leeuwis, C. (2010c). Reflexive Monitoring in Action. A Guide for Monitoring System Innovation Projects. Wageningen/Amsterdam, Communication and Innovation Studies, WUR; Athena Institute, VU**:** 104.

van Veen, S. C., de Wildt-Liesveld, R., F.G., B. J., & Regeer, B. (2014). Supporting reflective practices in social change processes with the dynamic learning agenda: an example of learning about the process towards disability inclusive development. International Journal of Learning and Change 7(3/4): 211-233.

Wiskerke, J. S. C., & van der Ploeg, J. D. (2004). Seeds of transition. Essays on novelty production, niches and regimes in agriculture. . Assen, the Netherlands: Royal Van Gorcum.

Outcome evidencing: a rapid and complexity-aware evaluation method

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

[1] *Independent Learning and Evaluation Researcher and Consultant*
[2] *Independent Evaluator*

Abstract: This paper describes the development and use of a rapid evaluation approach to meet programme accountability and learning requirements in an on-going research for development programme operating in five developing countries. The method identifies clusters of outcomes to which the programme has contributed, within programme 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 programme theory allows the programme 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 programme interventions are working. Repeated cycles of Outcome Evidencing can build a case for programme contribution over time that can be evaluated as part of any future impact assessment of the programme 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 programmes 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 programme offerings will work and in others they will not (Pawson, 2013). Evaluation methods therefore need to understand how different aspects of programmes work, for whom in different contexts. In other words, they need to unpack the causal black box between programme intervention and programme outcomes (Astbury & Leeuw, 2010). Most traditional impact evaluation methods do not dig into causality, but rather concentrate on establishing the worth of programme intervention, often evaluating it against whether its initial predicted routes to impact have come to pass (Mayne & 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 programme 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 & 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 programmes 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 programme of the CGIAR. The CGIAR is a worldwide

partnership addressing agricultural research for development carried out by 15 research centres through fifteen CGIAR research programmes. 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 programmes 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 Programme 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 programmes of work in five geographically defined hubs with an aspirational goal to make a positive difference to the livelihoods of 6 million poor and marginalised by 2023 (AAS, 2014). By the end of 2013 AAS was implementing programmes of work in the coral triangle of the 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 marginalised small-scale fishermen 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 & Glaser, 2011).

In the same period the programme 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 programme theory is based on the premise that agricultural research processes (e.g. multi-partner collaborations) and outputs (i.e. new technologies) work to catalyse and foster processes of rural innovation. It is these innovation processes that may be 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 programme 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 programme 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 programme 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 programme's beneficiaries. This is similar to Rogers' (2008) idea that programme theory can may be used to identify emergent outcomes that have the potential to make a big difference. To work in this way, the programme needed a method of quickly identifying emerging outcomes, both expected and unexpected.

The Outcome Evidencing method

We combined elements of Outcome Harvesting (Wilson-Grau & 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.

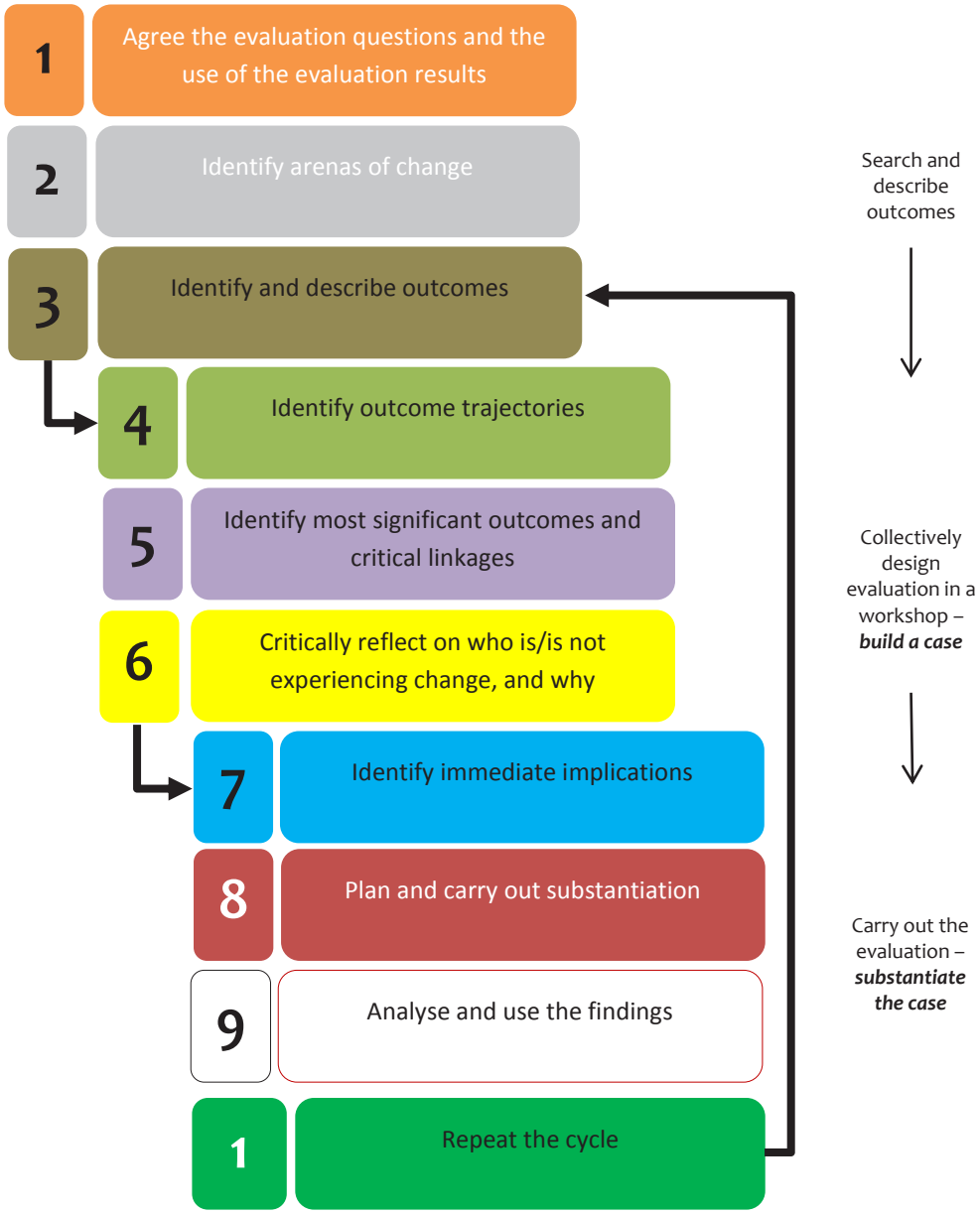

Figure 1: Ten steps of an Outcome Evidencing process

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

Step 1 involves programme 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 programme 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 programme is doing that is working.

Step 2: Identify key areas of change

In Step 2 knowledgeable programme staff identify areas of change to which the programme 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 programme 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 & Geels, 2008). Hence strategic niche management provides some detail to the AAS' programme theory described above, specifically that the programme creates, supports and guides socio-technical niches to be building blocks that come together to help achieve the programme's goal. Focusing rapid evaluation on if and how programme intervention is contributing to niches was a way of answering the second evaluation question relating to the credibility and workings of the AAS programme theory. Evaluation findings can guide how the programme 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 programme in the field. The workshop identifies outcome trajectories by which the programme is contributing to areas of change. The outcome trajectories, described as theories of change, identify and explain the causal links connecting programme 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 programme 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 recognisable. Just as identifying criminals' modus operandi can help catch them, so identifying programmes' modus operandi can help improve them by understanding how the programme 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 and involve a causal mechanism or mechanisms that produce an outcome or outcomes (Westhorp, 2014). Causal mechanisms are what intervenes between the delivery of programme service and the occurrence of outcomes of interest (Weiss, 1997). A mechanism is the response programme 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 programme'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 programme impact claims than others: these require greater scrutiny. The scrutiny helps clarify the programme's unique modus operandi - the distinctive set of underlying causal mechanisms that the programme is triggering. If the programme'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 programmes, is to benefit the poor and marginalised who are usually by-passed by mainstream development activity. Hence we include a step that involves analysing outcome trajectories in terms of social and gender equity, inclusion and power. This information helps AAS catalyse, support and modify outcome trajectories to favour poor, vulnerable and marginalised 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 analyse 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 marginalised groups are being, or could be, directly or indirectly affected by the change?
- does the outcome trajectory:
    - promote equal opportunities for vulnerable and marginalised 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; analyse 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 programme resources to generate outcomes. This is then compared to AAS' existing programme theory and action taken if required.

Step 9: Analyse and use the findings

The evaluator who has carried out the substantiation and other staff leading the Outcome Evidencing process analyse the findings from the substantiation to complete the evaluation report. Outcome Evidencing was designed to be repeated annually or bi-annually within a programme 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 programme 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 programme 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 & Munslow, 2014) and programme 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 Programme 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 Steps 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. Farmers 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. Farmers becoming self-confident leading to outcomes such as farmer-researchers taking up leadership roles and becoming recognised by other organisations.
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 to the market.
4. Influencing 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.

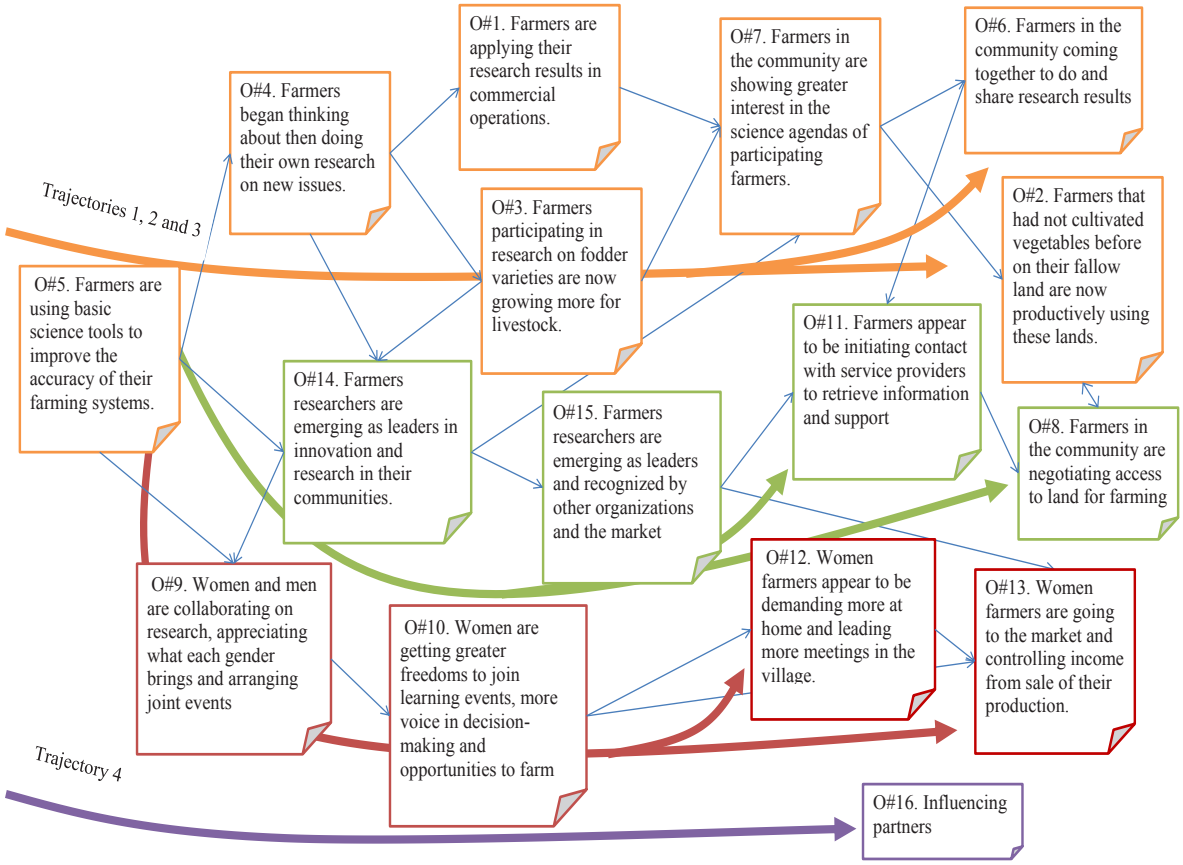

Figure 2. Outcome trajectories as identified during the workshop in Bangladesh

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 programme had contributed through implementing the RinD approach that had resulted in outcomes;
2. specify the causal links between the outcomes;
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.

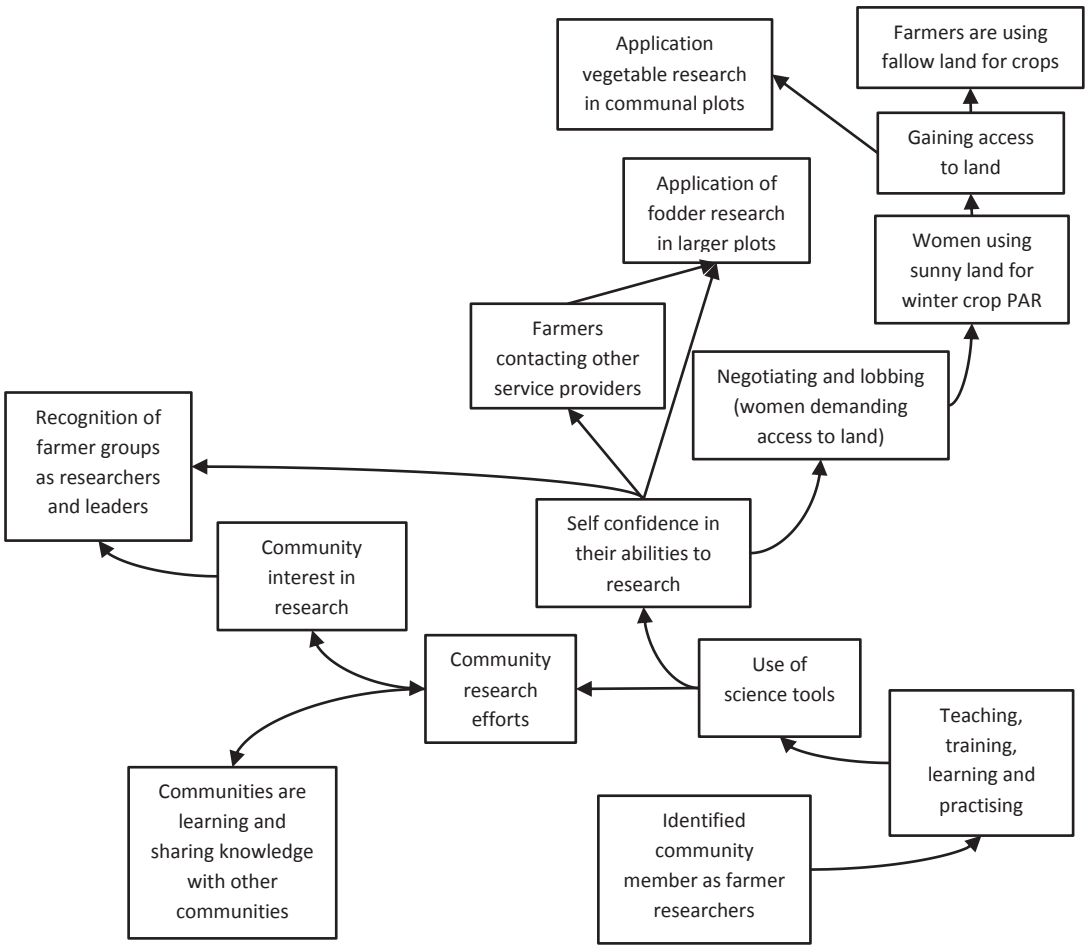

Figure 3. A multi-cause diagram depicting the theory of change for the capacity to do research trajectory

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 organisations 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 organisations showed interest in participating in different AAS programme 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 recognising the need to critically reflect on 'who' was experiencing change and to be sure AAS was reaching the poor and marginalised.

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 programme 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 (an approach introduced by AAS involving visioning, self-assessment, prioritising, 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 crop, 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 neighbouring 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 Programme 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-neighbour-farmers-(non-AAS) communicated with us (AAS farmer) about the research technology like line to line spacing, fertiliser dose, good quality seed and we (AAS farmers) assisted them in this regard. 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 organisation 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 summarises 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 organisations. – 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 organisations. 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 organised 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 recognising that the RinD approach is markedly different from their approaches and starting to adopt aspects of it;
3. communities recognising 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 programme are expanding – Partners using AAS outputs – Partners recognising the importance of participatory approaches – Key staff of partner organisations are becoming more aware of social inclusion issues and are more conscious of engaging the poor and marginalised particularly in conducting research activities	Endorsement of the AAS Programme by the Regional Development Council (RDC)	– RDC endorsements in Regions 7, 8, 10 – Letter 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 the RinD process of identifying community needs and use outputs of this process in targeting beneficiaries	– Focal group discussions and pre-testing 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 referenced with hyperlinks to documentary evidence held on an internal site.

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

Endorsement of the AAS Programme 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 Programme 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 programme in full RDC sessions. The 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 Programme 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 Programme 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 Programme 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 stakeholders to explore collaboration including the support of the endeavors tackling the HDC. At least $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 organisations to work towards a common goal through a number of initiatives. They realised that what it takes are communities that can organise and express their development requirements and an 'honest broker' able to link communities' visions and dreams and organisational mandates. They concluded that research organisations 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 realise 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 programme 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 realisation that better education about the damage done by illegal fishing methods was not working without illegal fishermen 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 programme value.

Reflection on novelty of Outcome Evidencing

As already mentioned, Outcome Evidencing is a hybrid of outcome harvesting and the modus operandi methods. Wilson-Grau and Britt (2012) describe Outcome Evidencing as an evaluation approach that starts with emerging outcomes and works back to establish if and how programme interventions have contributed by reconstructing and validating causal pathways. The steps in the outcome harvesting method are summarised in Box 4.

Box 4. Outcome harvesting in brief *(Adapted from Wilson-Grau & 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. Analyse and interpret: harvesters organise outcome descriptions through a database in order to make sense of them, analyse 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 programme 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 programme 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 is 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 programme 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 programme areas of change. It then seeks to understand, describe and verify these outcomes to support learning. The method is centred 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 programme log frames, or equivalent, allows the programme 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 programme intervention worked, for whom, to what extent and why. However, it is likely to be most useful as a central part of programme monitoring and evaluation. Repeated cycles of Outcome Evidencing build a case for programme contribution over time that can be evaluated as part of any future impact assessment of the programme 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 programme contribution within theories of change rather than programme contribution to discrete outcomes.

References

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

AAS (2014) CGIAR Research Programme 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., & 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: Cambridge University Press.

Halliday, A., & 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., & 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). The Netherlands: Springer.

Mayne, J., & E. Stern. (2013). Impact evaluation of natural resource management research programmes: a broader view. ACIAR Impact Assessment Series Report No, 84. Australian Centre 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). Maximising 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). Naturalising sensemaking. In K. L. Mosier and U. M. Fischer (Eds) Informed by Knowledge: Expert Performance in Complex Situations pp. 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 June, 2012.

Small-scale farmers' perspectives on what enhances capacity to innovate

Wettasinha, C.[1], Diop, J-M. [2], van Veldhuizen, L.[3], Waters-Bayer, A.[4] and Douthwaite, B.[5]

[1]*PROLINNOVA International Support Team, Netherlands*
[2]*PROLINNOVA International Support Team, Belgium*
[3]*Royal Tropical Institute, Netherlands*
[4]*PROLINNOVA International Support Team, Germany*
[5]*Independent consultant, Ireland*

Abstract: Agricultural research and development (ARD) agencies are now more aware of the importance of enhancing the capacity of small-scale farmers to innovate and to become better able to adapt to new conditions, problems and opportunities. Challenges for these agencies include: i) monitoring and evaluating changes in capacity to innovate (C2I) at individual and community level as a result of their interventions and ii) using the monitoring and evaluation (M&E) process as a means for all stakeholders in ARD to learn about what favours and constrains local innovation. The intervening ARD actors usually develop the M&E approaches, criteria and indicators to use. In order to better understand the factors that influence C2I from the perspective of small-scale farmers, a mini-study was carried out among 12 such farmers who showcased their innovations at the West African Farmer Innovation Fair in May 2015. The study explored what they saw as the main factors that strengthened local C2I. Semi-structured interviews revealed that many factors identified by the farmers were similar to those identified by intervening agencies, but other factors were mentioned only by farmers, e.g. the role of supportive family members, neighbours and others in their social networks in the innovation processes. Although very limited in scope, this mini-study indicated that there is more to C2I than intervening ARD agencies may expect. This paper calls for attention to this essential yet neglected aspect – the perspectives of small-scale farmers – in evaluating programmes that seek to build C2I as part of their theory of change.

Key words: Agricultural innovation systems, capacity to innovate, emic perspective, farmer-led research, local innovation, monitoring and evaluation

Introduction

Agricultural research and development (ARD) agencies are now becoming increasingly aware of the importance of enhancing the capacity of small-scale farmers and their communities to innovate (e.g. FAO, 2014; Leeuwis et al., 2014; Atta-Krah et al., 2015) and thus become better able to adapt to new conditions, problems and opportunities – in other words to become more resilient. A relatively new challenge for these agencies is how to monitor and evaluate changes in this capacity as a result of their interventions. How can outcomes be measured in terms of changes in capacity to innovate at individual, family and community level? At the same time, how can one use a process of monitoring and evaluation (M&E) as a means for all stakeholders in ARD – including the farmers and farming communities – to learn about what is stimulating and favouring and what is undermining or constraining capacity to innovate?

Developing relevant M&E approaches, methods and tools starts with clearly defining what C2I entails and what the key factors and parameters are that enhance or influence this C2I, in order to identify key indicators around which the M&E could be organised. Three system-oriented research programmes of the CGIAR group of international agricultural research institutes – Humidtropics, Dryland Systems and Aquatic Agricultural Systems (AAS) – had

included enhanced capacity to innovate (C2I) as an intermediate development outcome in their respective theories of changes. Leeuwis et al. (2014) therefore attempted to define C2I from the perspective of these research programmes, drawing on the work of a range of stakeholders, including civil society organisations. They identified some core capacities at the level of individual stakeholders and further capacities at the level of facilitators of system innovation that, together, would form a system's capacity to innovate. Derived from their definition, the core capacities that contribute to an enhanced C2I at the level of farmers and farming communities would be:

- the capacity to continuously identify and prioritise problems and opportunities in a dynamic systems environment;
- the capacity to take risks, experiment with social and technical options, and assess the trade offs that arise from these;
- the capacity to mobilise resources and form effective support coalitions around promising options and visions for the future;
- the capacity to link with others in order to access, share and process relevant information and knowledge in support of the above; and
- the capacity to collaborate and coordinate with others during the above, and achieve effective concerted action.

Leeuwis et al. (2014) stress that the facilitators of system innovation need to have a conceptual understanding of how change comes about and how to intervene effectively in order to enhance a system's C2I. They were looking at the capacity of a larger innovation system – involving not only farmers and farming communities but also other actors in research, extension, private sector and government administration. Based on this view of core capacities that need to be strengthened, the CGIAR system-oriented research programmes began to design how they would measure changes in C2I within an innovation system.

This entry into designing an M&E system comes from the perspective of those seeking to facilitate innovation at a fairly high level of an agricultural innovation system. The international network PROLINNOVA (www.prolinnova.net) focuses on farmers at the grassroots level who are developing their own innovations – new and better ways of farming and managing natural resources locally – not because they are 'pushed' by a project or by research or extension staff but rather on their own initiative, drawing on their own creativity to combine information and ideas from multiple sources, including their own ideas. In this approach to promoting innovation, the innovators are not primarily those who adopt what outsiders are instructing them to do. The C2I lies not only in the 2.5% of the farming population called "innovators" in Roger's (1962) influential model on diffusion of innovations but rather, to differing degrees, among all farmers in the course of their "performance" (Richards 1989) in day-to-day farming. They innovate for a variety of reasons and differ from each other depending on their situation, needs and opportunities and can be divided into several categories on this basis (see Box 1 for a categorisation developed by the authors of this paper). PROLINNOVA seeks to recognise and enhance C2I in all farmers – also the very poor – and to help farmers 'perform' better through improved communication with each other and with outside actors and through greater self-confidence to take and keep the lead in participatory research and development.

Box . Which farmers innovate and why?

The work of the PROLINNOVA network and similar initiatives (e.g. in the ISWC, JOLISAA and PFI[1] programmes) and research by anthropologists (e.g. Nielsen, 2001) have revealed a wide array of farmer innovators in different socio-economic situations and with different motivations. Some reasons for innovation that have been encountered and have also been articulated by farmers themselves are:

- Innovation out of dire necessity, motivated by extreme poverty, e.g. immigrants who have been allocated heavily eroded land and are struggling to produce something for their family to live on from this land are obliged to innovate (for examples from ISWC, see Reij & Waters-Bayer, 2001);
- Innovation out of curiosity or 'by accident', often done by small-scale and resource-poor farmers, including women, but not very obvious to outsiders (e.g. Abay et al., 2001). In northern Ghana, Tambo (2015) found that 35% of the farmers innovated out of curiosity;
- Innovation to deal with a specific challenge or problem, such as the woman innovator in Kabale, Uganda, who developed a low-cost solution to kill ticks and mites, derived from a leguminous tree that extensionists in the area promoted for improving soil fertility (Critchley & Lutalo, 2006);
- Innovation to improve the local economic or ecological situation, such as the numerous farmer innovators encountered by PROLINNOVA who set up backyard botanical gardens to domesticate fast-disappearing wild plant species (Abay et al., 2010), or a community in Senegal that set up a system for providing food for poorer families (Agrecol-Afrique, 2013);
- Innovation as a pastime – these are usually better-off farmers who have the time and money to try new ways of doing things and, because they can afford to take risks, can innovate in a bigger and more obvious way. For them, innovation is almost a 'hobby' or 'game'. They are perhaps closest to what Rogers (1962) refers to as "innovators".

This categorisation does not mean that individuals or groups that innovate continue doing it for the same reason as when they started. Often, farmers who initially innovated out of dire necessity and managed to improve their livelihoods then started to try out new things simply out of curiosity. Similarly, better-off farmers who can afford to innovate as a pastime, if faced with a challenge, use their innovative capacity to find ways of getting around it. If C2I among farmers has been nurtured and strengthened, exploring new possibilities becomes second nature to them.

As a network trying to recognise and promote endogenous innovation processes in agriculture and to support farmer-led processes of research and development, PROLINNOVA saw a necessity to explore the 'insider' perspective of farmers on C2I. How do innovative farmers, groups or communities regard themselves? What do they see as the main attributes of an outstanding local innovator? What do they see as the main factors that favour or constrain their own innovation processes? How do they assess as individuals, groups or communities whether they have become stronger in terms of C2I (in whatever way they may express this capacity themselves)? The PROLINNOVA International Secretariat therefore carried out a small pilot study to explore views of small-scale farmers recognised as outstanding innovators in order to find out what factors form and influence local capacity to innovate in their own reality.

[1]ISWC: Indigenous Soil and Water Conservation (Reij & Waters-Bayer, 2001); JOLISAA: Joint Learning in Innovation Systems in African Agriculture (www.jolisaa.net); PFI: Promoting Farmer Innovation (Critchley et al., 1999).

This mini-study was supported by the CGIAR Research Programme on AAS, with co-funding from McKnight Foundation.

Pilot study on local perspectives on capacity to innovate

In May 2015, 50 small-scale farmer innovators from eight countries in West Africa gathered in Ouagadougou, Burkina Faso, to take part in the West African Farmer Innovation Fair. This provided a unique opportunity to start finding out what the elements and factors are that determine local C2I from the perspective of such farmers. Small multi-stakeholder teams in each country had selected these farmers as being particularly innovative. In the fair, they were able to present and share their innovations, learn from other innovators and interact with professionals in formal ARD, including policymakers, as well as the general public visiting the fair. More information on the fair can be found at www.fipao.faso-dev.net.

The pilot study was conceived and planned by members of the PROLINNOVA support team based at the International Secretariat in the Netherlands. Given the explorative nature of the study the team decided to base it on focused, semi-structured interviews so as to elicit the views of the farmer innovators on aspects related to local (small-scale farmer) innovation and C2I. The interviews were conducted by a Belgian-Senegalese researcher, a team member with longstanding experience in supporting farmer-led innovation development in Africa. In some cases, he could conduct the interviews in the local language of the farmer innovators; in other cases, he had to work through translators.

Using short descriptions prepared by the fair organisers about each of the 50 farmers selected by national committees to exhibit their innovations at the West African Farmer Innovation Fair, the study team selected 12 innovators for the interviews. It sought as diverse a group as possible in terms of the country of origin, sex and age of the innovator and type of innovation. It shared its initial selection with members of the national committees for their review. Only in the case of Ghana did the national committee propose an alternative innovator to be interviewed. The farmer innovators selected were five women and seven men, who were interviewed in the midst of the fair for 1–2 hours each.

Table 1. Main characteristics of small-scale farmer innovators interviewed during the West African Farmer Innovation Fair *(n.k. = not known)*

No.	Country of origin	Age	Sex	Innovation presented at fair
1	Benin	28	F	Using compost in *zaï* pits and on broadcast plots (production technique)
2	Burkina Faso	38	F	Biopesticide for vegetable plants (plant treatment product)
3	Burkina Faso	62	M	'*Manegre*' or cellar or storage silo (technologies for preserving potato, onion and yam)
4	Cameroon	41	M	Awareness-raising and facilitation: creating a producers' association (institutional innovation)
5	Ghana	54	M	Fish feed (production technique)
6	Mali	52	F	Biopesticide (plant treatment product)
7	Mali	38	M	Incubator made of *banco,* i.e. mud mixed with straw (poultry production technology)
8	Niger	28	F	Community radio (communication technique)

9	Niger	50	M	Clearing aquatic weeds from ponds (natural resource management technique)
10	Niger	40	M	System of rice cultivation outside the irrigation scheme (production system)
11	Senegal	n.k.	M	Promoting and transforming family farms (institutional innovation)
12	Senegal	n.k.	F	Processing cashew nuts (production technique)

The study team had developed a short interview checklist in French to ensure consistency in terms of the information gathered. The checklist had five central questions, each of which was meant to generate discussions relevant for understanding farmers' views on C2I:

1) What are the characteristics of a good farmer innovator?
2) What supports and facilitates local innovation processes?
3) What limits or constrains local innovation?
4) What would you recommend to strengthen local innovation processes?
5) What would you recommend to address constraints to local innovation?

In the first part of the interview, the farmer was invited to talk about his/her specific innovation (what it is, how it works, what it does and what results it brings) as well as about the process of developing the innovation over the years (why, what, how, when). This was followed by a deeper probing into the farmer's view on his/her C2I (without using this term) and finding out what had helped and/or hindered him/her in the process and what he/she thought could support and facilitate the innovation process. The key question in this regard posed to the innovators was: "what does a farmer innovator need to have in order to be or become a better or more efficient innovator?"

The study team compiled all responses and further comments made by the 12 farmers pertinent to each of the main questions in the checklist and recorded these in tables per farmer and per question. It identified the main issues most frequently mentioned by farmers under each question and, from this listing in descending order of frequency of mention, drew out the views of the farmers related to C2I and their suggestions for enhancing the process of local innovation.

Findings of the pilot study

As this was an exploratory study to discover different elements of farmer innovators' perceptions on C2I and involved a very small sample of only 12 farmers, it did not lend itself to quantitative analysis beyond the use of frequency tables. The responses of the farmers are clustered here according to the five central questions mentioned above:

1) ***Characteristics of a good farmer innovator.*** When describing good farmer innovators, the interviewed farmers referred to personality traits of innovators, their interest and skills in 'research', their willingness to share and their ability to communicate and collaborate with others. As indicated in Table 2 they gave importance to personal characteristics that reflected research capacities in terms of analytical skills and the systematic comparison of alternatives, if needed, through experimentation. They emphasised the importance of communication skills for farmers to be able to access new ideas from various sources.

Table 2. Analysis of farmers' views on key characteristics of effective innovators

Key characteristics	Times mentioned
Personality traits	**12**
Pro-active, self-confident, persevering	6
Desire for continued development in his/her work	1
Dares to take risk, not afraid of critics	4
Follows intuition	1
Interest and skills in "research"	**9**
Observation; analysis of problems and options; comparing / weighing alternatives; experimentation; able to link past practice with current conditions	9
Interest in and capacity to communicate and share	**7**
Communicating with and convincing others	5
Looking for/accessing new ideas; language capacity to access information	2
Openness and capacity for (facilitating) collaboration	**4**
Open to others; collaborating with others to experiment; bringing people together; dialogue within family	4

2) ***Factors supporting and facilitating local innovation processes.*** The responses of the farmers (Table 3) revealed the particular importance they attached to the support received from people in their immediate social networks such as family members, neighbours and cooperative members. They also regarded advice, training and funding from external agencies as important elements of support. They rarely mentioned supportive policies at any level.

Table 3. Synthesis of farmers' views on key factors supporting local innovation

Key factors	Times mentioned
Individual	**8**
Own interest, insight, open spirit	7
Own funds generated from innovation	1
Family	**5**
Assistance, encouragement from family members	5
Community	**9**
Integration in farmers' group; experimentation in a group	2
Encouragement from neighbours; villagers asking advice	3
Spread of innovation by cooperative or other villagers	2
Support, encouragement, technical advice by farmer cooperative or group members	2
External agencies	**20**
Training support; visit by technical staff; advice in organising and managing group	8
Recognition by government agency	1
Provision of equipment	4

Funding	5
Participation in innovator fairs; support to increase visibility	2
Policies	**1**
Agriculture and park management policies	1

3) ***Factors limiting or constraining local innovation.*** The farmers mentioned six main factors as constraints in developing their innovations, as summarised in Table 4, which were – in descending order of frequency of mention – i) limited access to resources (land, labour, materials etc); ii) limited access to capital; iii) negative attitude of some external actors such as formal researchers; iv) lack of knowledge and skills such as literacy; v) inability to use certain kinds of equipment; and vi) opposition from parties within the community who feel that local innovation is a threat to their interests and established ways of doing things.

Table 4. Synthesis of farmers' views on key factors constraining local innovation

Key factors	**Times mentioned**
Resource-related constraints	**9**
Access to land; access to other materials required (availability, distance, costs)	5
Lack of labour	2
Others: protection of plots from animals (fences), rainfall	2
Funding	**6**
Lack of funds; short-term funding only; high bank interest rates	6
Role and attitude of external agencies	**7**
Lack of recognition by researchers; their attitude of superiority; danger of researchers or other experts hijacking the farmers' innovations	4
Lack of research support to improve innovation; research support expensive and risky	2
Lack of pathways to disseminate innovations	1
Opposing commercial interests	**4**
Local officials whose vested interests are threatened; opposition from entrepreneurs who control the market; scarce materials controlled by entrepreneurs/middlemen	4
Lack of knowledge or skills	**4**
Poor mastery of equipment needed for experiments; inability because of illiteracy to monitor and evaluate innovation well; lack of training in various aspects that could improve the process of innovation	4
Community attitude	**3**
Sabotage by community members; reluctance; group members not following	3

4) ***Recommendations to strengthen local innovation processes.*** As indicated in Table 5 many of the farmers gave high priority to the wider recognition by other development actors that local innovation is relevant for development. They called for changes in project design, M&E, reporting and impact assessment that make space for 'genuine' participatory research. They mentioned the importance of creating opportunities for learning, sharing

and networking such as innovation fairs, exchange visits and training sessions to enhance local innovation.

Table 5. Synthesis of farmers' recommendations for strengthening local innovation

Key recommendations	**Times mentioned**
Promote relevance of local innovation	**10**
General: Change in mentality of local authorities and leaders to accept local innovation; general promotion of local innovation; local innovation as relevant as formal research; lobby with donors for promoting local innovation; give recognition and space to farmer innovators	7
Specific: Encourage women to innovate; improve documentation of local innovation; involve innovators in schools and in teaching	3
Provide funding	**1**
Create funding support for innovators	1
Change role of external agencies	**6**
Research knowledge should support farmers in the field, all actors to collaborate with innovators in participatory research	2
Transparent project design; improved project monitoring and evaluation; correct reporting; prevent power politics interfering with development; post-project assessments built in to measure impacts	4
Facilitate access to and sharing of knowledge	**8**
Training	2
Farmer innovation fairs; exchange visits; space for innovators to explain their work; networking between innovators	6
Other	**2**
Reflection is needed on how to support local innovation and innovators	1
Promote spread and use of specific innovations	1

5) ***Recommendations to address constraints to local innovation.*** As can be seen in Table 6 the recommendations of the farmer innovators generally went in the same direction as under point 4 (above), but two recommendations stand out: i) removing barriers to accessing the resources (land, labour, transport etc.) needed by farmer innovators to carry out their work and ii) changing the roles of external agents to be truly collaborative and supportive of farmer innovation processes, specifically that researchers should support local innovation processes through engaging in joint research, adding value and providing training and coaching in relevant topics. In this regard farmers mentioned the need for training to bring about attitudinal change in external agents so that they can better support local innovation processes.

Table 6. Synthesis of farmers' recommendations to address constraints to local innovation

Key recommendations	Times mentioned
Undertake initiatives to address resource-related constraints	**6**
Use of local transport (not depend on external sources); find ways to get access to land; ensure availability of material (e.g. planting material) to continue innovation	6
Promote relevance of local innovation	**3**
Local awareness raising on relevance of local innovation; argue complementarity between local innovation and science-based innovation	3
Give attention to level and form of funding	**4**
Government payments to farmer innovators (as given to government extension staff); rewarding innovators when their innovations are widely spread; creating funding window to support local innovation	3
Funding support preferably in relatively small amounts but for longer periods of time	1
Change role of external agencies	**9**
Value addition by researchers to local innovation; validation of local innovation for easier spreading by agencies; research results better linked to farmer innovators; more participatory research	4
More interaction with innovators to address challenges; do not leave innovators to work in isolation; include farmer innovators in all development strategies	2
Training and coaching in financial management; training linked to local innovation to add value	2
Training for researchers and extension agents to open them up for recognising local innovation and to change their attitude	1
Promote community action	**1**
Promote collective action at community level	1
Create enabling legal and policy frameworks	**2**
Ensuring intellectual property rights for farmer innovations; legal changes to allow community radio to operate and be funded by the government	2
Create learning/training opportunities	**2**
Learning centres for young farmers interacting with innovators, literacy training	2

Overall, the farmer innovators referred mostly to local factors that directly influence their work and livelihood. Although they clearly defined how they thought the agricultural support system (making specific mention of extension agents and researchers) could best interact with farmer innovators, they made hardly any reference to the role of government policy. It is quite possible that relevant policies in the area of ARD indeed do not impact on their work, particularly in cases where the state is relatively weak and where formulated policies may not be implemented. In other cases, the farmers may not be aware that certain government policies are in fact partly responsible for the non-supportive behaviour of the ARD professionals they encounter.

Discussion: implications for M&E to enhance C2I

The brief study provided some insights into the individual perceptions of outstanding local innovators invited to an international farmer innovation fair. At such an event, the individual farmers are in the limelight and it is natural for them to focus on their particular technologies or niches, although they were asked to reflect on what favoured and constrained their capacity to innovate in a more general sense. Thus their personal bias should be acknowledged in framing the findings in the discussions on C2I.

However, from the experience of PROLINNOVA partners engaged in longer-term interaction with farmer innovators in their community settings, the discussions in such settings throw light on local innovation also from the perspective of the group or community and refer more to the processes (rather than only the specific technologies) involved in experimentation, innovation and wider sharing. It was more difficult in the setting of a large fair in a country that is foreign (for most of the interviewees) to probe beyond issues related to the outstanding innovations that had been their 'entry passes' to the fair.

Nevertheless, the study did reveal what these farmers viewed as being important in terms of favouring C2I at their level. Many of the key factors identified by the farmers that strengthened or constrained local capacity to innovate were similar to the factors that had been previously identified by field-based researchers and development actors, including those in the CGIAR system-oriented research programmes. However, the farmer innovators gave much more attention to the role of experimentation in innovation and acknowledged the role that supportive and appreciative family members, neighbours and others in local social networks played in the local innovation processes.

They also highlighted the importance of access to locally controlled resources to support their innovation processes (including locally controlled innovation funds). They pointed to aspects of their interaction with formal researchers and extension agents that indicate whether or not the local innovation system is functioning well. These aspects included: a) other ARD actors do not ignore local innovators and leave them to work in isolation; b) interaction between innovators and outsiders is in a participatory manner; c) outsiders add value to the innovators' own work, at least by giving scientific validation to what they have developed; d) interaction between research scientists and farmer innovators becomes more frequent and continues to address new challenges; and e) interaction with staff in ARD agencies also involves training and coaching in financial management and in how farmers could derive more value from their local innovations.

The views of the interviewed farmers on what is important in terms of C2I at their level have important implications for programmes seeking to build C2I as part of their theory of change, particularly with regard to areas of interest, criteria and indicators to be taken into account when monitoring and evaluating these programmes. For example, one or more indicators in the M&E might reflect whether resources for experimentation and innovation are locally controlled, as the farmers explicitly highlighted the importance of this.

Nielsen (2001) stresses the anthropological distinction of the etic view (the view from outside) and the emic view (the perspective from inside), in this case the perspective of farmer innovators themselves. In order to have a balanced picture of how C2I has been and can be enhanced, it would be necessary to seek both the etic and emic views. Moreover, it is quite likely that local perspectives on farmer innovation and on how innovative behaviour reveals itself and has been strengthened are specific to a country or ethnic group (cf. Nielsen, 2001). For this reason, attempts to monitor and evaluate C2I always need to seek the internal

perspective at the site in question. Moreover, examining the change in capacity together with the farmers and farming communities becomes part of the process of reflecting and learning at that level about the importance of innovation for the community and how different actors at the grassroots level can better play their roles in contributing to innovation processes. This reflection process, in itself, can contribute to improving the functioning of the local innovation system.

Programmes seeking to facilitate C2I need to take the diversity among farmer innovators (see Box 1) into account, especially in monitoring C2I. The perceptions of diverse farmer innovators would need to be explored in depth in order to gain a balanced picture of what C2I is and how it can be enhanced and measured. Multiple and mixed methods would need to be applied to capture this diversity and to integrate it into an M&E system. Interviewing 12 outstanding farmer innovators at the West African Farmer Innovation Fair was an initial attempt to harvest an emic perspective. These farmers represented only a part of the diversity of innovators among small-scale farmers in West Africa. Judging by their innovations and the reasons the farmers gave for having developed them, these farmers were mainly in the third and fourth categories of innovators listed in Box 1: those who were innovating to deal with a specific challenge or problem, and those who were innovating to improve the local economic and ecological situation.

The interviews stimulated these individual farmers to reflect on what helped and hindered farmer-led innovation processes. Other methods that would stimulate wider, community-based reflection and learning could involve focus-group discussions with self-formed groups of farmer innovators in a given community, plus joint discussions by farmer innovators and other community members on this topic. It would also be possible to ask resource persons in the community to identify different types of farmer innovators in their midst and then to carry out case studies of the innovation pathways and factors that led to the enhanced C2I of these different types of innovators. All of these methods would require that the facilitators of such farmer-led M&E processes are skilled in building rapport with small-scale farmers and their communities, in stimulating discussion and probing, and in finding appropriate spaces to be able to have unhurried conversations with various stakeholders at community level, including women, elderly and youth.

Finally, it should be emphasised that the small study presented in this paper is exploratory work that needs further and deeper investigation. It is simply meant to flag the importance of including farmers' perceptions on C2I in designing relevant M&E systems. The small sample of farmers' perceptions does not allow for specific suggestions as to how such systems should be designed and which indicators could be used. The PROLINNOVA network plans to use this approach and the initial findings to continue to develop a system of monitoring C2I that is inclusive, regarding the perspectives of farmer innovators to be just as important as those of other ARD stakeholders. The network hopes this will contribute to understanding and strengthening C2I at grassroots level.

Acknowledgements
The authors gratefully acknowledge the 12 farmer innovators for their willingness to participate in the interviews and to share their opinions and experiences. They also thank McKnight Foundation and AAS for co-funding this study and AgriProFocus, the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), the Collaborative Crop Research Program of McKnight Foundation, Misereor and the Swiss Agency for Development and Cooperation (SDC) for generously co-funding the West African Farmer Innovation Fair.

References

Abay, F., Lemma, M., O'Flynn, P., & Waters-Bayer, A. 2001. A challenge and an opportunity: innovation by women farmers in Tigray. In: C. Reij and A. Waters-Bayer (Eds.) Farmer innovation in Africa: a Source of Inspiration for Agricultural Development pp155–167. London: Earthscan.

Abay, F., Lemma, M., Gandarillas, E., Shrestha, P., Waters-Bayer, A., & Wongtschowski, M. 2010. Fundamentals of agrobiodiversity: promoting local innovation in managing agrobiodiversity. Issues Paper People, Food and Biodiversity Bas12-10-e. Eschborn: German Agency for International Cooperation (GIZ) Sector Project Sustainable Management of Resources in Agriculture.

Agrecol-Afrique. 2013. Caractérisation d'une innovation locale: lutte contre la soudure (insécurité alimentaire cyclique) dans le village de Pout Dagné, arrondissement et communauté rurale de Notto, Région de Thiès. Thiès, Sénégal: Agrecol-Afrique (www.prolinnova.net/sites/default/files/documents/Senegal/2013/gestion_de_la_soudure_a_pout_dagne-ag_profeis-ccafs.pdf).

Atta-Krah, K., Poire, V., & Koper, E. 2015. Humidtropics, a CGIAR Research Program on Integrated Systems for the Humid Tropics. Tropentag "Management of land use systems for enhanced food security: conflicts, controversies and resolutions", 16–18 September, Berlin, Germany

Critchley, W. (Ed.) 1999. Promoting farmer innovation: harnessing local environmental knowledge in East Africa. Regional Land Management Unit (RELMA) Workshop Report 2. Nairobi: RELMA / Swedish International Development Agency.

Critchley, W., & Lutalo, R.S. 2006. Controlling ticks and influencing policy. LEISA Magazine 22 (3): 24–25.

FAO, 2014. The state of food and agriculture: innovation in family farming. Rome: Food and Agricultural Organisation of the United Nations.

Leeuwis, C., Schut, M., Waters-Bayer, A., Mur, R., Atta-Krah, K., & Douthwaite, B. (2014). Capacity to innovate from a system CGIAR research program perspective. Penang, Malaysia: CGIAR Research Program on Aquatic Agricultural Systems. Program Brief: AAS-2014-29.

Nielsen, F. 2001. Why do farmers innovate and why do they not innovate more? Insights from a study in East Africa. In: C. Reij and A Waters-Bayer (Eds.) Farmer innovation in Africa: a Source of Inspiration for Agricultural Development pp.92 –103. London: Earthscan.

PROLINNOVA International Secretariat and Jean-Marie Diop. 2015. Understanding "capacity to innovate and adapt" from a smallholder perspective. Leusden, Netherlands: PROLINNOVA International Secretariat / ETC Foundation.

Reij, C., & Waters-Bayer, A. 2001. An initial analysis of farmer innovators and their innovations. In C. Reij and A. Waters-Bayer (Eds.) Farmer innovation in Africa: a Source of Inspiration for Agricultural Development pp. 77–91. London: Earthscan.

Richards, P. 1989. Agriculture as a performance. In R. Chambers, A. Pacey and L.A. Thrupp (Eds.) Farmer First: Farmer Innovation and Agricultural Research pp. 39–43. London: Intermediate Technology Publications

Rogers, E.M. 1962. Diffusion of Innovations. New York: Free Press of Glencoe.

Tambo, J.A. 2015. Farmer innovation in rural Ghana: determinants, impacts and identification. PhD dissertation, University of Bonn.

Monitoring & evaluation for research for development - building a results-based management system for climate smart agriculture

Förch, W., Schuetz, T. and Thornton, P

CCAFS, Kenya

Abstract: Making farming systems more climate smart requires taking different disciplines, sectors and scales into account, at the same time as facilitating farming system innovation within the context of climate change. Here we present a research-for-development programme's case of the evolution from a logframe approach to an outcome and results-based management oriented Monitoring, Evaluation and Learning (MEL) system. The CGIAR Research Programme on Climate Change, Agriculture and Food Security (CCAFS) is designing an impact pathway-based MEL system that combines classic indicators of research quality with innovative process and outcome indicators of developmental change. CCAFS has developed a methodology for evaluating with stakeholders factors that enable or inhibit progress towards behavioral outcomes in study sites and regions. Impact pathways represent the programme's best understanding of how engagement can bridge the gap between research outputs and outcomes in development. Strategies for enabling change include a strong emphasis on partnerships, social learning, gender and social inclusion, capacity building, communication and a MEL that focuses on progress towards outcomes. The importance of working with next-users in the development of impact pathways is highlighted as well as consistent engagement with partners and users of research outputs throughout the life of the programme. Theory of change can be used to balance the drive to generate new knowledge in agricultural research with the priorities and urgency of the users and beneficiaries of research results. Research alone may not lead to impact, but it can generate knowledge that can be put into practice to generate development outcomes.

Keywords: Impact pathway, innovation, theory of change, research for development, climate-smart agriculture.

Introduction

While global poverty has been reduced over the past 25 years, much remains to be done to reach the targets for 2030 as articulated in the Sustainable Development Goals (UN, 2015a). With an expected extra 2-3 billion people to feed over the next 40 years, this will require targeted efforts to achieve making 70% more food available to consumers to keep up with rapidly rising demand (WWAP, 2012). Climate change is already affecting agriculture in many developing countries and the effects will become increasingly challenging in the future. Climate change impacts are increasing the vulnerabilities of populations that are already struggling with food insecurity and poverty, even in the relatively conservative scenario of a global two degree rise in temperature (Thornton et al., 2014a).

Agricultural research for development (R4D) has played a significant role in reducing food insecurity over the last decades and will continue to play a critical role in addressing the above challenges (Raitzer, 2008). But it has not realised its full potential: the world food system continues to face challenges of persistent food insecurity and rural poverty in many parts of the developing world (FAO, 2014). Many studies have shown that 'scientifically proven'

technologies alone are not the only key to achieving an impact (Hartman & Linn, 2008; Pachico & Fujisaka, 2004).

In this paper we outline a R4D approach based on theory of change and impact pathway thinking for programme implementation, monitoring, evaluation and learning (MEL). This was undertaken by the CGIAR[1] Research Programme on Climate Change, Agriculture and Food Security (CCAFS). We describe the CCAFS case, where a theory of change approach combined with impact pathways and learning were employed to build an outcome-focused results-based management (RBM) for R4D. We discuss the experience, focusing on programme design and systems for planning and reporting. The paper concludes with lessons for required institutional change and for MEL practitioners, researchers and policy makers.

Background and Approach

CGIAR science is carried out by 15 research centres with 10,000 scientists working in 96 countries and a host of partners in national and regional research institutes, civil society organisations, academia, development organisations and the private sector (CGIAR, 2015). Its work contributes to the global effort to tackle poverty, environmental degradation, hunger and major nutritional imbalances (Raitzer & Kelley, 2008).

The challenges of demonstrating wide-reaching impact through R4D are compounded by a rapidly growing human population, climate change and other complexities of our time. To address this challenge CGIAR has broadened its portfolio of new initiatives for strategic research as part of a far-reaching reform process. A key part of the reorientation of the R4D portfolio was the move from an output to an outcome focus. Success is now measured in terms of research's contribution to behavioral changes, manifested in changes in knowledge, attitudes, skills and practices of a wide set of non-research next users, including development practitioners, farmers and policy makers. In 2014 CCAFS was one of four programmes tasked with developing a comprehensive results-based management approach for R4D. Accordingly CCAFS developed an approach to implementing RBM focusing on outcome delivery (Figure 1). The theory of change defines several activities, such as developing the impact pathways for thematic research and regional work, trialing RBM with a subset of projects, training key partners in building impact pathway, and analytical systems support. These led to tangible outputs such as facilitation guidelines (CCAFS, 2015a), a RBM MEL strategy (CCAFS, 2015b), and an online platform. This involved engagement with key next-users such as programme partners, with the intention that these outputs would be both useable and an incentive to overcome existing barriers in the system and as such would facilitate changes in current practices via proactively changing organisational norms. For example, project leaders were trained in designing their projects from a demand driven and outcome focused perspective. They needed to ensure that the research outputs would be enabling and incentives to support the practice changes that are required to achieve positive impact through their projects.

CCAFS started life using a logframe approach (LFA), but it became increasingly clear that the programme's vision of contributing towards development outcomes increasingly required an approach that acknowledged the importance of stakeholder engagement and capacity development. While a wide range of MEL approaches and methodologies with an outcome

[1] *CGIAR is a global partnership that unites organisations engaged in research for a food secure future.*

focus exist, none provides a blue-print solution. To adapt these approaches to a new programme, the right mix of elements needs to be selected, creating a conceptual framework in support of the programme's specific theory of change and MEL requirements. Springer-Heinze et al. (2003) advocate a holistic approach to impact evaluation and programme monitoring with quantitative and qualitative elements, based on an impact pathway that can accommodate different stakeholder views, allows for reflection and emphasises institutional capacity. Mixed methods provide opportunities to address the respective shortcomings of any single method as applied in practice.

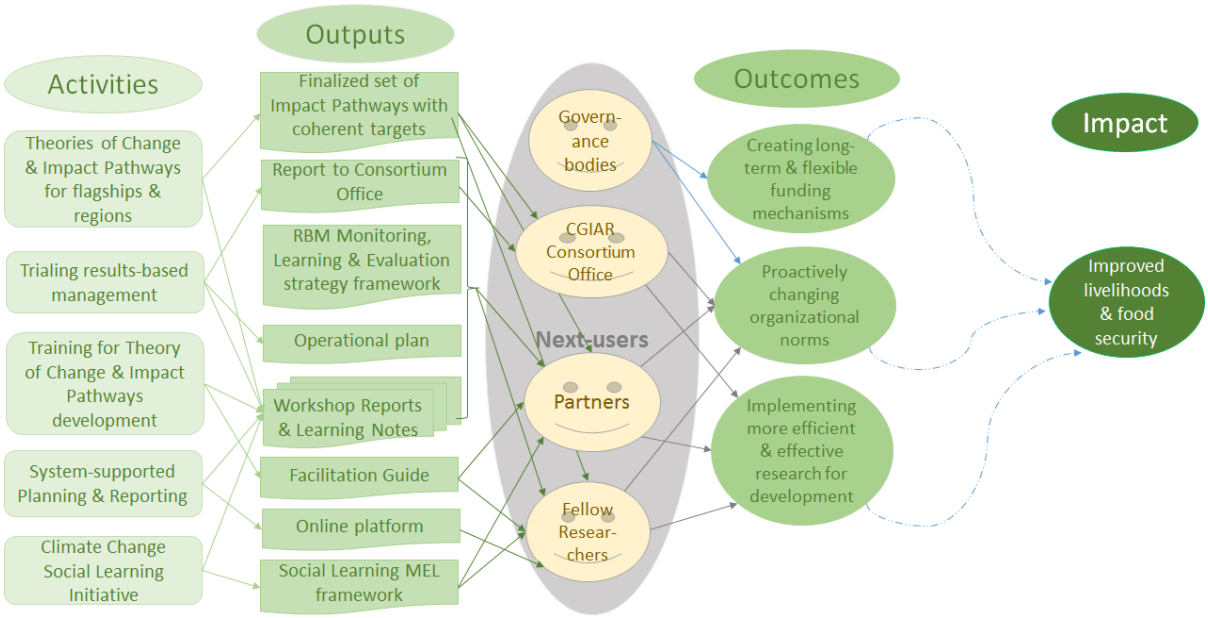

Figure 1. CCAFS' Theory of Change for its RBM approach and components

Findings and Analysis

Moving beyond the logframe approach

A logframe approach (LFA) has been widely used for project management; it adheres to a relatively rigid framework. It tends to prescribe a hierarchy of objectives converging on a single goal, a set of measurable and time-bound indicators of achievement, checkable sources of information, and assumptions of other impinging factors (Gaspar, 2000). In R4D the assumption is that development agencies, communication units, ministry staff and other people who could use the findings are able to source the scientific evidence, understand it, know how to implement and apply it, and convey this to people who they think need them. While this has been a useful approach, it is debatable whether it is entirely suitable for ensuring the use of research results and their translation into outcomes (Crawford & Bryce, 2003). The LFA does not pay enough attention to involving key stakeholders and their networks to achieve impact, provide managers with information to learn and report to donors, and establish a research framework to examine the change processes that projects seek to initiate (Douthwaite et al., 2008).

In line with donor requirements, CCAFS initiated its programmatic management approach on the basis of a logframe in 2009. Annual milestones were defined that were largely focused on producing scientific outputs and evidence of their achievement, which would then lead to

developmental impact. CCAFS has gone through several iterations of the logframe that was employed for planning and reporting (CCAFS, 2015c). In 2010 a limited version was used (CCAFS, 2010) while more elements were added in the following years. Planning and reporting elements were predetermined to some extent by requirements from CGIAR, though for internal purposes additional elements were added in response to the limitations that were identified from year to year.

Trialing results-based management with theories of change in CCAFS

In 2013, CCAFS's portfolio expanded to include project work in two new target regions and opportunities arose to implement and test a theory of change approach (Jost & Sebastian, 2014; Jost et al., 2014a). A new portfolio of six multiannual regional projects was set up and these were tasked with designing their approaches using a theory of change approach within a results-based management trial (Schuetz et al., 2014a).

There is no single definition of a theory of change, and no set methodology, as the approach assumes flexibility according to its respective user needs (Vogel, 2012). A theory of change provides a detailed narrative description of an impact pathway (a logical causal chain from input to impact, see Figure 2) and how changes are anticipated to happen, based on assumptions by people who participated in describing these trajectories. As such they provide an ex-ante impact assessment of a programme's anticipated success. Theory of change is at its best when it combines logical thinking and critical reflection; it is both process and product (Vogel, 2012).

RBM builds on the same logical causal chain as the LFA but is more explicit about output-use. Within R4D output-use refers to strategies that directly engage the next-users in the research process, e.g. through stakeholder platforms and user-oriented communication products. At the turn of the century, many development agencies and donors, including USAID, IDRC, UNDP and the World Bank, reformed their performance management systems and M&E approaches, to a RBM approach (Binnendijk, 2000; Bester, 2012; Mayne 2007a and 2007b). These experiences with RBM have informed further development of the approach.

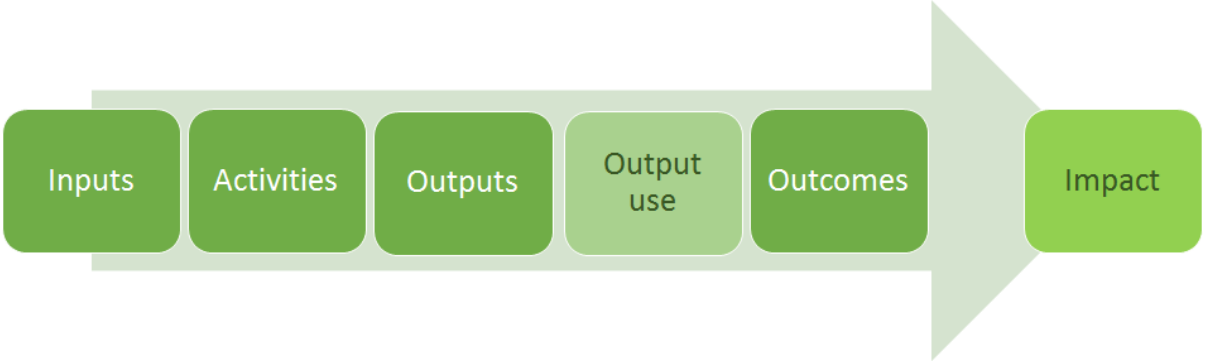

Figure 2. Theory of change logical causal chain

It is a particular challenge to show that R4D contributes to the desired behavioral changes (i.e. outcomes), that enable long-term positive impacts, as it requires qualitative monitoring rather than dealing with quantitative means of measuring alone (Young & Mendizabal, 2009; Springer-Heinze et al., 2003). Evaluators generally agree that it is good practice to first formalise a project's theory of change, and then monitor and evaluate the project against this 'logic model' (e.g. Chen, 2005). The theory of change is a mental model made explicit by

involving as many people as possible in its design. Key principles of the Participatory Impact Pathways Analysis method also include reflecting on these models, regularly validating the assumptions that were made and adjusting programme management accordingly (Douthwaite et al., 2013).

Within the CCAFS RBM, this theory of change approach to project planning helped position the R4D agenda further along the impact pathway (Schuetz et al., 2014a). Projects expanded their skill sets by bringing on board other partners that would help implement output-to-outcome strategies and thus create more clearly defined causal logical chains (Figure 2; Schuetz et al., 2014b and 2014c). This is not to take over the work of development agencies, but it is to ensure that research findings are adequately contextualised to be a good fit for the demand and given purpose. The CCAFS RBM projects have thus challenged the common thinking that good science and publications are enough and by themselves will lead to impact - rather, they are necessary but not sufficient.

Building capacity and learning within the programme for a theory of change approach
The RBM trial project teams were thrown in at the deep end. Used to a more traditional LFA, they were tasked with shifting to a theory of change and learning-based approach for planning their projects within the trial. It was quickly apparent that capacity to plan projects using this new approach had to be built within CCAFS and its partners.

Using theory of change approaches within R4D requires building the capacity of scientists to do research differently and work with non-research partners for impact, but also of institutions to facilitate such a shift. Several authors highlight the importance of building capacity for institutional learning (Hall et al., 2003; Horton & Mackay, 2003; Eade, 1997; Springer-Heinze et al., 2003). Johnson et al. (2003) show that participation of non-research stakeholders early on in the research process is important, as it can inform institutional learning in research organisations to change priorities and practices. It can also enhance the relevance of agricultural technologies and the capacity of these stakeholders to design their own action research processes (Johnson et al., 2003). Horton and Mackay (2003) outline the links between M&E, learning and institutional change and highlight the importance of institutional learning as a means of developing the capacities of the organisation and of individual researchers, and empowering non-research partners as key stakeholders in the process.

CCAFS worked with expert facilitators and trainers from Participatory Impact Pathways Analysis to implement a one week training course on using theory of change for project and programme planning (Alvarez et al., 2014). Initially about 20 participants were chosen strategically so that capacity would be available in CGIAR Centers at the point in time when CGIAR proposals would need to be developed following theory of change principles. In addition to project representatives, CCAFS science officers representing all themes and regions participated to build in-house capacity. The training, in combination with theory of change facilitation guides (version 1: Jost et al., 2014b; version 2: Schuetz et al., 2014d) and learning notes (CCAFS, 2015a), helped highlight the opportunities (and constraints) of rolling out RBM to a whole R4D programme. An online community of practice (and wikispace) was established and allowed for continued documentation and exchange of experiences.

CCAFS' Results-based management trial - insights from researchers and partners

CCAFS' approach to RBM is centred on adaptive management, regular communications between programme and projects and facilitated learning within projects. Besides periodic virtual meetings, trial participants were surveyed for a more in-depth and standardised reflection and for capturing lessons and achievements from their experience (Schuetz et al., 2014b and 2014c). These lessons and the programmatic perspective by the CCAFS Management Committee were documented in reports (Thornton et al., 2014b, CCAFS annual progress report) and a series of learning notes (CCAFS, 2015a). The approach to developing the impact pathways was simplified over time, mostly by reducing the number and type and number of indicators and level of complexity so that the wider group of people who were expected to work with them would continue to buy-in to the approach (Schuetz et al., 2014d). For example we focused on indicators at the outcome level and dropped any further development of detailed output progress indicators.

There are many people within CGIAR Centres and CCAFS partners who are willing to take on the challenge of developing new ways of collaborating and working beyond delivering outputs towards outcomes (Schuetz et al., 2014b). After one year of the trial, projects had made considerable progress. Project leaders and teams, for example, became more reflective in their project planning and reporting, identifying opportunities to adapt to new insights and questioning users, use and usefulness of research outputs to facilitate and encourage development outcomes. Another area where we saw some progress was the improvement of narrative qualitative descriptions of progress towards, and achievements of, outcomes. However, making fundamental shifts in the way of working takes time and (initially at least) additional resources. It requires iterative and continuous processes. Staffing, or the profile of project team members and project team composition are emerging as key factors for success. Project staff have acknowledged that they may require additional skills beyond disciplinary expertise, such as skills in coordination, facilitation, engagement, communications, participatory and learning-oriented M&E. We are exploring how additional support can be provided in areas such as engaging with stakeholders and using RBM.

Rolling out results-based management for CCAFS as a whole

Opportunities for changing the programmatic approach to project planning, implementation and MEL emerged when CCAFS was approaching the end of its first phase in 2014. Theories of change were developed and defined for all four research and five regional programmes in CCAFS as a first step in putting together the new programme portfolio (Schuetz et al., 2014e). Figure 3 provides an illustration of one research theme's theory of change component with its regional elements, indicators and outcome targets.

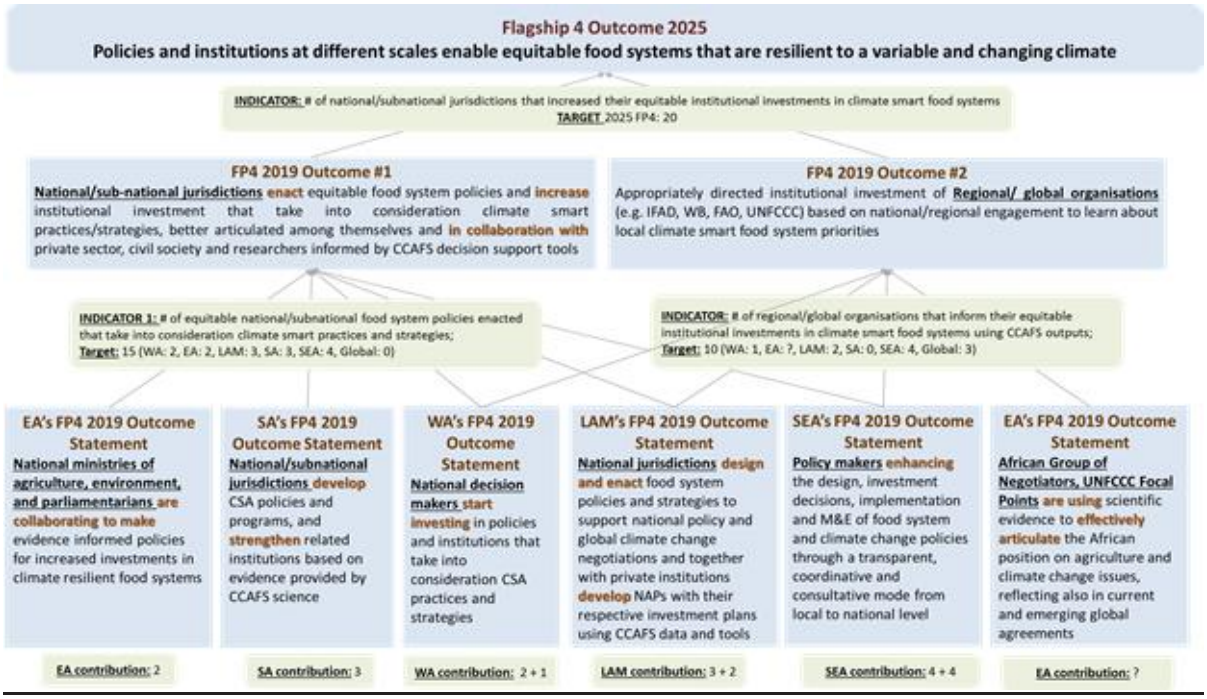

Figure 3. Illustration of a CCAFS thematic IP component *(drawn from the Flagship Programme on Policies and Institutions for Climate Resilient Food Systems)*

While it took a considerable amount of effort, the iterative development of the CCAFS theory of change was done with a view to attempting to be as efficient as possible. At the start most interactions were virtually facilitated and built on previous engagement and regional priorities. For completion, key next-users and stakeholders participated in five regional face-to face meetings (Schuetz et al., 2014e and 2014f).

The theory of change development and facilitation process and the guidance documentation were revised to make them leaner, more contextualised and easier to implement (Schuetz et al. 2014d). The theory of change building process is one key component in the CCAFS MEL system that was developed to support the new approach in a comprehensive manner (CCAFS, 2015b).

Building on the above, a CCAFS Monitoring, Learning & Evaluation Strategy was developed (Schuetz et al. 2014g), with the overall goal of developing an "evaluative culture" within CCAFS that encourages self-reflection and self-examination, seeks evidence, takes time to learn and encourages experimentation and change so that MEL becomes an integrated mechanism. The strategy includes a conceptual framework, guided by overall programme principles for partnership, engagement and communications in a modular way, to best meet the demands of the programme as a whole, its projects and the wider CGIAR system (Thornton et al., 2014c). For project implementation this led to some built-in and on-going monitoring and documentation of project activities on the outputs use, i.e. on engagement, partnerships and communication. Some elements are prescribed by CGIAR governance bodies, including the carrying out of baselines, independent impact assessments and periodic external evaluations. Programmatic flexibility exists within the day-to-day operational MEL, as a system is required that allows enough flexibility and adaptability to be applied to the different types of projects and programmes.

Implications for policy, practice and research

Working with theories of change and impact pathways has major implications for MEL (Schuetz et al., 2015). It implies a move to contribution rather than attribution, i.e. to show our contribution we acknowledge the role and inputs of partners and other actors both in achieving outcomes and in providing evidence for those outcomes i.e. lots of factors caused the change, rather than trying to claim a change due to our intervention alone (AIDLEAP, 2015, CCAFS glossary). Building in triple-loop learning[2] helps enable people to distil key lessons from reflection (hindsight) and make best use of those lessons (insights) for future planning (foresight) and can make a major contribution to reflection and to supporting adaptive management, so that project teams can better deal with uncertainty. At the same time, not everything can be measured; this highlights the need for narratives that can complement and support more quantitative information.

As part of creating a programme-enabling environment, CCAFS embraced the three-thirds principle, whereby one third of effort is spent on engaging with partners to decide what needs to be done and how; one third on doing the actual research, often in partnership; and one third on sharing results in appropriate formats and strengthening capacity of next users to utilise the research to achieve outcomes and impact. This implies different budgeting and funding structures, so that appropriate levels of resources are allocated to capacity building, communications and engagement with the wide range of different partners likely to be needed (CCAFS, 2014). These elements need to be budgeted for explicitly within a project life-cycle, rather than as an after-thought. At the same time, there is still much work to be done on how to monitor outcomes effectively, evaluate the real share of contribution towards the observed change and assess value for money. Similarly, delivery of outcomes, especially at scale, may take time for research-for-development programmes. Longer funding cycles could be expected to facilitate this considerably.

The CCAFS experience has highlighted several operational principles for RBM implementation. First, there is a need to focus on people and users, on utilising M&E as a tool to help achieve outcomes and on accountability - it is the people within organisations that make behavioural and practice changes happen. Second, there should be an emphasis on learning through M&E activities, i.e. an M&E system for R4D needs to integrate both monitoring and impact evaluation real time. Robust knowledge needs to be generated that can feed into developmental policy and investment decision making and this in turn requires a cumulative and catholic approach to choice of impact assessment methods at different levels (Maredia, 2009). Third, adaptive management needs to be encouraged as a key element of RBM. As a tool that is based on learning processes it can improve long-run management outcomes. The challenge in using it is to find the balance between gaining knowledge to improve management in the future and achieving the best short-term outcome based on current knowledge. Fourth, planning, reporting and evaluation procedures need to be as simple as possible while still providing (most of) the information needed for effective and timely programme management. In this the development and implementation of an online platform was an investment to support and guide programme participants in their contribution to the

[2] *Triple loop learning is a series of learning steps, from receipt of information (single loop), to reflecting on what activities will be more effective (double loop), through to behaviour change as a result of that reflection by multiple stakeholders (triple loop) (Carlile et al., 2013).*

programme's results-based management system and developed their capacity at the same time.

Sharing findings along the way is a good way to foster the inclusive involvement of as wide a range of stakeholders as possible in project planning and implementation. Encouraging researchers to get early drafts of findings out to potential users for feedback from early on is one way to build a learning culture and to encourage open-mindedness. Lessons have come from surveys (e.g. with the trial project teams) and via collective reflections and evaluations (e.g. from the workshop series with participants and the programme management team).

Rigid application of just one specific approach will most likely not work. Whether it is the adoption of a technology, an M&E methodology, a learning approach or a scientific result, it is often not the whole package that is attractive to users but specific pieces. We need to allow users to cherry pick while ensuring that the relevant linkages remain intact so that the context is not lost for others who may want other cherries.

Solutions that are good enough rather than optimal. In many domains of knowledge and practice there is no best practice or option, particularly when the problem is complex and resources are constrained. CCAFS made considerable changes once it had started to implement an approach based on theory of change and impact pathways and in time moved towards a leaner and simpler model One of the key messages from the RBM trial process was the need for systems that cover most users' needs, rather than aiming for completeness that could add unwanted complexity.

Addressing tensions across scale. CCAFS is still in the process of embedding theories of change for the different organisational units of the programme, to provide a flexible framework that allows for aggregation of output, outcomes and targets across the different units. For example, targets need to be framed locally with users and beneficiaries and voiced in such a way as to allow the flexibility to deal with uncertainty and emerging priorities and opportunities. New investments of time and effort may be needed to identify and work with non-traditional partners to promote behavioural change in shared IPs.

Providing value for money. Many members of the donor community now require that grantees demonstrate value for money. For instance, the Deutsche Gesellschaft für Zusammenarbeit states that its "work is systematically geared towards results, the yardstick by which we measure the success of our work. We want to help achieve tangible positive changes on the ground" (GIZ, 2015). Some have critiqued the whole notion of payment by results as applied to development and research-for-development on the basis that it provides perverse incentives that actually diminish cost-effectiveness (see Chambers, 2014). As noted above, there is much work still to do on appropriate measurement mechanisms, but this does not diminish the need to demonstrate accountability.

Balancing science and outcomes. Research is often curiosity-driven, and traditional indicators of success centre on peer-reviewed publications in high-profile academic journals. In today's highly competitive research environment another crucial success factor relates to fundraising: the ability to write and win competitive research proposals. Neither of these motivations for research is guaranteed to deliver development outcomes. For CGIAR and its research programmes it is still early days, but preliminary results suggest that 'successful

RBM' relates to effective and efficient research leading to outcomes, with a minimum of perverse incentives (Thornton et al., 2015). The building of an impact pathway with a narrative theory of change forces researchers to give some thought to what lies between solid science, great technologies and their positive developmental impact. A mix of an outcome-focused theory of change with people and partners at the core and a RBM approach that allows us to monitor, reflect, evaluate and learn, are key elements for a programmatic MEL strategy - coupled with great science.

Conclusions

Requests by donors for a move towards outcome-oriented research programmes are having considerable impact on the way in which research is conceived, planned, implemented and evaluated. A key requirement for such work is flexibility - the flexibility to adjust so that the outcome orientation works as a support mechanism and enabler rather than a one-size-fits-all straitjacket without space for innovation, serendipity and creativity. Results-based management offers this kind of flexibility. The shift to a R4D approach based on theory of change is fostering massive change, much of it for the better in our view. However, it also comes with considerable challenges. Defining the necessary changes and developing new processes and mechanisms, including monitoring and more built-in evaluation in real time, requires effort and resources, which are often grossly underestimated and inadequately planned for. Some of these challenges arise because of the nature of research: the results are not known from the start, unlike in engineering where the outcomes are generally much less uncertain. Another challenge is that CGIAR is a R4D and not a development organisation, and it is still striving to balance the need to do great science with the need for impact. We need to avoid the results-based focus being to the disadvantage of the science and development being seen as being in competition with the science. Rather, they need to be seen as complementary, enabling and liberating.

References

AIDLEAP. 2015. Contribution vs Attribution – A Pointless Debate, November 23, 2015.

Bester, A. 2012. Results-Based Management in the United Nations Development System: Progress and Challenges. A report prepared for the United Nations Department of Economic and Social Affairs, for the Quadrennial Comprehensive Policy Review Final Report.

Binnendijk, A. 2000. Results-based Management in the Development Co-operation Agencies: a review of experience. DAC Working Party on Aid Evaluation Report. OECD.

Carlile, L., Ballantyne, P., Ensor, J., Foerch, W., Garside, B., Harvey, B., Patterson, Z., Thornton, P.,& Woodend, J. 2013. Climate Change and Social Learning (CCSL): supporting local decision making for climate change, agriculture and food security, CCSL Learning Brief No.1, May 2013 Copenhagen, Denmark: CGIAR Research Programme on Climate Change, Agriculture and Food Security (CCAFS).

CCAFS, 2015a. How we Work: Impact Pathways.

CCAFS, 2015b. Planning, Monitoring, Learning and Evaluation. ccafs.cgiar.org/planning-monitoring-learning-and-evaluation.

CCAFS, 2015c. Key Strategic Documents.

CCAFS, 2014. Annual Report 2014.

CCAFS, 2010. Activity Workplans and Budgets for 2010.

CGIAR, 2015. www.cgiar.org

Chambers, R. 2014. Perverse payment by results: frogs in a pot and straitjackets for obstacle courses. Blog.

Chen, H.T. 2005. Practical Programme Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness. CA: Sage Publications.

Crawford, P., Bryce, P. 2003. Project monitoring and evaluation: a method for enhancing the efficiency and effectiveness of aid project implementation. International Journal of Project Management 21(5): 363–373.

Douthwaite, B., Kamp, K., Longley, C., Kruijssen, F., Puskur, R., Chiuta, T., Apgar, M., & Dugan, P. 2013. Using theory of change to achieve impact in AAS. AAS Working Paper.

Douthwaite, B., Alvarez, S., Thiele, G., & Mackay, R. 2008. Participatory Impact Pathways Analysis: a Practical Method for Project Planning and Evaluation. Colombo, Sri Lank: CPWF. Eade, D. 1997. Capacity-Building: An Approach to People-Centred Development. Oxford: Oxfam.

Earl, S., Carden, F., Smutylo, T. 2001. Outcome Mapping, Building Learning and Reflection into Development Programmes. IDRC Publications.

Eyben, R. 2013. Gasper, D. 2000. Evaluating the 'logical framework approach' towards learning-oriented development evaluation. Public Administration & Development 20:17-28.

FAO, 2014. The State of Food Insecurity in the World 2014. FAO, Rome.

GIZ, 2015. Results-Based Monitoring.

Hall, A., Sulaiman, V.R., Clark, N., & Yoganand, B. 2003. From measuring impact to learning institutional lessons: an innovation systems perspective on improving the management of international agricultural research. Agricultural Systems 78(2): 213–241.

Hartmann, A., & Linn, J.F. 2008. Scaling Up: A Framework and Lessons for Development Effectiveness from Literature and Practice. Wolfensohn Centre for Development, Working Paper 5. Brookings Institute.

Horton, D., & Mackay, R. 2003. Using evaluation to enhance institutional learning and change: recent experiences with agricultural research and development. Agricultural Systems 78: 127-142.

Johnson, N., Lilja, N., & Ashby, J.A. 2003. Measuring the impact of user participation in agricultural and natural resource management research. Agricultural Systems 78: 127–142.

Jost, C., Sebastian, L., Kristjanson, P., & Förch, W. 2014a. Lessons in theory of change: CCAFS Southeast Asia Research for Development Workshop. CCSL Learning Brief No. 8. CCAFS, Copenhagen, Denmark.

Jost, C., Alvarez, S., & Schuetz, T. 2014b. CCAFS Theory of Change Facilitation Guide. Copenhagen, Denmark: CGIAR Research Programme on Climate Change, Agriculture and Food Security (CCAFS).

Jost, C., & Sebastian, L. 2014. Workshop on Mapping out a CCAFS R4D Agenda and Strategy for Southeast Asia. CCAFS, Copenhagen, Denmark.

Maredia, M.K. 2009. Improving the proof - evolution of and emerging trends in impact assessment methods and approaches in agricultural development. IFPRI Discussion Paper. IFPRI: Washington DC.

Mayne, J. 2007a. Best Practices in Results-Based Management: A Review of Experience. A Report for the United Nations Secretariat Volume 1: Main Report.

Mayne, J. 2007b. Challenges and lessons in implementing results-based management. Evaluation 13(1): 187-109.

Pachico, D. and Fujisaka, S. (Eds. 2004. Scaling up and out: achieving widespread impact through agricultural research (Vol. 3). CIAT, Cali.

Raitzer, D., & Kelly, T. 2008, Benefit-cost analysis of investment in the international agricultural research centres of the CGIAR. Agricultural Systems 96: 108-123.

Schuetz, T., Förch, W., & Thornton, P. 2015. CCAFS Reporting and Evaluation in a Results-Based Management Framework. CCSL Learning Brief No 15. Copenhagen, Denmark: CGIAR Research Programme on Climate Change, Agriculture and Food Security (CCAFS).

Schuetz, T., Cramer, L., Foerch, W., Jost, C, Alvarez, S., Thornton, P., & Kristjanson, P. 2014a. Summary for the CCAFS Flagship 4 Projects Kick-off Meeting 28-29 January 2014: Result-based Management Trial.

Schuetz, T., Schubert, C., Thornton, P., Förch, W., & Cramer, L. 2014b. Lessons and Insights from the CCAFS Results-Based Management Trial. Summary from CCAFS Flagship 4 RBM trial projects survey 2014. CGIAR Research Programme on Climate Change, Agriculture and Food Security. Report

Schuetz, T., Förch, W., Schubert, C., Thornton, P., & Cramer, L. 2014c. Lessons and Insights from CCAFS Results-Based Management Trial. CCSL Learning Brief No 12. CGIAR Research Programme on Climate Change, Agriculture and Food Security (CCAFS). Copenhagen: Denmark.

Schuetz, T,, Förchm W., & Thornton, P. 2014d. Revised CCAFS Theory of Change Facilitation Guide. Copenhagen, Denmark: CGIAR Research Programme on Climate Change, Agriculture and Food Security (CCAFS).

Schuetz, T., Förch, W., Thornton, P., Wollenberg, L., Hansen, J., Jarvis, A., Coffey, K., Bonila-Findji, O., Loboguerrero Rodriguez, A.M., Martínez Barón, D., Aggarwal, P., Sebastian, L., Zougmoré, R., Kinyangi, J., Vermeulen, S., Radeny, M., Moussa, A., Sajise, A., Khatri-Chhetri, A., Richards, M., Jost, C., & Jay, A. 2014e. Lessons in Theory of Change from a Series of Regional Planning Workshops. CCSL Learning Brief No 11. Copenhagen, Denmark: CGIAR Research Programme on Climate Change, Agriculture and Food Security (CCAFS). Link:

Schuetz, T., Förchm W., Thornton, P., Wollenberg, L., Hansen, J., Jarvis, A., Coffey, K., Bonilla-Findji, O., Loboguerrero Rodriguez, A.M., Aggarwal, P., Sebastian, L., Zougmoré, R., Kinyangi, J., Jost, C., & Jay, A. 2014f. Lessons in theory of change from a series of regional planning workshops. CCAFS, Copenhagen, Denmark. Full report.

Schuetz, T., Förch, W., & Thornton, P. 2014g. CCAFS Monitoring and Evaluation Strategy. Copenhagen, Denmark: CGIAR Research Programme on Climate Change, Agriculture and Food Security (CCAFS).

Springer-Heinze, A., Hartwich, F., Henderson, J.S., Horton, D., & Minde, I. 2003. Impact pathway analysis: an approach to strengthening the impact orientation of agricultural research. Agricultural Systems 78(2): 267–285.

Thornton, P.K., Vermeulen, S.J., & Campbell, BM. 2015. Can research be results-based? A response to Chambers‘ Blog

Thornton, P.K., Ericksen, P.J., Herrero, M., & Challinor, A.J. 2014a. Climate variability and vulnerability to climate change: a review. Global Change Biology 20(11): 3313-3328.

Thornton, P., Schuetz, T., Förch, W., & Campbell, B. 2014b. The CCAFS Flagship 4 Trial on Results-Based Management: Progress Report

Thornton, P., Förch, W., Schuetz, T., Barahona, C., Abreu, D., & Tobón, H. 2014c. Guiding principles or "propositions" on monitoring and evaluation in CCAFS. Copenhagen, Denmark: CGIAR Research Programme on Climate Change, Agriculture and Food Security (CCAFS).

UN., 2015. Transforming our World: the 2030 Agenda for Sustainable Development.

UNDG., 2011. Results-Based Management Handbook: Harmonizing RBM Concepts and Approaches for Improved Development Results at Country Level. United Nations Development Group, October 2011.

Vogel, I. 2012. 'Theory of Change' in International Development. Review Report for the UK Department of International Development.

WWAP., 2012. The United Nations World Water Development Report 4: Managing Water under Uncertainty and Risk. Paris, UNESCO.

Young, J., & Mendizabal, E. 2009. Helping researchers become policy entrepreneurs. ODI Briefing paper No. 53.

Adaptive management intentions with a reality of evaluation: getting science back into policy

Hasselman, L.

Institute for Governance and Policy Analysis, University of Canberra, Australia

Abstract: Adaptive management was initially proposed to address system uncertainty in natural resource management. In theory, adaptive management integrates scientific experimentation in policy planning and implementation to discover and gather knowledge from across a system's stakeholders. It systematically tests hypotheses with the results redirecting or improving policy, applying a paradigm of scientific problem solving.

This paper uses a case of water management in Australia's Murray-Darling Basin. Water reform has been contentious as government attempts to reconcile historical over allocation of water to irrigation with the use of water to protect and restore wetlands of international biodiversity significance. In areas scientific knowledge of the system is either imperfect, incomplete or system responses are unpredictable. In this case there are legislative requirements for both adaptive management and evaluation. Evaluation looks to achievement of policy objectives, as determined through monitoring of system response and value judgements, in a structured framework of action, outputs, outcomes and objectives.

The intentions for adaptive management are compared to the reality, as determined through legislation, public speeches, government reports and semi-structured interviews with government policy makers and implementers. The findings demonstrate contradiction between intent and reality, with adaptive management subsumed by evaluation. The loss of adaptive management as a distinct concept is seen as a loss of science and discovery from the policy process. Despite intentions for adaptive management, the dominance of evaluation is discussed as limiting innovation, a 'muddling through' process of improvement and meeting political and accountability needs.

Key words: Adaptive management, evaluation, science, uncertainty, politics, conflict, accountability

Introduction

Adaptive management is now widely accepted as a necessity in the management of natural resources, such as water, soil and biodiversity (Allan, 2009; Dovers, 2009; Pahl-Wostl, 2009; Pahl-Wostl et al., 2013). In theory, adaptive management integrates scientific experimentation in policy planning and implementation in order to develop new knowledge and gather knowledge from across a system's stakeholders (Walters & Holling, 1990). It uses a systematic process of hypothesis testing with the resulting scientific discoveries redirecting or improving policy, applying a paradigm of scientific problem solving. However empirical examples of adaptive management are scarce in practice (Eberhard et al., 2009), and the very meaning of adaptive management remains debated (Allen et al., 2011; Scarlett, 2013).

In the 1990's, with the lack of evidence of adaptive management, Lee suggested that adaptive management needed to be included in legislation to ensure it actually occurred (Lee, 1993) and these calls for prescription in legislation have continued since (Dovers, 2009). The argument is that adaptive management as a legal requirement will provide the additional impetus for organisations to overcome implementation challenges. In addition, a legislated definition of adaptive management could be expected to provide clarity and direction to what has become a confused and misunderstood enigma. The reasons for adaptive management being theoretically championed whilst being rare in practice are discussed in this article.

In 2012 adaptive management became a defined term in Australian water legislation in the Murray-Darling Basin Plan (Basin Plan), providing a fit case to examine the differences in intention and reality of adaptive management. This article uses the case of the Basin Plan to test the argument that prescription in legislation is required to overcome the challenges to adaptive management.

First, a brief literature review of adaptive management is provided, followed by a description of the method used and an introduction to the case. The intentions for adaptive management in the Basin Plan, as provided by legislation and policy documentation, are compared to the reality as determined through semi-structured interviews with government policy makers and implementers and published government reports. In analysis and discussion of the results it is demonstrated that the true barrier to adaptive management is not the absence of legal mandate, but confusion of adaptive management with evaluation, with this further marginalising science from policy development and implementation.

Literature review

Adaptive management

Adaptive management emerged from the field of ecology in the late 1970s. Over time its meaning has been debated, with adaptive management referred to as 'experimental management' (Walters, 1997), 'learning by doing' (for example, see Schreiber et al., 2004) and 'structured decision making' (Allen & Gunderson, 2011; Gunderson & Light, 2006). Forms or types of adaptive management distinguish between active adaptive management, with multiple hypothesis testing, statistically sound experimentation and technical modelling, and passive adaptive management that looks to observation and response in single treatments. The label of passive adaptive management was first used by Walters & Holling (1990) but does not imply a lack of effort or resourcing requirements. It is planned, participatory and requires monitoring and analysis to test a single best hypothesis with a single treatment. Over time the role of partnerships to bring together socially-held knowledge has been increasingly emphasised, further highlighting the misnomer of the label 'passive'.

Active and passive adaptive management both emphasise systematic and planned hypothesis testing, involve stakeholders working across knowledge disciplines and remain strongly motivated by the need to increase knowledge of system function and address uncertainty (Hasselman, forthcoming). However, there are three broadly recognised types of uncertainty (Berkes, 2007; Brugnach et al., 2008). This includes uncertainty that results from imperfect knowledge (undiscovered science), incomplete knowledge (knowledge that cannot be held by an individual but is collectively held across stakeholders) and unpredictability (unforeseeable futures with unknown society and ecosystem responses). In addition to these three types of

uncertainty, Pagan and Crase (2005) also note unforeseen changes to community preferences and government objectives over time.

Active adaptive management seeks to reduce imperfect knowledge with experimentation to discover new knowledge and determine the optimal solution (Walters & Holling, 1990), viewing knowledge as absolute and uncertainty as something to remove. In comparison passive adaptive management seeks responsiveness to unpredictability. Management is seen as experimentation with socially-held knowledge, applying an approach that accepts unpredictability (Berkes 2007; Brugnach et al., 2008; Huitema et al., 2009).

The context to which adaptive management is applied is important; influencing which form of adaptive management is most suitable. These differences mean that adaptive management is hardly a single thing or panacea, but a pluralist concept and practice. The evolving history of adaptive management with its varying emphasis on experimental management, learning by doing and structured decision making, along with the varied contexts to which policy makers and implementers have sought to apply adaptive management has contributed to confusion about the meaning of adaptive management (Allen et al., 2011; Loftin, 2014). Following an extensive review, Hasselman (forthcoming) has proposed a definition of adaptive management that acknowledges the different types of underlying uncertainty that may be motivating adaptive management. This definition accepts the need for pluralism and context specific application and can be applied to the different forms of adaptive management.

> *Adaptive management is a systematic process for improving policy and its implementation. It seeks to address at least one type of uncertainty with varying emphasis on experimentation to discover new knowledge; deliberative processes to engage multiple perspectives in decision making; and monitoring of outcomes and changes with responsive adjustment of decisions and implementation.*

In this definition adaptive management remains a scientifically based activity to increase collectively held knowledge and experience, in order to make better management decisions. The essence of adaptive management remains applied science with the learnings used to gain ecological outcomes. The ability to change decisions based on new information is just as critical to adaptive management as the ability to gain new knowledge or gather knowledge.

Evaluation

In natural resource management there are other approaches to learning, such as evaluation that play a significant role in policy implementation and development. Evaluation has been described as an appraisal or systematic assessment of merit and/or worth (Guba & Lincoln, 2001). It is considered applied social research that is both transdisciplinary and an autonomous discipline (Scriven, 2013). The purpose of evaluation has been variably identified as performance improvement, organisational learning, accountability, learning about persistent social problems and how to address them, informing decision making and to democratise decision making (Alkin, 2013; Greene, 2013).

Evaluation involves evidence collection, often referred to as monitoring, and a process of applying judgement to an evaluand; or the subject of the evaluation. Evaluations can be formative to improve a programme with feedback gained on processes and factors that may affect achievement of objectives, done during a programme or policy implementation. In comparison summative evaluations seek to determine a programme's merit and worth, taken as the extent to which objectives have been achieved and the contextual factors that have

affected the results; often done after completion of a programme or policy (Patton, 2013). Regardless of the timing, both formative and summative evaluations look to achievement of stated policy or programme objectives.

Scriven (2013) argues that a widely held misunderstanding is "that the difference between evaluation and research is that *research* is aimed at the acquisition of new knowledge whereas *evaluation* is aimed at developing information for decision making." Scriven instead draws a distinction between evaluative research and non-evaluative research, focusing on the distinction of value judgements used in evaluation.

Evaluation in Australia

In Australia, evaluation has been shaped strongly by public administration reforms in the 1980s, including the 1988 Evaluation Strategy (Rogers & Davidson, 2013). Australian evaluations have been described as concentrated on ongoing management of programmes, commonly using theory driven approaches such as programme theory or programme logic, with emphasis on stakeholder participation (Rogers & Davidson, 2013). Programme theory and programme logic approaches to evaluation work within frameworks of causal pathways that articulate how programme and project activities lead to achievement of desired outcomes, with these in turn leading to achievement of objectives (Funnell, 2000). Assumptions underpinning the causal relationships may be stated, with monitoring and evaluation seeking to test these assumptions. In the testing of assumptions, causal pathways are confirmed, achievement or contribution to achievement of outcomes is deduced and eventually objectives are proven as being met.

Inclusive participation in the evaluation process from planning through to final judgement is seen as having a role in validity and credibility, not just to promote use and implementation of findings. The result is a close integration between evaluation and management with an expectation for programme evaluation as driven by performance improvement and accountability, with results often structured as outputs and outcomes, as evidenced by indicators.

The performance improvement and accountability view of evaluation is most notable in the Australian Government's Caring for our Country Programme, as commenced in 2007. This nationwide, large scale investment introduced mandatory requirements for monitoring, evaluation, reporting and improvement (MERI) and has significantly influenced government natural resource management since. MERI has become an accepted part of the evaluation lingo. MERI is defined as "*simple concepts that, when applied, help us understand what is being achieved and help identify possible improvements, for projects and programmes*" and evaluation specifically as "*analysing the monitoring data and assessing what it means. Based on this, informed judgements can be made about the success of a project or programme and improvements can be identified*" ('Monitoring, Evaluation, Reporting and Improvement Strategy', 2014). This has further established evaluation as performance management and accountability, as structured by generally linear programme logics of input-output-outcome level indicators. Evaluation is readily accepted as a part of good governance, serving a managerial mandate and providing accountability for the use of public resources.

Method

Case study – The Murray-Darling Basin

In 2012 adaptive management became a defined term in Australian legislation with the passing of the Murray-Darling Basin Plan (Basin Plan). At over 1 million square kilometres, or 14% of Australia's land mass, the Murray-Darling Basin is Australia's largest water catchment and river system (Figure 1). It supplies drinking water to over three million people, contains wetlands of national and international importance, including 16 Ramsar sites and a World Heritage site. The Murray-Darling Basin also contains 40% of Australia's farms and produces over a third of Australia's food, valued between $10 billion and $15 billion. Water reform has been contentious as government attempts to reconcile historical over allocation of water to irrigation with the use of water to protect and restore wetlands of international biodiversity significance. The following overview of governance arrangements of the Murray-Darling Basin is brief and limited to key points necessary for this paper. Others have elsewhere published more detailed accounts of the governance history, ecological and social challenges and conflict (Connell, 2007).

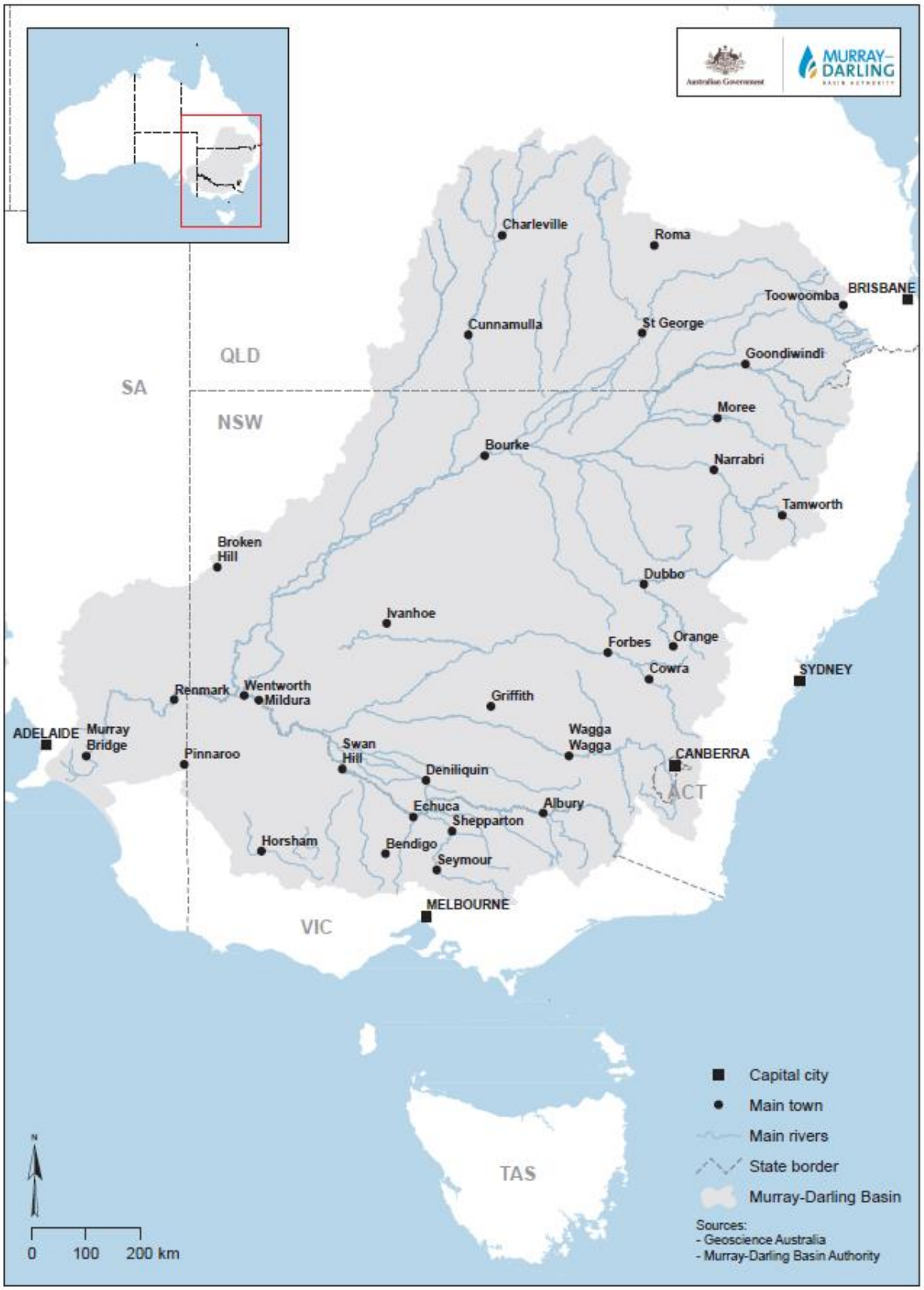

Figure 1. Murray-Darling Basin. *Source: http://www.mdba.gov.au/sites/default/files/cartographicmapping/8_Murray-Darling_Basin_Boundary.pdf, accessed 31/5/2016*

Currently water governance and management is done by six jurisdictions, being each of the four States and the Territory that the Murray-Darling Basin spans and the Commonwealth. The Basin Plan determines the maximum volume of water that can be sustainably extracted for urban, industrial and agricultural use (Sustainable Diversion Limits), provides the latest reform on water trading rules and sets a framework for the planning and use of environmental water. The Basin Plan is said to be based on the 'best available science', with this also recognising that there remain unknowns in the system. In parts of the Murray-Darling Basin, scientific knowledge of the system is either imperfect, incomplete or system responses are unpredictable. In particular, the ecological, social and economic responses to the Basin Plan are uncertain and remain points of contention and debate. To some the very idea of setting sustainable diversion limits for the Basin as a whole and assigning contributing volumes to parts of the Basin in this context of uncertainty is "folly" (Crase, 2012).

The State and Territory governments administer water licences that provide irrigators with a share of the water resource (also called entitlement), with the actual volume determined (share allocation) based on seasonal conditions. Water use and trade is governed through rules established in 36 regional scale Water Resource Plans, as developed by the State governments. These Water Resource Plans must align with the Sustainable Diversion Limits set by the Basin Plan and require accreditation by the Commonwealth. State and Commonwealth governments also hold water licences, as purchased from the market. This water is held by government in order to supply water to environmental assets such as wetlands and is referred to as environmental water. The use and trade of environmental water is governed by the States' Water Resource Plans and State and Commonwealth level Annual Environmental Watering Plan and Long-term Environmental Watering Plans.

Method

The research was qualitative, using document analysis and interviews. Document analysis included legislation, planning documents, published reports, policy statements and published speeches. Documentation provided insight into government intentions for adaptive management, outlining commitments, processes and projects underway.

Interviews were conducted with Commonwealth and NSW policy makers and implementers to understand the reality of adaptive management, under the framework set by the Basin Plan and associated governance arrangements. The interviewees included past government employees involved in the development of the Basin Plan and those currently involved in its implementation. Regional stakeholders with roles within the governance arrangements for developing and implementing the Basin Plan and associated State planning instruments were also interviewed. In total, 30 interviews were conducted, with this sample covering different aspects of water reform within the Murray-Darling Basin. Saturation point, when no new themes were emerging, was reached. The interviews were semi-structured and explored views and experiences of adaptive management in water management. The interviews included questions on the definitions of adaptive management, example cases of adaptive management in practice and challenges to adaptive management. The interviews were transcribed before being thematically coded using Nvivo software. The research takes a constructivist view and applies an inductive logic, with examples used to infer broader principles, meaning the research is descriptive.

Results

Intentions for adaptive management

The intentions for adaptive management have been determined from government documents, including the Basin Plan, reports published by the Murray-Darling Basin Authority and public speeches by the Chair of the Murray-Darling Basin Authority.

With the passing of the Basin Plan, "adaptive management" became a defined term in legislation for the first time in Australia[1]. In the Basin Plan ***"adaptive management*** is taken to include the following steps:

(a) setting clear objectives;

(b) linking knowledge (including local knowledge), management, evaluation and feedback over a period of time;

(c) identifying and testing uncertainties;

(d) using management as a tool to learn about the relevant system and change its management;

(e) improving knowledge;

(f) having regard to the social, economic and technical aspects of management." (Section 1.07 Basin Plan 2012, Cth)

With respect to the differences between active and passive adaptive management, the Basin Plan definition is taken here as predominantly passive, due to the statement of 'management as a learning tool', the use of evaluation as feedback in a context of set objectives, identified uncertainties and acknowledged social, economic and technical 'aspects'. The definition makes no reference to modelling or experimentation; it states 'objectives' rather than hypotheses and 'evaluation' rather than science or research and emphasises local knowledge. While it can be argued that in doing several of the steps a) to f) experimentation and other active adaptive management processes can be used, it is not required by the definition. While there is some ambiguity, the definition points towards passive forms of adaptive management. In looking at how the term is applied to confirm this interpretation as passive, adaptive management as a defined term is used nine times throughout the Basin Plan. It is found in the:

[1] Other legislation, such as the NSW Water Management Act 2000 ambiguously refers to the "principles of adaptive management" (Section 5 Water Management Act 2000 No 92 Chapter 2 Water management planning Part 1 General Division 1 Water management principles), "objectives of adaptive management" (NSW Water Sharing Plans made under the Water Management Act 2000) or "an adaptive management framework" (Murray-Darling Basin Amendment Act 2002 – Schedule 1 http://www.austlii.edu.au/cgi-bin/sinodisp/au/legis/qld/bill_en/waolab2010362/waolab2010362.html?stem=0&synonyms=0&query=%22adaptive%20management%20%22)

- objectives for the Basin Plan[2;]
- purpose of Chapter 8 - Environmental Watering Plan[3];
- objectives of environmental management framework[4] with a note that "the application of adaptive management will enable various triggers to be responded to" (Section 8.11 Basin Plan 2012 Cth);
- principles of environmental watering, stating that "adaptive management should be applied in the planning, prioritisation and use of environmental water" (Section 8.40 Basin Plan 2012 Cth); and
- Chapter 13 – Programme for monitoring and evaluating the effectiveness of the Basin Plan with monitoring and evaluation to include adaptive management processes[5];
- principles of monitoring and evaluating the effectiveness of the Basin Plan as "monitoring and evaluation findings, including in respect of progress towards meeting targets and trends in the condition and availability of the Basin water resources, should enable decision-makers to use adaptive management" (Section 13.04 Basin Plan 2012 Cth); and
- key evaluation question "to what extent has the programme for monitoring and evaluating the effectiveness of the Basin Plan contributed to adaptive management and improving the available scientific knowledge of the Murray-Darling Basin?" (Section 13.06 Basin Plan Cth)

Across these references, adaptive management is intended to contribute to decision making, evaluation is to contribute to adaptive management and adaptive management will be evaluated. Adaptive management is about responding to triggers and applying a management process, confirming the passive interpretation. The relationship between science and adaptive management is not apparent, instead, monitoring and evaluation of the Basin Plan is seen to contribute both to adaptive management and to improving scientific knowledge.

The Murray-Darling Basin Authority describes the Basin Plan as an adaptive plan that "*is dynamic, and will be refined and updated with the knowledge gained from ongoing monitoring and evaluation framework*" (A Guide to the Murray-Darling Basin Plan, 2016). In a speech to the United Nations, the Chair of the MDBA described the Basin plan and its implementation as "*based on adaptive management*" because "*it's meant to be a flexible plan because in nature, things change. As we discover better ways to do things, we need to respond. Equally, we need to be ready to adjust to things like seasonal and climate changes*." In this speech the Chair also stated "*but, it's not just a 'science experiment'…the plan recognises the need to make judgements and decisions based on social and economic impacts*" and "*not everything happens in one day…communities need time to adjust to change and, for scientifically valid reasons, introducing a plan over time allows us to monitor, evaluate and adjust based on the*

---

[2] "to establish a sustainable and long term adaptive management framework for the Basin water resources" (Section 5.02 Basin Plan Cth)

[3] "enabling adaptive management to be applied to the planning, prioritisation and use of environmental water" (Section 8.02 Basin Plan 2012 Cth)

[4] to "enable adaptive management to be applied to the planning, prioritisation and use of environmental water" (Section 8.11 Basin Plan 2012 Cth)

[5] "processes for reviewing and evaluating the Basin Plan, conducting audits, and assessing the condition of the Murray-Darling Basin, contributing to adaptive management" (Section 13.01 Basin Plan 2012 Cth)

new knowledge and evidence that confronts us as we move into the future." (MDBA Chair Speech to UN, 2016). This managerial and passive interpretation is repeated in other public speeches and corporate documents, such as annual reports and provides insight into the organisation's internal view of adaptive management. The organisation and leadership has provided a strong rhetoric of adaptive management as necessary and intended.

Despite adaptive management as a defined term in legislation and strong public leadership level support for adaptive management the Basin Plan itself can be viewed as limiting adaptive management. For example, the legislated process for adjusting the Sustainable Diversion Limits is widely viewed by policy makers and implementers as a great example of adaptive management provisions. The Basin Plan acknowledges that the figures used to determine the Sustainable Diversion Limits were based on river management infrastructure expected to be in place and "*the level of scientific understanding of the Basin hydrology and ecology at that time*." As a result of these limitations, Chapter 7 of the Basin Plan outlines how the Sustainable Diversion Limits can be adjusted, with the process regarded as enabling adaptive management. However, for surface water the Basin Plan only permits adjustment on the basis of improved efficiency and supply of water. There are no provisions for new information on river systems, ecology or unpredicted negative impacts (social, economic or environmental). In addition, any experiential learning gained by river operators and Basin Plan implementers is not recognised as cause for change. The final limitation on adaptive management of the Sustainable Diversion Limits is that the net adjustment, Basin wide cannot be greater than 5%[6].

Lastly, and possibly most significantly, the coupling of adaptive management and evaluation in the Basin Plan can be viewed as limiting to adaptive management. Chapter 13 describes adaptive management as making a change as a result of an evaluation, effectively limiting adaptive management to occurring at a set time of the five and ten year reviews or through a political process as raised by a State Minister. Changing the Basin Plan requires an amendment to the legislation, or as one interviewee bluntly stated *"the bloody Act of course has to go through Parliament so it's not a trivial manner to change the Basin Plan."*

Reality of adaptive management

In interviews with policy makers and implementers of the Basin Plan, adaptive management was widely espoused as necessary and important. The reasons included demonstrating the success of the plan, improving implementation and to provide accountability. Statements of importance and support included *"It's really important, it's expensive to collect, but it would be just irresponsible not to do it"* and *"You obviously have to do it, and you have to do it as well as you possibly can. It's part of the accountability*".

However, in interviewing those involved in developing and implementing the Basin Plan for their definitions, the interviews confirmed its lack of meaning in reality. Adaptive management was described as *"more of a buzz word and an ideal rather than reality*" and *"It's a bit of an overused term, and I don't think we do it particularly well. I guess it's so overused I'm a bit over it to be honest*." One interviewee even stated *"the difficulty is though it's very, very, very hard to define. Nearly everybody you talk to, and you will probably find this, has got a different*

[6] See Basin Plan, Section 7.19 "Note: this section allows a supply contribution or an efficiency contribution of more than 5% of total surface water SDL to each be given full effect in an adjustment, provided that the net effect across the Basin is within the 5% limit"

idea of what it is." The definitions provided by interviewees commonly referred to learning by doing, checking progress to objectives, monitoring, review and management. One interviewee used the word hypothesis, but in a passive adaptive management way stating *"adaptive management is about coming up with probably a hypothesis, and then putting things in place to review that and adapt accordingly from the management outcomes."* In this, there was no discussion on putting experiments in place to test any hypotheses. Some interviewees acknowledged the difficulties of implementing large scale, replicated experiments in a varied landscape noting *"when you start talking about social-ecological systems, it becomes much more difficult, and you can't avoid your approach to adaptive management being more passive."* Evaluation language featured strongly, with some specifically noting efficiency and effectiveness with a typical definition provided by the interviewees being *"continuing to review, monitor, evaluate and rethink about how things should go and then adapt according to how you monitor things."*

Adaptive management, despite its definition and prescription in legislation remains an enigma in reality. In both intent and reality there is a dominance of passive forms of adaptive management, a loss of experimentation to gain scientific knowledge and a dominance of monitoring and evaluation language. The prevailing understanding of adaptive management by those implementing the Basin Plan is that adaptive management is part of or follows from an evaluation. Adaptive management is about achieving the policy's objectives, not changing or testing objectives; "*adaptive management works on the basis of seeking to achieve the outcomes that were originally set, and having a robust and transparent process in place to make adjustments along the way, if needed, to achieve those outcomes."*

Examples of adaptive management and implementation challenges

To further understand adaptive management, the interviewees were asked for good examples of adaptive management. Several interviewees responded that they could not think of a good example of adaptive management and the same example was mentioned by a number of interviewees, confirming a shortage of adaptive management in practice. For those that could identify actual examples of adaptive management, in contrast to the prevailing passive view of adaptive management, the examples also included active adaptive management. These examples mentioned experimentation at a site scale, for example to understand fish spawning. The presence of active examples is incongruent with the passive definitions. Others gave examples of passive adaptive management with adjustment of policy instrument or methods used to achieve a set objective, for example adjustments through efficiency projects or a scenario planning approach looking at what happened last year and current conditions to make decisions for the coming year. Passive adaptive management experiments that involved applying single treatments at a time to test a single hypothesis, such as weir pool height manipulations, were also noted.

In seeking explanation for the limited availability of good examples of adaptive management, the challenges to adaptive management were questioned. The challenges identified by interviewees largely confirmed the literature. The 'usual suspects' of cost, unclear responsibilities, lack of information, organisational culture, time, data complexity, landscape differences and scale of implementation were identified (Carter & Ross, 2013). For example, landscape differences and scale were seen to limit replication and transferability, with statements such as *"There are rivers where we are relaxing constraints, but we're not setting up an experiment. Could you compare, say, the Lachlan to the Gwydir? I just don't know if you*

could do it effectively" and *"The bigger the area, the more the people, the less active it can be and the more you move into a passive."*

In describing challenges, evaluation featured strongly in responses. For example, cost was identified on the basis that *"in a tight fiscal environment, the monitoring and evaluation programmes are the ones that tend to get dropped off, unfortunately."* Similarly, challenges associated with data complexity and quality of information was described as *"Having good data and information that's feeding into that. You've got to have a good monitoring programme on the ground. You've got to be able to have an effective way of evaluating the outcomes from that."* In these cases evaluation is seen as adaptive management, in line with the noted merging of the concepts.

In addition to the usual suspects, or easy scape goats, conflict and politics were raised as challenges to adaptive management. These challenges surfaced as public support for decisions and cross jurisdiction politics and accountabilities. For example, public support for decisions were seen as changing over time, creating a time limit or lifespan on legitimacy; *"even if you can accurately reflect community values and take them broadly into SDLs and you put it in, even if you could get that right, the following day you would be wrong, because community values are always changing."* and *"if you think some of the information maybe comes from the scientists then you've got to translate that information into a way that can win the public and bring the public along. We shouldn't underestimate the role of that, I don't think, in adaptive management in the long term being successful because ultimately you don't do anything unless you get the social licence to do it."*

When the Basin Plan was in development, it was argued that centralised governance of water resources would prevent local management and flexible responses (Crase et al., 2011). Interviews with those working to implement environmental watering have confirmed that this is now a reality. To overcome the time lags and other difficulties with gaining departmental and ministerial approval for environmental watering, the Annual Watering Plans are drafted as options papers, outlining a number of possible environmental scenarios ranging from dry to wet. This enables some flexibility to be retained by environmental water managers.

The interviewees also commented on how conflict over water use restricts adaptive management, particularly when a very specific detail, such as a volume for Sustainable Diversion Limits is negotiated and then legislated. For example, *"I think it's also at odds with our political process and also what the community expect when they want finite outcomes to be clearly defined and delivered."* In this context, adaptive management becomes limited and any change to policy is a point of conflict; *"anytime those policies, particularly the ones that are legislated, that they need to be changed there's always going to be conflict. In the sense that you may wish to change policy because of learnings that have occurred over the past 3 or 4 or 5 years or a decade, then yeah I reckon conflict is inevitable and it will make it a bit more difficult to get it through"* and *"There are too many people, too many vested interests who don't want to change things and they always take longer, hugely longer than you expect."*

Conflict and politics were also identified as impacting on cross jurisdiction politics and accountabilities. Recent government changes to responsibility for conducting external review and audit of Basin Plan implementation were described as reducing the imperative for the States to respond and change, *"There needs to be a reporting process that both the States and the Commonwealth are committed to. The National Water Commission was a product of a COAG decision, the Productivity Commission isn't. The State's don't necessarily have buy-in to the Productivity Commission. … there's nothing to oblige the States to actually make*

changes as a consequence of the Productivity Commission's reporting." Here, a weakening of political will to change, through a loss of accountability, is seen as curtailing adaptive management.

Discussion

The definition of adaptive management provided by the Basin Plan connects adaptive management to evaluation. The prevalence of adaptive management as passive, initially in the legislated definition and most certainly in interpretation and implementation of adaptive management, lends the concept to redefining as evaluation. This connection or redefining was echoed by the interviews. It could be regarded that linking to evaluation is a broadening of adaptive management, in line with that implied by the more recent references to adaptive management as structured decision making.

Adaptive management and evaluation in the Basin Plan

Adaptive management, perceived as evaluation in a performance improvement cycle, has some compatibility with the Basin Plan's definition. In both passive adaptive management and evaluation the results of the policy are monitored and results used to inform a management response. In the Basin Plan, adaptive management is the change that occurs following an evaluation of progress towards the policy objectives. This is problematic on two fronts, firstly the loss of knowledge discovery and secondly value judgements.

Adaptive management, as a decision that follows an evaluation, reinforces the passive approach of monitor and respond as the reaction to monitoring data. The policy or programme itself becomes the single hypothesis that is being tested, limiting adaptive management to resolving unpredictability. The adaptive management ideals of scientific discovery to address uncertainties of incomplete and imperfect knowledge remain incongruent with evaluation. The role of knowledge discovery in adaptive management and evaluation differs significantly. Evaluation is not regarded as generating new scientific knowledge on ecosystems or natural resources. Instead evaluation seeks to confirm or refute the results of policy or programme, with respect to the effectiveness, efficiency or appropriateness of its intended objectives. In evaluation, science may be used to confirm an assumption underlying a causal link from action to outcome or outcome to objective, whereas in adaptive management knowledge on system functioning is explicitly sought.

The role of value judgements differs significantly between adaptive management and evaluation. Adaptive management, taken as a process for improving policy and implementation through increasing knowledge of system behaviour, does not involve passing judgement or assessing merit. In stark contrast (and as noted by Scriven as a key distinction between research and evaluation), value judgements are central to evaluation that assesses the merit or worth of a policy or programme. Particularly within the Australian culture of natural resource evaluation, evaluation focuses performance assessment and is managerial. In the dominant approach of programme logic or programme theory, the only role for new discovery is in the testing and confirmation of assumptions with monitoring used to confirm and gauge expected policy results. One interviewee noted this difference in evaluation and the scientific discovery of adaptive management, stating "*Data can kind of provide some of the script for the thinking about those choices, but the choices are so inherently a value choice. To suggest that its adaptive management gives it a scientism which I think isn't there."* Adaptive management, interpreted as following from evaluation, dramatically changes the role of science in the policy making process.

Implications to the role of science in policy

Both adaptive management and evaluation seek to learn, with the ultimate purpose being to gain improved policy outcomes. Adaptive management takes a view of scientific hypothesis testing to discover new knowledge while evaluation focuses on the experience gained in policy implementation to identify recommendations. In evaluation, and it could be argued also in passive forms of adaptive management, monitoring seeks to confirm existing beliefs of system operation. Monitoring does not aim to explore alternatives. Instead, it is structured to a confirmation and validation bias. The misinterpretation or reinterpretation of adaptive management to evaluation, redirects adaptive management to a performance management concept, as a managerial tool. It also means a weakening of scientific inquiry, with incremental improvement of policy towards its objectives. The reinterpretation of adaptive management to evaluation may merely be reflective of the most recent adaptive management pseudonym of 'strategic decision making'. However, in effect, the result is a marginalisation of science from the policy making process.

Whether or not this marginalisation is intentional, accidental or through ignorance remains somewhat debateable. It could be argued that the logistical challenges associated with adaptive management have steered adaptive management towards evaluation. It could also be argued that the conflict and political challenges have made evaluation a much more attractive prospect. The risk associated with science providing proof of poor or incorrect decision making by government may to too great. A few interviewees specifically spoke about the role of science, stating *"we've got to get the science out of it"* and explaining the passive approach to adaptive management in the Basin Plan as "*the talk here about experiment and science, it wasn't unacceptable.*" To these interviewees politics and science were not compatible and as a result only certain forms or applications of adaptive management were palatable.

A fuller approach to adaptive management that systematically seeks to address more than just unpredictability through monitoring of outcomes is needed, or policy development will be limited to incrementalism and first loop learning. The scientific testing associated with adaptive management, that seeks to experiment to discover new knowledge and deliberative processes to engage multiple perspectives in decision making, pushes towards questioning objectives and values. In the absence of science, exploration and innovation of alternative solutions is limited.

Recommendations

A refinement to both adaptive management and evaluation practices is required to reinforce their respective contributions to policy planning and implementation. This solution recognises that scientific problem solving and performance improvement are both essential to governance of natural resources. For this adaptive management must remain as a distinct concept to evaluation, as both can generate significant knowledge and learning, with the findings ultimately used to improve policy. The common step of changing policy or making decisions based on findings is not an adequate reason to merge these concepts. That management may change, or in other words adapt, on the basis of findings, merely draws attention to the poor and ambiguous naming of the concept of adaptive management.

To overcome this, a clear statement of uncertainties at the outset of policy design is required, as originally intended in the Basin Plan's definition of adaptive management. The most appropriate form of investigation can then be used to address each uncertainty. Imperfect knowledge is most likely to require some form of experimentation, incomplete knowledge to require some form of social inquiry and learning, while an unpredictable response to policy may be best addressed through evaluation to test if expected causal impacts occurred, and indeed contributed to, achievement of objectives.

However, the politics remain most challenging to the practice of adaptive management and inclusion of science in policy. The political risks associated with 'being wrong', particularly in a high conflict context such as water, are significant. There remains a strong political need to remain accountable to highly negotiated and specific outcomes, such as a volume of Sustainable Diversion Limit, making change unlikely. In addition, the bipartisan support for the Basin Plan, passing through parliament with 95 for and only 5 opposing (along with the review and amendment processes required to adjust the legislation), makes any significant change unlikely. In this political environment the science involved in adaptive management represents risk. To overcome this significant challenge to adaptive management requires a change in governance structures and practices, particularly to maintain accountability alongside adaptive management.

Conclusion

Adaptive management and evaluation are two distinct concepts and practices. Despite the differences there has been a coupling or merging of the two, both in intent and reality. The dominance of evaluation and its paradigm of performance improvement designed to test the achievement of set objectives, acts to confirm policy choices and contributes to decision accretion and 'muddling through'. It fails to test alternative hypotheses and overlooks questioning the underlying values that contributed to initial decision making. Over time it leads to a narrowing of choices with incremental muddling through.

There are a number of logistical challenges that may have contributed to the merging of adaptive management and evaluation, but the underlying causes are conflict and politics. In the conflict context of the case study, adaptive management poses a political risk, with science having the potential to question the wisdom of past decisions. Prescription of adaptive management in legislation may provide adequate impetus for organisations to overcome the logistical challenges to adaptive management, but fails to address the underlying conflict and politics of the policy in focus. An evaluation focus on validating objectives remains preferable, confirming policy choices and providing accountability. As a result science is being marginalised from policy, with significant implications to discovery and innovation.

References

Alkin, M. C. (Ed.) (2013). Evaluation Roots: a Wider Perspective of Theorists' Views and Influences (2nd edition). Los Angeles: SAGE Publications.

Allan, C. (2009). Reviewing adaptive management through a wicked lens. Contested Country: Local and Regional Natural Resources Management in Australia, 215.

Allen, C. R., Fontaine, J. J., Pope, K. L., & Garmestani, A. S. (2011). Adaptive management for a turbulent future. Journal of Environmental Management 92(5): 1339–1345. http://doi.org/10.1016/j.jenvman.2010.11.019

Allen, C. R., & Gunderson, L. H. (2011). Pathology and failure in the design and implementation of adaptive management. Journal of Environmental Management 92(5): 1379–1384. http://doi.org/10.1016/j.jenvman.2010.10.063

Basin Plan 2012 Cth Accessed 2016, May 31. Retrieved from https://www.legislation.gov.au/Details/F2012L02240

Berkes, F. (2007). Understanding uncertainty and reducing vulnerability: lessons from resilience thinking. Natural Hazards 41(2): 283–295. http://doi.org/10.1007/s11069-006-9036-7

Brugnach, M., Dewulf, A., Pahl-Wostl, C., & Taillieu, T. (2008). Toward a Relational Concept of Uncertainty: about Knowing Too Little, Knowing Too Differently, and Accepting Not to Know. Ecology & Society 13(2): 30. [Online] URL: http://www.ecologyandsociety.org/vol13/iss2/art30/

Carter, R. W., & Ross, H. (2013). Tyrannies of effective monitoring and evaluation. Australasian Journal of Environmental Management 20(4): 263–269. http://doi.org/10.1080/14486563.2013.860872

Connell, D. (2007). Water Politics in the Murray-Darling Basin. Federation Press.

Crase, L. (2012). The Murray–Darling Basin Plan: an adaptive response to ongoing challenges. Economic Papers: A Journal of Applied Economics and Policy 31(3): 318–326.

Crase, L., O'Keefe, S., & Dollery, B. (2011). Some observations about the reactionary rhetoric circumscribing the guide to the Murray–Darling Basin Plan. Economic Papers: A Journal of Applied Economics and Policy 30(2): 195–207.

Dovers, S. (2009). Normalizing adaptation. Global Environmental Change 19(1): 4–6.

Eberhard, R., Robinson, C. J., Waterhouse, J., Parslow, J., Hart, B., Grayson, R., & Taylor, B. (2009). Adaptive management for water quality planning - from theory to practice. Marine and Freshwater Research 60(11): 1189–1195.

Funnell, S. C. (2000). Developing and using a programme theory matrix for programme evaluation and performance monitoring. New Directions for Evaluation 2000(87): 91–101.

Greene, J. C. (2013). Making the world a better place through evaluation. In Evaluation Roots: a Wider Perspective of Theorists' Views and Influences (pp. 208–217). Los Angeles: SAGE Publications.

Guba, E. G., & Lincoln, Y. S. (2001). Guidelines and checklist for constructivist (aka fourth generation) evaluation. Retrieved June 18, 2008.

Guide to the Murray-Darling Basin. (2016, April 10). Retrieved from http://www.mdba.gov.au/basin-plan/plan-murray%E2%80%93darling-basin

Gunderson, L., & Light, S. S. (2006). Adaptive management and adaptive governance in the everglades ecosystem. Policy Sciences 39(4): 323–334. http://doi.org/10.1007/s11077-006-9027-2

Huitema, D., Mostert, E., Egas, W., Moellenkamp, S., Pahl-Wostl, C., & Yalcin, R. 2009. Adaptive water governance: assessing the institutional prescriptions of adaptive (co-)management from a governance perspective and defining a research agenda. Ecology and Society 14(1): 26.

Lee, K. N. (1993). Compass and gyroscope. Island Press, Washington, DC.

Loftin, M. K. (2014). Truths and governance for adaptive management. Ecology and Society 19(2): 21.

MDBA Chair Speech to UN. (2016, April 10). Retrieved from http://www.mdba.gov.au/media/mr/mdba-chair-speech-un

Monitoring, Evaluation, Reporting and Improvement Strategy. (2014, July 1). Australian Government. Retrieved from http://www.nrm.gov.au/publications/meri-strategy

Pagan, P., & Crase, L. (2005). Property right effects on the adaptive management of Australian water. Australian Journal of Environmental Management 12(2): 77-88.

Pahl-Wostl, C. (2009). A conceptual framework for analysing adaptive capacity and multi-level learning processes in resource governance regimes. Global Environmental Change 19(3): 354–365.

Pahl-Wostl, C., Conca, K., Kramer, A., Maestu, J., & Schmidt, F. (2013). Missing links in global water governance: a processes-oriented analysis. Ecology and Society 18(2): 33.

Patton, M. Q. (2013). The Roots of utilisation-focused evaluation. In Evaluation Roots: a Wider Perspective of Theorists' Views and Influences (2nd edition), pp. 293–303. Los Angeles: SAGE Publications.

Rogers, P. J., & Davidson, E. J. (2013). Australian and New Zealand evaluation theorists. In Evaluation Roots: a Wider Perspective of Theorists' Views and Influences (2nd edition), pp. 371–385. Los Angeles: SAGE Publications.

Scarlett, L. (2013). Collaborative adaptive management: challenges and opportunities. Ecology and Society 18(3): 26.

Schreiber, E.S.G, Bearlin, A.R., Nicol, S.J., & Tod, C.R. (2004). Adaptive management: a synthesis of current understanding and effective application. Ecological Management and Restoration 5 (3): 177-182.

Scriven, M. (2013). Conceptual revolutions in evaluation: past, present and future. In Evaluation Roots: a Wider Perspective of Theorists' Views and Influences (2nd edition), pp. 167–179. Los Angeles: SAGE Publications.

Walters, C. (1997). Challenges in adaptive management of riparian and coastal ecosystems. Conservation Ecology 1(2): 1.

Walters, C. J., & Holling, C. S. (1990). Large-scale management experiments and learning by doing. Ecology 71(6): 2060–2068.

Developing a monitoring, evaluation and learning system for the Food Systems Innovation initiative: an experience of 'three steps forward, two-steps back'

Stone-Jovicich, S.[1], Kelly, J.[2] and Hall, A.[2]

[1] *Land and Water, Commonwealth Scientific and Industrial Research Organisation (CSIRO)*
[2] *Agriculture, Commonwealth Scientific and Industrial Research Organisation (CSIRO)*

Abstract: Most monitoring and evaluation approaches and tools available to practitioners and organisations have been designed either to monitor tangible outputs and evaluate performance, or to capture unquantifiable processes and dynamics of change. While numerous are excellent, they are limited in their capacity to meet the multiple accountability, learning, and adaptive management needs and rhythms of complex initiatives aimed at fostering system-level innovation. In this paper we share our experience developing and trialling, over a three and a half year period, a monitoring, evaluation and learning (MEL) system for the Food Systems Innovation (FSI) initiative in Australia. A partnership between three Australian government departments, FSI was highly ambitious and complex. Its main goal was to improve the impact of ODA-supported agriculture and food security programmemes in the Indo-Pacific region by fostering innovation via partnership-building; the exchange of knowledge and experiences; cross-sectoral, systems thinking; hands-on-experimentation with novel approaches; and an embedded MEL system. The challenge the FSI team encountered was not difficulty in finding 'best practices' in MEL. Rather, it was identifying a suite of approaches and tools, whilst also attending to other dimensions of the MEL system, that were a 'best fit' with the complex nature of the agriculture and food security issues FSI was trying to tackle, and with the complexity of FSI itself. In this paper, we discuss the challenges, failures, and promising leads encountered while designing and implementing a flexible and multi-purpose MEL system that tried to balance, on the one hand, a rigorous theoretical foundation drawing on complex systems and innovation thinking with, on the other hand, the production of non-cumbersome and easy to implement and communicate MEL tools.

Keywords: Monitoring, evaluation, learning, MEL system, complexity, innovation, fit-for-purpose

Introduction

Over the past decade, the aid and development sectors have come under increasing pressure to acknowledge and address the failures of dominant models of aid investments and interventions predicated on linear understandings of change, technological solutions, and accountability applied in a world that is complex, highly dynamic and unpredictable (Boulton et al., 2015; Burns & Worsley, 2015; Ramalingam, 2013). This largely intellectual debate has begun to influence the way we think about and design aid practices, most notable in recent efforts in monitoring and evaluation (M&E) to develop approaches and tools that are 'complex-aware' (Douthwaite et al. 2003; Guijt, 2011, 2010a, 2007; Patton, 2011; Ramalingam et al., 2014; USAID, 2013) and aimed at catalysing learning and innovation (Hall et al., 2003; Kusters et al., 2015; USAID, 2016).

While this shift in M&E to move beyond simply measuring efficiency and effectiveness is significant, there are numerous challenges. The first revolves around the identification of 'best practice' monitoring and evaluation frameworks and tools that are well-matched for helping donors, practitioners, and partners navigate complexity and for catalysing collective learning and building system-level innovation. The fact that there exist so many M&E approaches and tools in many ways makes this task harder. Discerning among the vast 'toolkits' and 'guides' which ones are the right fit, or have the potential to be adapted to be suitable, is a slow process. The majority of M&E frameworks and methods are designed to either monitor and assess tangible outputs and evaluate performance for primarily planning and accountability purposes (mainly through quantitative methods and linear, or stepwise, perspectives of change) or to capture less tangible (i.e. less quantifiable) processes and dynamics of change (primarily qualitative or narrative-based approaches) (for succinct overviews and constructive critics of M&E see Guijt, 2008, 2010a; Hughes et al., 2013). Moreover, most M&E frameworks and tools are designed to be implemented at the level of discrete units or scales, most commonly at a project or organisational level, or for a specific sector; less so across different components, dimensions, and scales of a system (Lynam, 2012; Odame et al., 2012). Also, whilst there is a rapidly expanding body of tools and processes focused on learning (e.g. social learning approaches, appreciative inquiry) many of these are being developed and practised in other arenas (e.g. facilitation, knowledge management and brokering, adaptive co-management) although, in recent years, there has been a push to more explicitly link M&E and learning (Guijt, 2010a, 2007; Oswald & Taylor, 2010; Woodhill, 2007). However, many learning-focused methods remain time- and data-intensive, expensive, highly context-dependent, and their value-add a challenge to communicate and embed within M&E systems.

While progress is being made in identifying a set of best practice 'complexity and innovation aware' M&E and Learning (MEL) tools, this is not sufficient. For MEL approaches and tools to be effective in navigating complexity and enabling innovation they need to be, by design, flexible and adaptable (Ramalingam, 2011). As such there are no fixed 'best practices' but rather a set of 'best-known practices' (or promising practices) that need to be continuously assessed, probed, and revised to ensure relevance and effectiveness. Thus, the second challenge, which remains underexplored, is the systematic collation of lessons learned that emerge from the application of these MEL tools in complex systems and innovation systems - i.e., which approaches and methods are effective (or not) in catalysing greater capacity to navigate unknowns and uncertainties, to foster collective and cross-scale learning and to lead to fundamental changes and system-level innovation? Capturing patterns of how and why certain MEL approaches work, while others do not, across different contexts and points in time in policy/programme/project cycles is fundamental to enhancing understanding and practices of MEL as vehicles for building and measuring system capacity to learn and innovate.

Last but not least, for MEL to be effective in enhancing system capacity to innovate it needs to extend beyond simply discrete approaches and tools and, rather, embrace a wider assemblage of interlinked components or elements, i.e. the whole-of-MEL-system. While the term 'M&E systems' is widely used, it is not uncommon that the predominant focus remains on tools - for planning M&E, identifying indicators, collecting information, and so on. Over the past decade, more attention is being paid to other elements seen as critical for M&E to function effectively, including the need for sufficient human capacity and diversity of skillsets; partnerships for planning, coordinating and managing M&E; organisational structures for embedding and sustaining M&E; and dimensions associated with communication, advocacy

and culture (see, for example, UNAIDS, 2009). Scholars and practitioners engaging explicitly with innovation systems thinking have also highlighted the need to move beyond approaches and tools that focus on individual organisations, or the relations between two organisations, to 'systemic instruments' that operate at and support system-level functions and change (see van Mierlo et al., 2010b). In practice, working across all of these dimensions of an MEL system is incredibly challenging.

We bring the points raised above to the fore by sharing, in the remainder of this paper, our experience in designing and implementing a MEL system for the Food Systems Innovation (FSI) initiative, a project which aimed to improve the impact of Australian-supported aid investments in agriculture and food systems. Our greatest challenge was trying to transform M&E from being a 'tick-the-box' exercise to a set of M&E *and* Learning tools and processes that could effectively support our team and partners not only to plan short- to medium-term activities and demonstrate accountability, but also to engage in co-learning, be responsive to emergent issues and opportunities, and ultimately help bring about significant and lasting changes. We largely failed but learned a lot along the way.

Background

FSI's story

The Food Systems Innovation (FSI) initiative was a three and a half year research for development (R4D) project with the aim of enhancing the impact of Australian-supported international aid investments in agriculture and food (FSI, 2016a). It was founded on a partnership between three Australian government agencies/organisations: AusAID/Department of Foreign Affairs and Trade (DFAT), the Australian Centre for International Agricultural Research (ACIAR), and the Commonwealth Scientific and Industrial Research Organisation (CSIRO). DFAT funded FSI; ACIAR was a partnering organisation; and CSIRO was the project lead responsible for co-designing (with DFAT and ACIAR), managing and implementing the initiative.

FSI was launched in 2012 as a 1-year 'inception project' and was initially called FSIFS (Food Security through Food Systems Innovation). FSIFS emerged from an existing partnership between CSIRO and AusAID (which, in 2013, was integrated into DFAT). It had a focus on improving the design and implementation of food security R4D interventions. This Inception Phase ended in mid-2013 and Phase 2 was launched, with FSIFS being renamed FSI. In this second phase of the initiative, greater emphasis was placed on innovation and, accordingly, FSI aimed to contribute to an enabling environment to support continuous innovation within the Australian agriculture and food aid programme.

FSI in its second year had three components: 'analysis and field application' (primarily R4D activities); 'knowledge management and communications'; and 'capacity building'. A mid-term external review and a change in leadership at the end of the second year resulted in a reorganisation of FSI into three 'focal areas'. The first was 'Markets and Partnerships' which centred on market-driven approaches to agricultural development that are inclusive and leverage private sector investment for maximum impact. The second, 'Agricultural Linkages', focused on strengthening the linkages between agriculture and other development priorities such as nutrition-sensitive agriculture. And the third focal area was 'Managing for Impact' with an aim of supporting agriculture projects to have sound Theories of Change, good management and robust monitoring and evaluation. These three focal areas broadly guided

the direction and objectives of the initiative's key activities and products which were: learning events (which ranged from training courses to reflection workshops), knowledge products (conventional research-based reports to learning-oriented 'Practice Notes' and blogs), expertise and practice networking activities (among Australian and international experts and development practitioners), in-country engagement (primarily in Southeast Asia but also in Africa), and outreach and external visibility activities (from 'Brown-Bag' seminars and conference presentations to the FSI website) (see FSI, 2016a for more details).

Across the three focal areas, activities and products were designed and implemented with the objective of creating conditions for supporting innovation. Thus a key focus was on maximising collaborations and alliances; enabling the exchange of knowledge and experiences, analysis, reflection and learning; and fostering cross-sectoral, systems thinking. In addition, within the FSI project team emphasis was placed on working with uncertainty and emergent opportunities, and brokering of ideas, knowledge, networks and differences in perspectives, including regarding how to best achieve FSI's goals. Engaging in 'boundary work' was seen as critical, i.e. building and managing the interfaces among FSI partners to effectively mobilise knowledge and networks into actions. Accordingly, FSI team members took on multiple roles, including those of boundary workers/brokers/spanners, managers, communicators and researchers. The FSI project team also endeavoured to embed (where possible) its activities and products within the broader Australian and international agriculture development environment.

With CSIRO responsible for the operationalisation and implementation of the initiative, the FSI project team (leader and implementation staff) was comprised predominantly of CSIRO staff, the majority of whom were scientists. Key partners - DFAT and ACIAR - contributed to FSI through their role as members of the FSI Steering Committee and Management Committee, and through engagement with and active involvement in focal area activities and contributions to MEL. Over the course of the three and a half years, the number and composition of people involved changed (particularly between the Inception Phase and Phase 2). Nonetheless, there remained a core of approximately 10 individuals, from across the different organisations, engaged in the last two and a half years of FSI. In line with FSI's aim to build an enabling environment for innovation, there was a progressive effort to bring in external experts (i.e. not affiliated with CSIRO, DFAT, or ACIAR) to fill in gaps in expertise, knowledge and networks seen as critical for fostering innovation within the Australian agriculture and food aid programme.

The Monitoring, Evaluation and Learning (MEL) team and system

M&E was not explicitly discussed during the Inception Phase. Rather, in the first year of the initiative there was a strong focus on inter-organisational learning and capacity-building. Three team members, all of whom were CSIRO staff, were tasked to develop and implement a series of 'learning activities' (e.g. workshops, training courses, and presentations) with the aim of fostering the exchange of knowledge and skillsets among DFAT, ACIAR and CSIRO.

M&E was formally introduced in the second year of the initiative. DFAT and ACIAR feedback on the Inception Phase and a change in leadership led to the decision to bring in a formal MEL system. The view was that it would serve as a mechanism to more explicitly track progress as well as support the learning agenda that had been initiated during the Inception Phase. This resulted in the creation of a MEL team, headed by one of the CSIRO staff previously involved

in the ‘learning activities’ in the Inception Phase. While formally incorporated into FSI and loosely linked to the ‘Managing for Impact’ focal area, the MEL team and activities remained throughout FSI slightly in the margins as the majority of the FSI team and budget was allotted towards the above-mentioned three focal areas. The MEL ‘team’ consisted of one individual exclusively allocated to the M&E and learning activities, with other FSI team members and a DFAT staff member contributing to MEL strategic discussions and/or information gathering and reflection processes. While these individuals did not have any of their time formally allocated to MEL, most actively and frequently engaged with MEL. Nonetheless, their official roles necessitated them focusing on delivering to the initiative’s focal areas (FSI team members) or other priorities (in the case of the DFAT collaborator). This meant that the MEL ‘team’ was essentially comprised of one individual, a scientist with limited experience in the *practice* of M&E and learning.

The primary task of the MEL team was to design and implement a MEL system that could effectively support FSI progress its purpose and goals. This system, as envisioned by the MEL team, comprised a conceptual framework that delineated the assumed pathways to system-level innovation and, thus, key areas of focus for MEL (e.g. a programme logic or Theory of Change; indicators or domains of change); approaches and tools, along with guidelines and plans, for information gathering, processing, and reporting, and for collective sense-making, learning, and feedback into FSI management; key people (i.e., team members with roles in planning, coordination, and management of the MEL system; ‘producers’ and ‘users’ of information and participants/catalysers of learning processes); aligned structures (i.e., project and organisational structures in which to embed/link MEL); and supportive social dynamics and culture (i.e., relationship building processes and conducive environment for M&E, reflection, learning, experimentation and change). Another critical element was a budget for MEL which in the case of FSI was within the range of what the partners typically invest in M&E (i.e. less than 10% of the overall project budget).

What this MEL system needed to serve and how it was going achieve that - in terms of specificities regarding the various bits and parts comprising the system - for such a complex and ambitious initiative was not immediately apparent. It took almost two years, and a number of trials and errors, for the MEL team to come to realise that the MEL system for FSI had to be multi-layered and flexible enough to: support the initiative’s agree-upon short, medium and long term objectives, whilst also help nurture innovative ways of thinking and practice; meet the different needs and emergent information, learning and governance requirements and rhythms of partner organisations, and those of the initiative’s management and implementing teams; and be responsive to the continuous changes in the broader Australian ODA context (and hence to changes in FSI’s enabling environment and impact pathways). Designing and implementing a MEL system that could effectively support these different objectives, uses and decision-making expectations and needs proved to be a significant challenge. In the remainder of the paper, we share what we did, what worked and did not work, and why. We focus primarily on our MEL approaches and tools, as we found that the majority of our efforts and time was consumed by those elements of the MEL system.

Our journey from 'best practice' to 'fit-for-purpose' MEL system

A start with 'best-practices'

In the first two years (Inception Phase and Phase 2; July 1012 – May 2014), the FSI team drew on 'best practice' approaches and tools in the field. Given that FSI was a partnership initiative between AusAID/DFAT, ACIAR and CSIRO, and a key initial objective was enhancing inter-organisational collaboration and learning, the team chose to focus on social learning approaches, drawing on single-, double-, and triple-loop learning theory (Argyris & Schön, 1996; Fabricius & Cundill, 2014; Flood et al., 2002; Guijt, 2010a; Romm, 1996). The rationale was that providing mechanisms for these three diverse organisations to share experiences, perspectives, and interests would lay the foundation for a stronger partnership, support the development of a shared vision for FSI, assist in institutionalising learning and adaptive management, and elicit cross-organisation relevant avenues for 'thinking and doing out of the box'.

While the formation of the FSI partnership could be seen as a form of social learning in itself (i.e. the recognition of the need - if greater international development impacts are to be achieved - for a closer synergetic relationship between the three main Australian government agencies delivering agriculture and food aid investments, and the subsequent formation of an inter-organisational partnership) (see Guijt, 2007), embedding explicit iterative processes of critical reflection and learning in the partnership proved to be too challenging to implement at the time. There are many reasons for the difficulties we encountered, some of which are discussed later in the paper. Among these was our struggle to design tools and processes that were non-cumbersome, relatively easy to implement and, perhaps most significantly, that were perceived as cogent and pragmatic in the eyes of our partners.

This led, in year 2 of FSI, to the MEL team deciding to establish closer linkages with the M&E frameworks and tools that were either being used by, or were familiar to, our DFAT and ACIAR partners. We invited M&E specialists from our partnering organisations to contribute to our efforts. This resulted in the development of a programme logic-oriented, indicator-based M&E framework (Figure 1). We inevitably encountered problems with this form of conventional 'best practice' M&E. As highlighted by others (e.g. Woodhill, 2007) these types of approaches were too narrow and rigid to adequately capture the dynamic nature and complexity of initiatives such as FSI. And mechanisms to support learning and capacity to innovate were virtually absent.

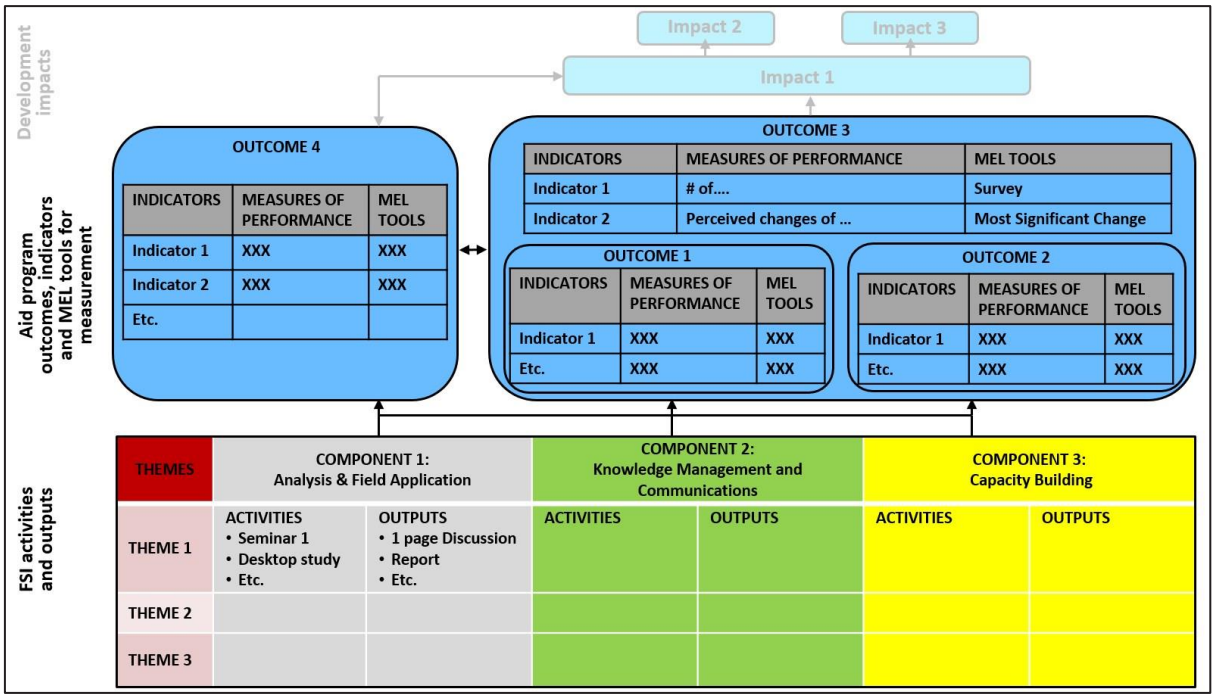

Figure 1. FSI's second attempt at an MEL framework (year 2): a programme logic, indicator-based framework

Experimentation with a 'home-grown' emergent approach

Towards the end of year 2 (mid-2014), we decided that we needed to take stock of where we were at, particularly why we were struggling with our MEL system. Around the same time, a small group of FSI team members met to jointly reflect on the achievements and challenges encountered over the previous year. Using a loosely-facilitated discussion format, and aided by a visual flow-chart methodology, we spent two days trying to make sense of what felt like a 'messy, incoherent and busy' year. The experience of sitting together and jointly reflecting led to the team pinpointing some of the key activities, outputs, and processes that had led to positive outcomes for FSI. Most importantly, it enabled them to 'connect the dots' between what appeared to have been disparate activities/outputs/processes and to identify why some had been successful and others not. The reflections and lessons learned in those two days culminated in the development of what we called a 'Learning Trajectory', a diagram that visually captured accomplishments (and bottle-necks) and the underpinning processes that led to key outcomes (Figure 2). The Learning Trajectories helped FSI team members succinctly capture, understand, and communicate achievements (and failures) and were subsequently found to be helpful in informing the work plan for the third year of FSI.

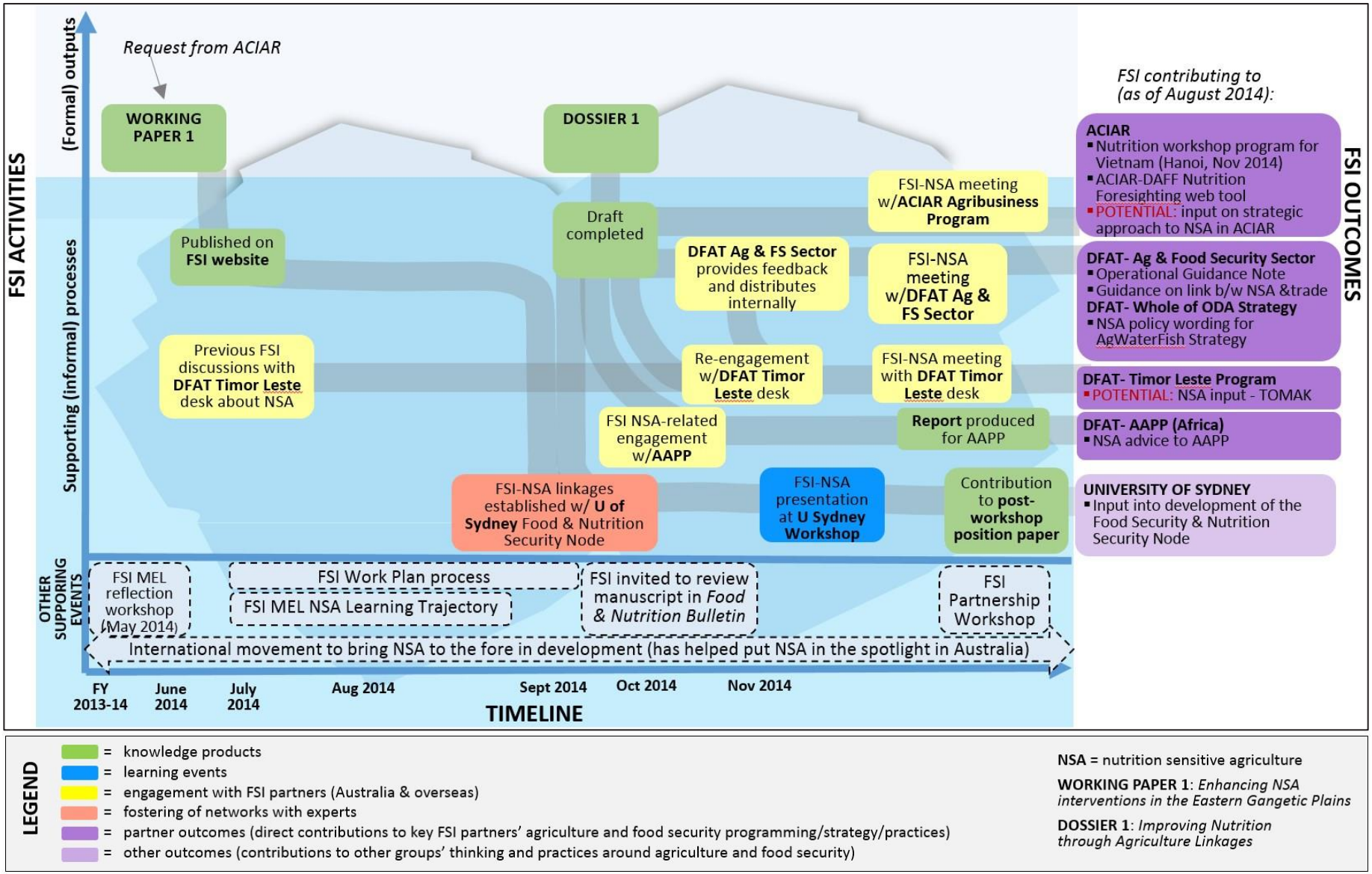

Figure 2. FSI's Learning Trajectory (end of year 2 of FSI): an example for the Agricultural Linkages focal area

Moving to a 'fit-for-purpose' approach

While developing the Learning Trajectories proved to be a valuable critical reflection and learning experience for the FSI team, it was a time-consuming exercise. It also was not easy to replicate as it was logistically too difficult and too expensive to bring partners together to engage in regular face-to-face interactions. Nonetheless, our experience trialling three very different MEL approaches - from social learning, to indicators, and then the Learning Trajectories - shed light on where we had gone wrong and what was needed, in terms of MEL approaches and tools. It highlighted for us that the MEL system needed M&E tool(s) capable of capturing the non-quantifiable, process-level aspects of FSI that had come to light in the Learning Trajectory exercise; it also needed to be able to incorporate, in a concise way, differences in perspectives and opinions regarding FSI's achievements, including those of our key partners as well as of others (e.g. people who participated in FSI activities). It also had to be easy to update on a regular basis, as well as be succinctly communicated to FSI's team and governance committees. Finally, and arguably most importantly, we desired a tool that was better able to monitor and evaluate *innovation*, both its more tangible elements (i.e. implemented activities, products, networks, etc.) and embedded dynamic processes and emergent outcomes so that we articulate consistent, yet varied and nuanced, stories about how system-level innovation can be effectively enabled and supported in the long-term. In line with this last criteria, the tool thus also needed to have the potential to be used as a platform for reflection and learning.

At this juncture, we decided that we needed to seek external expertise and approached, with the above criteria in mind, an MEL practitioner with experience working in complex projects/programmes (Irene Guijt). That engagement led the FSI team to consider a M&E approach called Rubrics, a qualitative descriptive assessment tool which involves articulating and clarifying 'the things that matter' in a project or initiative (King et al., 2013; Oakden, 2013a, 2013b; Oakden & Weenink, 2015). Some members of the FSI team had been exploring this approach with partners overseas, which provided some, albeit limited, internal experience and validation of the potential of this M&E tool for FSI. With the guidance of external M&E experts (Irene Guijt and Judy Oakden), we experimented with the Rubrics for FSI. We started with the initiative's programme logic and for each set of core activities, we revisited the outcomes and identified the key qualities or changes that would tell us that FSI was achieving its purposes and progressing towards its goals (Figure 3).

While the Rubrics approach met the criteria we had identified as crucial for the FSI MEL system, we found that we had to significantly adapt the approach. In the end it took us several attempts and modifications to design a Rubrics that was 'fit-for-purpose' for FSI (for more details and reflections on our experience see Stone-Jovicich, 2015). At around the same time (3rd year of FSI), we found that a regular e-mail update called 'FSI This Past Month' - prepared by the FSI Management Committee for members of the Steering Committee and other key individuals in DFAT and ACIAR, and comprising a detailed listing of the activities and outputs that had been completed and were underway - was a very effective monitoring tool. This also overlapped with the development of a knowledge product called 'FSI Practice Notes', compilations of succinct written reflections drawing on experience and aimed at facilitating shared learning and innovation to improve practice amongst research and development practitioners (see FSI, 2016b). These three approaches ended up forming the cornerstone of the approaches and tools of FSI's MEL system in the final year of the initiative (Table 1).

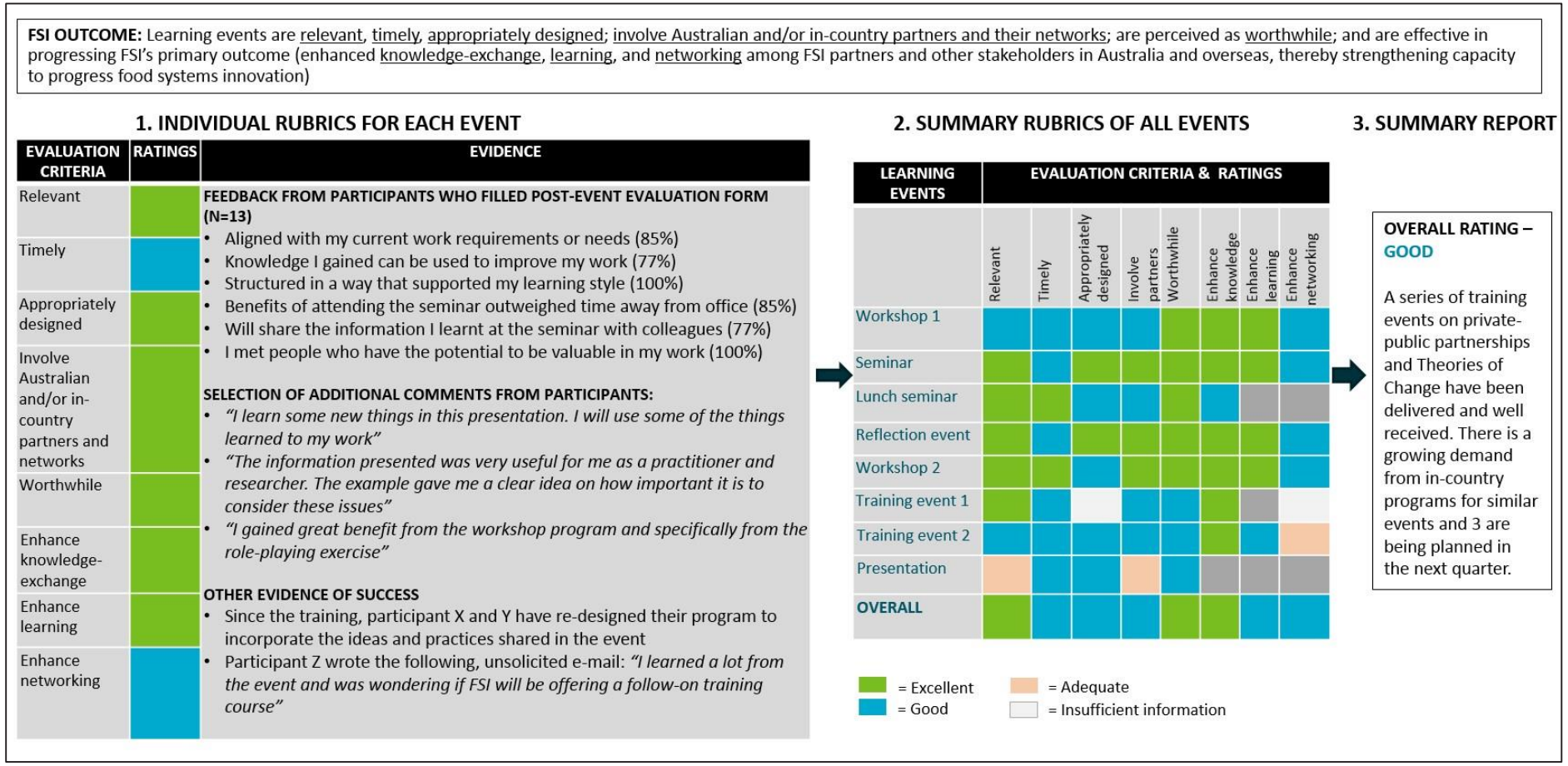

Figure 3. The FSI Rubrics approach (year 3 of FSI): an example for Learning Events (fabricated data)

Table 1. The final suite of MEL tools for FSI (end of year 3 of FSI)

MEL APPROACH/ TOOL	M, E, or L*?	M, E, or L OF WHAT?	FREQUENCY & FORMAT	PRIMARY PURPOSE
FSI This Past Month	Monitoring	FSI short-term to medium-term activities and outputs, planned and emergent	Approx. every month; Succinct e-mail message	Evidence-base for FSI Steering Committee & funders; Accountability; Sharing where everyone is at; Day-to-day management
FSI Rubrics Reports	Evaluation Learning (potentially)	Evaluation of completed FSI activities and outputs; Reflections of processes, and progress towards short and medium term outcomes ; Deviations from plan; wins & challenges; Lessons learned	Attended to be approx. every 3-4 months: FSI Steering Committee meetings; Succinct report comprised of rubrics tables & narratives	Evidence-base for FSI Steering Committee & funders; Accountability; Critical decisions and directions (management)
FSI Practice Notes	Learning	Practice-based experiences and lessons learned	As 'critical mass' of practices and associated lessons learned gathered; Succinct, reader friendly 'notes'	Synthesis and sharing of key lessons learned and insights from practice

** Our definition of learning drew on single-, double-, and triple-loop learning theory (Argyris & Schön, 1996; Fabricius & Cundill, 2014; Flood et al., 2002; Guijt, 2010a; Romm, 1996). We viewed learning as comprising both a process (i.e. involves and emerges from an individual and/or from a group of people interacting together over a sustained period of time and engaged in on-going deliberation and the sharing of knowledge in a trusting environment) and an outcome [i.e. changes in the way we think (cognition), see (framing), value (affect/emotions), and act] (Cundill & Rodela 2012; Cundill et al. 2014; Wals & Rodela, 2014). These learnings can serve multiple purposes, ranging from accountability to supporting institutional transformation (see Guijt, 2007).*

FSI officially ended before we could adequately test these MEL approaches and tools (for example, we never had an opportunity to experiment with Rubric's potential as a catalyst for collective learning and for enhancing system capacity to innovate). Nonetheless, this mix of MEL tools and associated reporting mechanisms and formats appeared promising in terms of

their capacity to support FSI's multiple needs, aims, and demands. As to whether they would have been effective for catalysing learning that transcended individuals and organisations and enhanced system-level innovation capacity, it is hard to say.

Critical reflections and lessons learned from our experience

1. Don't lose sight of the forest for the trees

One of the biggest lessons for the MEL team was that it is very easy to focus on approaches and tools and lose sight of the broader MEL system. While we intuitively knew that an effective MEL system for supporting innovation needed to be greater than the sum of its information-gathering and sense-making tools, we struggled to give sufficient attention to other critical elements. These included putting more time/effort/resources into: a collectively articulated whole-of-FSI Theory of Change (or other similar approaches); adequate human capacity for MEL; clearly identified organisational alignments for embedding or linking MEL (e.g. M&E structures and teams within our partner organisations); and the strengthening of a culture of and practices for inter- and intra-organisational (and institutionalised) critical reflection, risk-taking, learning and system-level change and innovation. Attempts were made but an insufficiently experienced MEL team leader, a small MEL budget, and a FSI team that was stretched-to-capacity were some of the major constraints encountered.

2. Start small and modest

Starting with social learning as the underpinning conceptual framework for developing M&E and Learning tools and processes for FSI was too 'big', i.e. too ambitious and unrealistic. A certain level of trust is needed for people to engage in critical reflection and sharing of opinions and perspectives (Pahl-Wostl & Hare, 2004; Pretty, 1995). When FSI started, the partnership among the three organisations was in its formative stage. It was thus unrealistic to assume that the conditions were in place to encourage the team and partners to partake in co-learning processes. Moreover, while the MEL team envisioned social learning workshops as mechanisms to inductively generate a set of indicators that could then be used for M&E purposes and for supporting learning, they were time consuming exercises which were difficult to fit into the time-constrained environments in which everyone operated. In hindsight, we should have started 'small' - e.g. informal, 'rapid' ice-breaking conversations; 'back-of-the envelope' M&E frameworks used as boundary object to select, with partners, tools that were perceived by team members and partners as not too time-consuming and cumbersome to engage with.

3. Don't overthink

The MEL team spent too much time on developing and trying to communicate the conceptual thinking behind the choice of M&E and Learning tools, and too little time on implementation. It was important to offer a solid rationale for why the team was selecting the types of tools it was - i.e., explaining the fit with complexity and innovation thinking - and thus build greater awareness around why conventional M&E approaches were not sufficient. However, in hindsight, we may have achieved this (and more effectively) by putting the tools into practice. Then, through processes of iterative feedback and refinement, we most likely could have demonstrated the value-add of these tools in terms of enhancing understanding of change in

complex projects and innovation pathways, and as mechanisms for capturing and communicating these storylines and impact.

4. Don't fall into the data collection trap

In similar vein to the first lesson learned discussed above, it is easy to get too focused on the need to collect data and, in the process, miss the bigger purpose of MEL. The Learning Trajectory exercise, carried out mid-project at a point where FSI felt a bit chaotic, brought this to the fore for us. At that stage we had been very focused on defining indicators and data collection methods. The Learning Trajectory exercise, which brought together a range of team members, forced the MEL team to rethink what was really critical for supporting FSI to progress. It was not having an exhaustive list of indicators, along with a suite of robust data collection methods. These could not adequately keep up with the pace of FSI management needs and changes in partner priorities, nor meet FSI's broader aim of fostering innovation capacity.

5. The MEL approaches and tools need to be 'fit-for-purpose'

Whilst FSI was complex and had a focus on building an enabling context for fostering innovation, the MEL tools had to meet multiple accountability, learning, and adaptive management needs and rhythms. That meant that we could not just focus on tools that were suited for fostering learning and enhancing capacity to innovate; we also had to consider tools that were best for measuring progress on delivery of agreed-upon outputs and for activity planning purposes. In other words, we needed to put together a mix of MEL approaches and tools that were 'fit-for-purpose' (Ramalingam, 2011; Ramalingam et al., 2014). Most of the struggles we encountered were because we either focused at one end of the spectrum (learning-focused approaches) or the other (conventional, accountability oriented methods), or tried to build a tool that could do it all.

We also had to be willing to let go of tools, even if they had worked well. This was our experience with the Learning Trajectory exercise. It ended up being one of the most useful MEL tools *at a particular point and time in FSI* as it helped capture a snapshot of how FSI was progressing overall at the time and helped guide the subsequent phase of FSI, both its overarching direction and the planning of activities. FSI subsequently gained greater focus (for a range of reasons, beyond the contribution of the Learning Trajectory exercise) and the Learning Trajectory was no longer the most suitable tool.

It took a while to develop a fit-for-purpose MEL system. However, by the end of the project, we seemed to have identified a reasonable mix of M&E and learning methods that were a right fit with the multiple objectives and needs of the initiative. The 'FSI This Past Month' updates proved to be very effective in supporting the FSI team report on progress of activities and agreed-upon deliverables. The 'FSI Rubrics' were promising as a mechanism for assessing and demonstrating progress towards outcomes as well as for making visible and communicating how the sum of, and interactions and dynamics among, the individual activities and outputs (both defined in the project workplan and emergent) was key to some of FSI's achievements and impacts. Also, in the 'FSI Practice Notes', we seemed to have finally found a learning tool that was seen as useful and not overly demanding.

6. Engaging those who will make use of the MEL information and lessons learned, along with MEL experts, is critical

After having failed several times, we realised that involving the people/groups who will make use of the information and lessons learned from the MEL approaches tools is critical to making the MEL system useful, appropriate, and effective (see Guijt, 2010b; Ramalingam, 2011). We should have engaged more closely, from the very beginning, with FSI team members, the Steering Committee and staff from partner organisations. We also would have benefited from linking earlier on with MEL experts who had experience working in complex projects. This would have led to a more balanced team, one made of team members who brought with them the diversity of knowledge, skills and experience needed. In this way, we would have most likely avoided what we did - i.e., 'hop' from one tool to another in search of the 'best practice' approach and tools - and arrived much earlier at a set of M&E and learning approaches and tools that were a 'best fit'. Having said this, it is not guaranteed to work. We attempted various co-designing approaches at various stages of FSI and were challenged by the divergent perspectives groups and individuals had about the purpose of FSI and how it would achieve what it was set out to do. We also found that our partners had many other competing demands that made it difficult for them to engage with the MEL team.

7. A flexible and adaptable MEL system is paramount but within bounds

While the FSI team was fortunate to have been given the time to experiment, it ultimately resulted in considerable delays and we ran out of time to demonstrate a functioning and running MEL system for FSI. This detrimentally impacted the FSI team's capacity to demonstrate and communicate progress. More critically, we did not have an opportunity to assess the extent to which our MEL approaches and tools could have contributed to FSI's innovation capacity and to that of the Australian ODA agriculture and food system.

This highlighted for us the need to have some MEL processes and tools up and running early on, whilst also having space and time to be experimental, flexible and innovative with our approaches (see Guijt, 2007). The latter is critical. As our experience showed (as have others who have been engaged in similar spaces), we found that as we implemented different tools we encountered problems and disjunctures but also new insights. Moreover, as FSI evolved, the activities and needs of the FSI team and partners also changed. Thus the MEL system needed to be flexible enough to respond to all of these changes in order to ensure that it remained 'fit-for-purpose'. However, this process consumed most of our time and energy and, in the end, the initiative ended before we could fully implement the suite of 'most promising' MEL tools.

Summary: key take-home messages

Monitoring and evaluation, with an emphasis on critical reflection and learning, is an essential enabler of innovation. However, for M&E to effectively support reflexivity, learning, and capacity to innovate it too necessitates an enabling environment (Ramalingam, 2011; van Mierlo et al., 2010a). The conditions required for M&E to be mobilised in ways that can catalyse and support learning-oriented processes and outcomes are many of the enabling factors needed for innovation. These include the right individual and institutional capabilities, appropriate resources and incentives, and policy frameworks and institutional arrangements that support genuine collaboration, non-linear perspectives of change, and experimentation with emergent, novel, and innovative approaches (Hall et al., 2007, 2010; Hawkins et al., 2009;

Ison et al., 2014; van Mierlo et al., 2010b).This flags the importance of not only unpacking which enabling conditions are critical for MEL in programmes/projects underpinned by innovation systems thinking and which are realistically achievable, but also considering whether some need to be prioritised or even sequenced.

Our experience with MEL in FSI also reinforced what many others have underlined: that the paradigms that govern aid and development investments, interventions and other practices (e.g. M&E) are 'slow-moving' institutions (Roland, 2004). In practical terms, this means that *"small incremental changes to existing systems might be more feasible and workable than radical and abrupt changes that seek to impose blueprints from outside"* (Holvoet et al., 2012, p. 751). This highlights that our 'failures' with MEL should not come as a surprise. It also points to some practical considerations, including that for M&E *with learning* to be institutionalised, i.e. embedded and sustained as part of the fabric of routine practice, the M&E framework and tools should at the very least not conflict with (and, at best, be complimentary or supportive of) other required (usually conventional) M&E systems. The M&E tools and associated processes also need to be 'simple', i.e. non-cumbersome to engage with, implement, and sustain as institutionalised practice. This does not imply that they should be simplistic, i.e. tools that dumb down the complex reality of the issues at hand and of innovation processes. As noted by Ang (2011, p. 779), *"the development of sophisticated and sustainable responses to the world's complex problems requires the recognition of complexity…[along with] simplification to combat the paralyzing effects of complexity…".* The challenge is how to simplify MEL tools and processes in ways that support individuals and organisations to effectively navigate complexity and support innovation, without rendering them simplistic. Finally, also critical is having 'champions' who are embedded in the institutions and who can, over the long term, support and adapt M&E approaches and tools gradually towards greater reflection, learning and innovation.

A final take-home message that emerged from our experience in FSI is that 'perfect is the enemy of good'. We spent much of our time and resources on developing a comprehensive and robust MEL system. We should have been more willing to go with an imperfect plan and less-than-perfect tools and approach MEL as an "evolving practice" (Guijt, 2007) where we learn and innovate along the way. This "learn(ing) in 'real time'" (Ramalingam et al., 2014) is critical if we are to better understand, build, and implement MEL tools and systems that can support learning and innovation. In other words, what is needed is not simply best practice guidelines but 'living repositories' of systematically-collated and shareable reflections, lessons learned and practical advice garnered from experiences of applying M&E and Learning in complex and innovation contexts.

Acknowledgements
We thank Irene Guijt and Judy Oakden for their constructive insights and suggestions that helped us design a 'fit-for-purpose' MEL. Colleagues at DFAT – Jim Woodhill, Kris Hendrickx, Scott Bayley, and Elaine Ward, and at ACIAR – Andrew Alford and John Dixon, are also gratefully acknowledged for their feedback on specific MEL approaches and tools trialled. We would also like to thank the FSI team for their constant support, tremendous patience and invaluable insights throughout the process. Thanks are also due to Michaela Cosijn and Barbara van Mierlo for helpful comments on earlier drafts of this paper.

References

Ang, I. (2011). Navigating complexity: from cultural critique to cultural intelligence. Continuum 25(6): 779-794. doi:10.1080/10304312.2011.617873

Argyris, C., & Schön, D A. (1996). Organisational Learning II: Theory, Method and Practice. Wokingham, England: Addison-Wesley Publishing Company.

Boulton, J.G., Allen, P.M., & Bowman, C. (2015). Embracing Complexity: Strategic Perspectives for an Age of Turbulence. Oxford: Oxford University Press.

Burns, D., & Worsley, S. (2015). Navigating Complexity in International Development: Facilitating Sustainable Change at Scale. Warwickshire, UK: Practical Action Publishing.

Cundill, G., & Rodela, R. (2012) A review of assertions about the processes and outcomes of social learning in natural resource management. Journal of Environmental Management 113: 7-14.

Cundill, G., Lotz-Sisitka, H., Mukute, M., Belay, M., Shackleton, S., & Kulundu, I. (2014). A reflection on the use of case studies as a methodology for social learning research in sub Saharan Africa. NJAS - Wageningen Journal of Life Sciences 69: 39-47. doi:http://dx.doi.org/10.1016/j.njas.2013.04.001

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. doi:http://dx.doi.org/10.1016/S0308-521X(03)00128-8

Fabricius, C., & Cundill, G. (2014). Learning in adaptive management: insights from published practice. Ecology and Society 19(1). doi:10.5751/ES-06263-190129

Flood, R.L., & Romm, N.R.A. (1996). Diversity Management: Triple Loop Learning. Toronto, Ontario, Canada: Wiley.

FSI. (2016a). Food Systems Innovation - Partnering and Learning for Impact. Retrieved from http://foodsystemsinnovation.org.au/

FSI. (2016b). FSI Practice Notes. Retrieved from http://foodsystemsinnovation.org.au/type/practicenotes/

Guijt, I. (2007). Rethinking monitoring for collective learning in rural resource management. Paper for workshop Farmer First Revisited 12-14 December, Institute of Development Studies, University of Sussex, UK.

Guijt, I. (2008). Critical readings on assessing and learning for social change: a review. Paper prepared for the Power, Participation and Change Group, Institute of Development Studies (IDS), UK.

Guijt, I. (2010a). Accountability and learning: exploding the myth of incompatibility between accountability and learning. In J. Ubels, N. Acquaye-Baddoo and A. Fowler (Eds.) Capacity Development in Practice pp. 277-292. UK: Earthscan.

Guijt, I. (2010b). Rethinking monitoring in a complex messy partnership in Brazil. Development in Practice 20(8): 1027-1044. doi:10.1080/09614524.2010.513729

Guijt, I., Brouwers, J.H.A.M, Kusters, C.S.L., Prins, E., & Zeynalova (2011). Evaluation revisited: improving the quality of evaluative practice by embracing complexity: conference report. Wageningen, Centre for Development Innovation, Wageningen UR.

Hall, A., Sulaiman, R., Clark, V. N., & Yoganand, B. (2003). From measuring impact to learning institutional lessons: an innovation systems perspective on improving the management of international agricultural research. Agricultural Systems 78(2): 213-241. doi:http://dx.doi.org/10.1016/S0308-521X(03)00127-6

Hall, A., Janssen, W., Pehu, E., & Rajalahti, R. (2007). Enhancing Agricultural Innovation: How to Go Beyond the Strengthening of Research Systems. Washington, DC: The World Bank.

Hall, A., Dijkman, J.R., & Sulaiman, V. (2010). Research into Use: Investigating the Relationship between Agricultural Research and Innovation. United Nations University - Maastricht Economic and Social Research and Training Centre on Innovation and Technology, UNU-MERIT.

Hawkins, R., Heemskerk, W., Booth, R., Daane, J., Maatman, A., & Adekunle, A.A. (2009). Integrated Agricultural Research for Development (IAR4D). A Concept Paper for the Forum for Agricultural Research in Africa (FARA) Sub-Saharan Africa Challenge Programme (SSA CP). Accra, Ghana, FARA.

Holvoet, N., Gildemyn, M., & Inberg, L. (2012). Taking stock of monitoring and evaluation arrangements in the context of poverty reduction strategy papers: evidence from 20 aid-dependent countries in Sub-Saharan Africa. Development Policy Review 30(6): 749-772. doi:10.1111/j.14677679.2012.00597.x

Hughes, B., Hunt, C.T., & Curth-Bibb, J. (2013). Forging New Conventional Wisdom Beyond International Policing: Learning from Complex, Political Realities. Leiden, Boston: Martinus Hijhoff Publishers.

Ison, R., Carberry, P., Jocelyn, D., Hall, A., McMillan, L., Maru, Y., Pengely, B., Reichelt, N., Stirzaker, R., Wallis, P.J., Watson, I., & Webb, S.(2014). Programmes, projects and learning inquiries: institutional mediation of innovation in research for development. Outlook on Agriculture 43:165-172.

King, J., McKegg, K., Oakden, J., & Wehipeihana, N. (2013). Rubrics: a method for surfacing values and improving the credibility of evaluation. Journal of Multidisciplinary Evaluation 9(21): 11-20.

Kusters, C.S.L., Guijt, I., Buizer, N., Brouwers, J.H.A.M, Roefs, M., van Vugt, S.M., & Wigboldus, S.A (2015). Monitoring and Evaluation for Responsible Innovation. A conference report for M&E for Systemic Change 19-20 March, Centre for Development Innovation, Wageningen UR, Wageningen, the Netherlands.

Leeuwis, C., & Pyburn R. (2002). Wheelbarrows Full Of Frogs: Social Learning in Rural Resource Management. Assen: Royal van Gorcum.

Lynam, J. (2012). Agricultural research within an agricultural innovation system. In T.W. Bank (Ed.) Agricultural Innovation Systems: An Investment Sourcebook pp. 261-344. Washington, DC: The World Bank.

Oakden, J. (2013a). Evaluation rubrics: how to ensure transparent and clear assessment that respects diverse lines of evidence. Melbourne, Better Evaluation. Retrieved from http://betterevaluation.org/sites/default/files/Evaluation%20rubrics.pdf

Oakden, J. (2013b). Evaluative rubrics – passing fad or an important contribution to evaluation practice? Paper presented at the ANZEA Conference, Auckland, New Zealand. Retrieved from http://www.anzea.org.nz/wp-content/uploads/2013/10/Judy-Oakden-2013-Evaluative-rubrics-fad-orhere-to-stay.pdf

Oakden, J., & Weenink, M. (2015). What's on the rubric horizon: taking stock of current practice and thinking about what is next. Paper presented at the ANZEA Conference, Auckland, New Zealand. Retrieved from http://www.anzea.org.nz/wp-content/uploads/2015/07/150718-Oakden-WeeninkANZEA-vxx.pdf

Odame, H.H., Hall, A., & Dorai, K. (2012). Assessing, Prioritising, Monitoring, and Evaluating Agricultural Innovation Systems. In The World Bank (Ed.) Agricultural Innovation Systems: An Investment Sourcebook pp. 539-628. Washington, DC: The World Bank.

Oswald, K., & Taylor, P. (2010). A learning approach to monitoring and evaluation. IDS Bulletin 41(6): 114-120.

Pahl-Wostl, C., & Hare, M. (2004). Process of social learning in integrated resources management. Journal of Community and Applied Social Psychology 14: 193-206.

Patton, M.Q. (2011). Developmental Evaluation: Applying Concepts to Enhance Innovation and Use. New York: The Guilford Press.

Pretty, J.N. (1995). Participatory learning for sustainable agriculture. World Development 23: 1247-1263.

Ramalingam, B. (2011). Learning how to learn: eight lessons for impact evaluations that make a difference. ODI Background Note April 2011. Retrieved from http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/7096.pdf.

Ramalingam, B. (2013). AID on the Edge of Chaos: Rethinking Internal Cooperation in a Complex World. Oxford: Oxford University Press.

Ramalingam, B., Laric, M., & Primrose, J. (2014). From best practice to best fit: understanding and navigating wicked problems in international development. London, ODI. Retrieved from http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9159.pdf

Roland, G. (2004). Understanding institutional change: fast-moving and slow-moving institutions. Studies in Comparative International Development 38(4): 109-131. doi:10.1007/bf02686330

Stone-Jovicich, S. (2015). FSI Practice Note: to rubrics or not to rubrics? An experience using rubrics for monitoring, evaluating and learning in a complex project. Australia: Food Systems

Innovation. Retrieved from http://foodsystemsinnovation.org.au/resource/rubrics-or-not-rubrics-experience-usingrubrics-monitoring-evaluating-and-learning-complex

UNAIDS (2009). 12 Components Monitoring and Evaluation System Assessment Tool. Geneva, UNAIDS. http://www.unaids.org/sites/default/files/sub_landing/files/2_MERG_Strengthening_Tool_12_Compon ents_ME_System.pdf

USAID (2013). Discussion Note: Complexity-Aware Monitoring. Washington, DC: USAID. Retrieved from https://usaidlearninglab.org/sites/default/files/resource/files/Complexity%20Aware%20Monitoring%20 2013-12-11%20FINAL.pdf

USAID (2016). Monitoring, Evaluation, Research and Learning Innovations Programme (MERLIN). Retrieved from https://www.usaid.gov/GlobalDevLab/about/monitoring-evaluation-research-andlearning-innovations-programme

van Mierlo, B., Arkesteijn, M., & Leeuwis, C. (2010a). Enhancing the reflexivity of system innovation projects with system analyses. American Journal of Evaluation 31(2): 143-161. doi:10.1177/1098214010366046

van Mierlo, B., Leeuwis, C., Smits, R., & Woolthuis, R.K. (2010b). Learning towards system innovation: evaluating a systemic instrument. Technological Forecasting and Social Change 77(2): 318-334.

Wals, A.E.J., & Rodela, R. (2014). Social learning towards sustainability: problematic, perspectives and promise. NJAS - Wageningen Journal of Life Sciences 69: 1-3. doi:10.1016/j.njas.2014.04.001

Woodhill, J. (2007). M&E as learning: rethinking the dominant paradigm. In J. de Graaff, J. Cameron and S. Sombatpanit (Eds.) Monitoring and Evaluation of Soil Conservation and Watershed Development Projects. CRC Press.

Taking farmers on a journey: experiences evaluating learning in Farmer Field Labs in UK

Reed, M.[1], Ingram, J.[1], Mills, J.[1] and MacMillan, T.[2]

[1] *The Countryside and Community Research Institute, The University of Gloucestershire, UK*
[2] *The Soil Association, Bristol, UK*

Abstract: This paper presents results from an evaluation of the Soil Association's Duchy Originals Future Farming programme which supported regular farmer group meetings (Field Labs) and problem based field experiments. Drawing on the theoretical ideas of the three learning loops (Argyris & Schön, 1996), the paper examines the nature and extent of farmer learning that can be attributed to participation in the initiative. Using data from a survey, detailed interviews and a discussion forum, the evaluation found that farmer understanding of Field Lab topics, practices and skills, and research methods has been enhanced to different extents as a result of participating in Field Labs. However, overall farmer learning is as much about being given the tools and the confidence to go away and try things, as it is about acquiring specific knowledge, skills and practices. The paper concludes by reflecting on methodologies employed for evaluation of learning approaches and arguing for a more embedded and reflective approach.

Keywords: Field labs, farmer learning, participatory approaches, learning loops, evaluation

Introduction

A shift towards participatory and demand driven extension (Leeuwis et al., 2004) together with an increasing policy interest in peer to peer learning to foster innovation has led to the emergence of a range of participatory initiatives (operational groups, farmer field labs) in Europe. These are modelled on more established approaches such as Farmer Field Schools (FFS) and Stable Schools. A central element of these strategies is a learner-centred process that relies on discovery-based and experiential learning and critical reflection in groups. A common assumption is that these approaches lead to improved skills and knowledge, problem-solving and critical thinking, and enhanced empowerment and capacity building amongst participating farmers, as well as increased adoption and diffusion of sustainable and innovative practices. Verifying such claims and evaluating such approaches can be challenging methodologically (Douthwaite et al., 2003; Waddington & White, 2014).

This paper presents results from an evaluation of such an initiative, the Duchy Future Farming Programme (subsequently called Innovative Farmers). The programme[1] supports 'Field Labs', which are regular farmer (organic and conventional) group meetings, and problem based field experiments guided by a facilitator with research expertise (MacMillan & Benton, 2014). Drawing on the theoretical ideas of the three learning loops (Argyris & Schön, 1996), this paper examines the nature and extent of farmer learning that can be attributed to participation in the initiative. The paper highlights the challenges of attributing learning to such interventions

[1] *This Programme was initiated by the Soil Association in partnership with the Organic Research Centre and UK retailer Waitrose, with funding from the Prince of Wales's Charitable Foundation.*

particularly in a retrospective evaluation and argues for an ongoing and reflective evaluation for the successor programme.

Understanding the nature and extent of learning or improving knowledge associated with interventions has been the subject of much scholarship (Coudel et al., 2011). This scholarship reveals that multiple elements are contributing to the process of learning and acquiring knowledge (Baars, 2011) which any evaluation of farmer-centred learning initiatives should consider. In FFS evaluations, learning tends to be seen as enhanced knowledge of farming technology such as Integrated Pest Management practices and pest identification. Often these studies provide evidence via adoption, or, taking a broader view, of improved analytical skills, critical thinking, ability to make better decisions and familiarity with practices, which lead to better decisions regarding inputs, yields and costs (Waddington & White, 2014). At a deeper level, empowerment and enhanced capacity to learn are also indicative of improved and more transformative learning (Duveskog et al., 2011). The methodology in this evaluation sought to examine and understand these different dimensions of learning.

The SA define field labs as farmer-led meetings, open to all (both organic and non-organic), where producers examine innovative approaches, share existing best practice, learn how to run effective producer-led trials and identify real gaps where academic research would make a crucial difference. Field Labs are designed to empower farmers as innovators, increasing the impact of their informal research and enabling them to influence the formal research conducted in their name. They aim to build knowledge of specific topics, an understanding of the research process and associated skills, and develop critical thinking amongst individual farmers. Typically, Field Lab groups meet 3-4 times to address these elements. The FL were evaluated against these aims. Specifically, the aims of the evaluation reported here were to assess the effect of participation on: (i) farmer learning in farming techniques/best practice; and (ii) farmer learning in research/innovation skills/critical thinking and other learning outcomes.

Literature Review

This review of the literature builds on insights that focus on understanding the context in which farm business operates, the way in which the literature on farm learning has tended to not be attentive to the significance of the context, and in this point to the novel elements of the subject of the study.

The focus on innovation and learning has recently become a policy prescription on the basis that through Schumpterian effort meaningful economic growth can be resumed after the crisis of 2008. This implies in part that agricultural businesses and farmers were previously deficient in innovation and knowledge. In a British context this can be seen in the focus on export based agricultural technologies as well as food commodities in the Taylor Report (Herbert & Lord Taylor of Holbeach, 2010). Rather, as we argue, this paper illustrates the agricultural industry in the UK through self-organisation is attempting to realise important sustainability goals that it has set for itself. The wider British agricultural industry has already made clear its preference for farmer-led participatory approaches which are divergent from government policy and provision (Defra, 2013; Gibbs, 2013; OECD, 2015; Sutherland et al., 2013). This may have wider implications as others emulate the process or learn from it, but also speaks to the scholarship around farmer learning and innovation.

Much of the literature on innovation and participatory learning has its roots in learning with large organisations and the process of transformation that lead to organisational innovation, suggesting that this is not simply a question of agency but also of enabling structures (Coudel et al., 2011). In order to transport some of the insights about such learning scholars concerned with agriculture and farming have had to adapt to what Coudel and colleagues identify as more "loosely structured environments" (Coudel et al., 2011, p.121). This reflects a broader totalising tendency in much of the literature to assume that all agricultural contexts are sufficiently similar that comparisons and analogies can be made between projects and enterprises (Waddington & White, 2014).

The project which is the focus of this study typifies much of the context of British agriculture but differs in significant elements. Firstly, the food supply chains in the UK are retailer led with their influence not only influencing commodity prices but also on-farm practices, especially those related to food safety (Marsden et al., 1999). Farm level autonomy is circumscribed by such relationships as is the flow of information through and within the farm business. Secondly, the state has a significant role within the agricultural sector through agri-environmental interventions both in terms of legislation, but also through payment schemes to ensure environmental improvement (Mills et al., 2011). Again this represents considerable flows of not just money but information between and within farms (Mills et al., 2016). Many farms remain multi-generational, family owned and operated enterprises, that have high levels of capitalisation and technology, which is reflected in the developmental trajectories they have adopted (Ingram et al., 2013). Innovation at the farm level in this context is modulated by a complex, but not necessarily sophisticated, interaction between major private sector corporations, the implementation of EU schemes by various state agencies and devolved government bodies and familial requirements.

To leave the account of the context at this point would also reduce the complexity of the context. The project is led by an organic farming organisation, and most, but not all, of the participating farmers were also organic. Organic farming has been noted throughout the literature as having particular epistemic practices that differentiate it both practically and philosophically from much of the wider agricultural sector (Morgan & Murdoch, 2000). The practices of certification and market creation for the organic sector have been reflected in a particular organisational form as well as a distinct market profile, with a much greater uptake of schemes as well as a greater tendency towards use of some short supply chains (Lobley et al., 2009). Further scholars have argued that organic farming is part of a wider social movement that has wider civic and political goals that challenge both how agriculture is practised and its status within broader society (Reed, 2010). In recent years this has seen organic farmers allied with other groups in protests about GM crops and debates about the future role of agriculture (Reed, 2008). This suggests that agricultural sustainability innovation may have a wider social and civic impact that often recognised.

Recently social movement scholars have come to focus on the productive, epistemic actions that result from the collective action that is undertaken by movements as participants work on what they want to replace the present (Crossley, Melucci). Rao, writing from business studies notes how the rise of organic food was related to the role of activists, and from technology studies Hess argues that organic agriculture is the product of a social movement (Rao, Hess). This is significant as it indicates other social influences, in terms of flows of information and values, that are not often accounted for the in the farmer learning/innovation literature but also

another type of organisation, the social movement, which has a very high capacity to foster learning and innovation. Castells (2012) points to the innovations introduced through the 'Occupy' protests and the importance of internet augmented deliberations in a relentless process of interaction. It may be that through a movement many people have experienced profound changes in values or societal understanding that require pragmatic changes. The innovation literature can help in the analysis of how these changes occur.

Significantly, in their study of organisational innovation with regards to gay and lesbian advocacy, Foldy and Creed (1999) note the importance of wider social movement activism in fostering these changes. Through their detailed account of the activism within and without the businesses in questions they observe that "*a closer look reveals an intertwining of single-, double-, and triple-loop approaches, a maze that resists simplification"* (p.214). They do not abandon the schema but refine it to be able to analyse the particular, suggesting that whilst it is necessary for the first loop to succeed before second and third loops can, this sequence is neither linear nor easily observed to be serial, *"They happen concurrently, sometimes cross-fertilising and sometimes at cross-purposes, but ultimately, it is that continuous interaction out of which change efforts grow" (*p.224). This suggests that we need to be attentive to the dynamic ways in which analytical frameworks are heuristic guides, not to be mistaken for the reality they interpret.

Methods

Given the theoretical understandings of learning and the evaluation context (ex post), client requirements and aims, the methodology focused on evidence of learning in individuals as an indication of the FLs' impact (but on the understanding that any individual learning was most likely connected to wider group learning), and aimed to establish whether different levels of learning could be distinguished.

The evaluation took place towards the end of the three-year programme and the method was somewhat prescribed by the programme requirements. Three main methods were used in the following order: detailed interviews with farmers/growers; a facilitators' discussion forum; and a telephone survey by the research team of farmers/growers. Results from detailed interviews were used to inform and steer questions in the facilitators' discussion forum, likewise results from detailed interviews and the facilitators' discussion forum were used to develop the survey questionnaire. Also, the authors attended Field Labs towards the beginning of the evaluation and a facilitator/research workshop towards the end of the evaluation.

Purposeful sampling from a list of participants from 22 Field Labs was used to select 12 interviewees. These interviews were semi structured. Random sampling was used to select 30 telephone survey respondents from a sample of 221 farmers/growers and advisors (representing 14%). Interviewees and respondents were asked: to what extent they had gained new knowledge and information; learned new farming practices and skills; come to understand some of the underlying principles beneath these practices; come to understand and acquire research skills; and reflected on their learning overall. Survey respondents ranked statements using a Likert-like scale.

Drawing on qualitative and quantitative methods allows patterns as well as processes to be explored; the telephone survey reveals patterns in responses whilst the detailed interviews and discussion forum provide some explanation for these patterns and insight into processes

involved. In this respect, detailed interviews allow some in-depth analysis of the nature of learning while the survey can be used to extend this analysis and obtain some idea of the extent of farmer learning. This mixed methods approach also allows some triangulation, for example, farmer self-reported learning or practice change could be validated by facilitators (Moran-Ellis et al., 2006).

Results

Good correspondence between the three data sources confirms that topic knowledge, practices and understanding, and research understanding has been enhanced to different extents. It is apparent that significant learning has occurred as a result of attending Field Labs. Selected survey results are presented in Figure 1. The differences in the nature of the learning can be described in terms of the concept of learning loops, originally developed by Argyris and Schön (1996)[2] to explain the types of learning that take place in organisations. The concept has been applied to natural resource management, adaptation and farming contexts (Duveskog et al., 2011; Eshuis & Stuiver, 2005). The intention here is to distinguish and frame the results using a simple interpretation of learning loops and not to provide a thorough analysis and critique of learning loops theory.

Single loop learning –improved learning about the Field Lab topics and practices

Single loop learning is understood here as changing the way of working within a set frame of thought through incremental learning. The focus is on techniques and practical and locally applicable answers to questions rather than questioning underlying principles. It is akin to the 'know-what' and the 'know-how' described by Lundvall and Johnson (1994).

The majority of survey respondents agreed that the Field Lab gave them a clearer understanding of the topic they were investigating; those who did not agree already possessed high levels of knowledge about the topic. Although the majority (80%) were satisfied with the level of learning in the Field Lab, only 37% agreed that the Field Lab they attended gave them a chance to learn new skills and practices such as how to grow cover crops effectively. The gap between these figures is accounted for by those who had not attended long enough to engage in learning or develop new skills.

Acquiring new knowledge from others at the Field Lab was mentioned by most of the interviewees who agreed that they had learned new facts about the Field Lab topic by sharing information with co-participants. According to the detailed interviews this learning depended on the nature of the topic, and on the baseline knowledge of the participant. Participants agreed that the format of the Field Lab in most cases allowed a good combination of technical, practical and financial information to be discussed. For some the Field Lab format, compared to other similar formats, was thought to provide a good context for learning about facts and figures. As one farmer remarked *"They are more focused than a farm demonstration or walks and as trials are undertaken in a semi-controlled environment they should produce better facts and figures".*

Larger groups with a "good healthy breadth of views and experiences," including some experts were considered to lead to most learning of this type compared to Field Labs described as not particularly participatory or with only a few participants. For more informed participants the

[2] Double and single learning loops are part of a broader concept of organisational learning theory developed by Argyris and Schon in the 1970s, and later expanded by several other organisational thinkers (to include for example triple learning loops) (Foldy & Creed, 1999)

extent of this learning was not that pronounced. For some it was described as more about "joining the dots" rather than picking up specific information about the topic.

Double loop learning – learning about underlying principles

Double loop learning refers to learning that alters underlying values, rules and assumptions. This evaluation found evidence of learning about the principles underlying the topics and practices in Field Labs and learning about how to apply experimental and research protocols on-farm. This is akin to the acquiring the 'know-why' dimension of knowledge described by Lundvall and Johnson (1994).

Four of the interviewees agreed that the Field Lab had enhanced their understanding of the basic principles underlying the new techniques and measures. For example, for the Foam Weeding Field Lab, one interviewee said *"Yes we did learn more about the underlying science behind the technology, particularly the role of the foam as a wetting agent, and the role that played".*

In some cases Field Lab participants felt they had been given the tools, and the confidence, to go away and try things for themselves, as this comment demonstrates: *"It was more theoretical, we walked and talked and had a look at things, although we did handle compost, and the final mix etc. We got the basic recipe to go away and experiment with. We would have been confident to try it out."* (Wood chip compost).

Figure 1. Responses to survey questions

This theme of taking control of details and feeling enabled was picked up by an interviewee participating in a weed control Field Lab: *"I was trained as a conventional farmer and we spent a lot of time looking at detail for organic farmers. The temptation is to think that you can't do anything other than sow your crop and hope for the best. This brought us back to the details which I think have been missing in many organic situations. What happens is that you think things are so huge, subject to so many vagaries, and out of your control, but this helped remind us there are things you can control."*

With respect to learning how to conduct research, survey respondents were less likely to agree that they had learned research skills and results suggested that learning about research methods was not something they had considered as relevant. However, some 40% agreed they wanted to get more involved in research as a result of attending a Field Lab. Also notably a number of respondents and interviewees were already well versed in research skills and understanding. The interviews provided more depth on this subject, revealing how some farmers valued, and learned from, carrying out trials. For some participants the practical hands-on measuring was a distinctive element of the Field Lab. As one interviewee who had attended the 'Compost Teas' Field Lab remarked: *"Basically we were physically doing the trial ourselves, actually doing it, not a researcher doing it and reporting back."*

Triple loop learning – learning how to learn

Triple loop learning here refers to learning how to learn. It allows participants to reflect on and learn how to evaluate and appreciate their own experiences and viewpoints, as well as those of others. This is aligned to transformative learning impacts which entails a deep seated shift in perspective (Duveskog et al., 2011; Percy, 2005).

The nature and extent of this deeper learning is hard to gauge from the survey, with responses to statements about the nature of learning inconclusive. The results from the interviews, facilitators' discussion forum and participant observations, however, show that learning in Field Labs is as much about changing perspectives as learning new facts or practical skills. Interviewees explained that the Field Labs made them question things and as one participant explained *"it introduced a different way of thinking about the problem".*

Another element of this learning is building confidence in decision making. For example, in a Field Lab experimenting with mastitis control in cattle some farmers felt empowered to manage mastitis more effectively by either using herbal treatments or being 'brave enough' to make decisive management changes. The openness and sharing ethos of working in a group were highlighted as important in instilling confidence and a sense of empowerment. As with first and second learning loops this was enhanced when groups were of a sufficient number (generally >10), had a good mix of participants and provided a good breadth of views and experiences. The opportunity for Field Labs to develop this learning over time relies on sufficient continuity within groups with participants committing and returning to events and reflecting together on outcomes. The commitment, enthusiasm, honesty and expertise of host farmers were also seen to generate effective group learning and inspire confidence.

Facilitators suggest that the whole Field Lab process is one of deeper learning. In the early stages, for example, farmers learn how to formulate and agree on ideas to test. As one facilitator remarked, this learning involves *"knowing to ask the right questions".* With respect to setting up research in the Field Labs, facilitators focused on leading the farmers through the research process. They described how some farmers are inclined to just test "with and without" rather than setting up randomised trials, or to be over ambitious with their research questions, one facilitator explains: "*it's taking them through this process, and the how do you measure it? What are the parameters you are measuring? It's saying to growers how are we going to do it then? It demonstrates to them that doing research isn't like falling off a log".*

According to facilitators this realisation is part of the Field Lab learning process that farmers should go through, although the extent to which different farmers do so is unclear and hard to gauge or measure directly. Overall, rather than point to specific learning achievements,

facilitators said that Field Labs were more about "taking them [farmers] on a journey to understand how to look at their farms". This again highlights the importance of building up continuity within groups, which facilitators noted was absent in many Field Labs.

Although facilitators have these aspirations for farmer learning, they voiced concerns that farmers did not always understand what a Field Lab entails. Several had clear examples of farmers not understanding what they were participating in: *"Is it what they thought it was when they signed up? It is clear in the information, but still people turn up and they don't understand the concept of the Field Lab – and why should they? It's the topic that's drawing them in, not the process."*

The management of expectations had two facets; as above those of the farmers, but also those of the researchers: *"There is a perception that farmers expect to get spoon-fed by researchers and once they understand it is not like that then they are much more likely to get involved …."*. In the more successful and longer term Field Labs there was a sense that those perceptions are changing, that dialogue and reflection between farmers and researchers is leading to more open-minded and appreciative learning.

Discussion

These results suggest that the Field Labs are supporting and building farmer innovation capacity. They do this by: facilitating collaboration for sharing and processing information and knowledge; enabling farmers to identify and prioritise problems and opportunities through experimentation; fostering confidence in practical and research skills; and nurturing new perspectives and outlooks.

The results highlight the difficulties in assessing changes in learning ascribed to Field Labs, especially in regard to evaluating improvements in farmers' analytical skills, critical thinking and ability to make better decisions; as well as changes at a deeper more transformative level. A recent meta-analysis of FFS programme evaluations noted the difficulty in ascribing changes specifically to the intervention. It concluded that the evaluations are broadly not sufficiently rigorous and there are dangers of 'systematic overestimation of impact' particularly with respect to diffusion and scaling up (Waddington & White, 2014). This review and others challenge the notion of a causal chain or linear outcomes in complex situations and identifies an 'attribution gap' (Douthwaite et al., 2003). The review of FFS also highlighted the challenge in evaluating softer outcomes such as empowerment and capacity development compared to other impacts where studies have shown improved farmer knowledge and adoption of beneficial practices with participants feeling more confident with problem solving and decision making (Duveskog et al., 2011; Van den Berg & Jiggins, 2007; Waddington & White, 2014).

The nature of the Field Labs programme and activities further challenged the evaluation, particularly in attributing farmer learning to participation in the Field Labs. This was because the labs were at different stages, a number having not yet developed sufficiently to result in measurable change in learning; and the diversity of contexts, activities, topics and goals made comparison difficult. Further, the majority of respondents had only attended one meeting and as such their perspectives about, and opportunities for learning were limited, The different levels of participation and engagement in groups clearly has an impact on learning, as has been widely noted in other studies showing how learning is embedded in social and cultural contexts, and that people learn through their ongoing participation in these contexts (Sewell et al., 2014). Participants also started from different levels of understanding. Significantly a

number were innovative farmers already experienced in (formal and informal) on-farm experiments, accustomed to finding solutions, and each with their own experiential knowledge to impart. Given these conditions, any evaluation has to recognise the situated and contingent environment within which farmers learn (Eshuis & Stuiver, 2005) and the contributions that farmers can make as co-learners (Baars, 2011).

The results of this study suggest the need for evaluation to be embedded within the initiative as a form of action research with in built M&E so that it can evolve, adapt and learn as it develops, allowing ongoing reflection to foster learning and innovation. In turn creating a virtuous loop of credible critical learning that will enable farmers to guide their farms to their ends within a context in which they are the experts.

Conclusion

The evaluation found farmer learning as a result of participating in Field Labs is as much about being given the tools and the confidence to go away and try things as it is about acquiring specific knowledge, skills and practices. The learning process within Field Labs is about "joining the dots" and learning how to formulate ideas and "to ask the right questions" rather than specific skills or techniques. Although single loop learning is occurring, double loop is more apparent. Field Labs also aim to enhance farmers' critical thinking and help them 'learn how to learn', yet, whilst there is some evidence for this, there is an indication that not all farmers understand the concept and ambition. Overall Field Labs were described as 'taking farmers on a journey' rather than achieving defined outcomes. This approach presents particular challenges for evaluation unless the evaluators are part of that journey in a continuous process of M&E and critical reflection. The successor programme (Innovative Farmers) has made provision for this by incorporating M&E from the outset through an action research approach. It has also brought in changes in response to this evaluation including a stronger emphasis on group continuity whereby farmers become member of groups that meets on an ongoing basis, running successive Field Labs.

References

Argyris, C., & Schön, D.A. (1996). Organisational Learning II. Theory, Method, and Practice. Boson: Addison-Wesley.

Baars, T. (2011). Experiential science; towards an integration of implicit and reflected practitioner-expert knowledge in the scientific development of organic farming. Journal of Agricultural and Environmental Ethics 24(6): 601-628.

Coudel, E., Tonneau, J.-P., & Rey-Valette, H. (2011). Diverse approaches to learning in rural and development studies: review of the literature from the perspective of action learning. Knowledge Management Research & Practice 9: 120-135.

Defra. (2013). Review of Environmental Advice, Incentives and Partnership Approaches for the Farming Sector in England. Retrieved from London:

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.

Duveskog, D., Friis-Hansen, E., & Taylor, E.W. (2011). Farmer field schools in rural Kenya: a transformative learning experience. Journal of Development Studies 47(10): 1529-1544.

Eshuis, J., & Stuiver, M. (2005). Learning in context through conflict and alignment: farmers and scientists in search of sustainable agriculture. Agriculture and Human Values 22(2): 137-148.

Foldy, E.G., & Creed, W.D. (1999). Action learning, fragmentation, and the interaction of single-, double-, and triple-loop change, a case of gay and lesbian workplace advocacy. The Journal of Applied Behavioural Science 35(2): 207-227.

Gibbs, C. (2013). The Value of Advice Report. Retrieved from Peterborough:

Herbert, N., & Lord Taylor of Holbeach. (2010). Science for A New Age of Agriculture (The Taylor Report). Retrieved from London: http://www.conservatives.com/~/media/Files/.../taylor-review-agriculture.ashx?

Ingram, J., Gaskell, P., Mills, J., & Short, C. (2013). Incorporating agri-environment schemes into farm development pathways: a temporal analysis of farmer motivations. Land Use Policy 31: 267-279. doi:10.1016/j.landusepol.2012.07.007

Leeuwis, C., Leeuwis, C., & Ban, A. (2004). Communication for rural innovation: Wiley Online Library.

Lobley, M., Butler, A., & Reed, M. (2009). The contribution of organic farming to rural development. Land Use Policy 26(3): 723-735. doi:10.1016/j.landusepol.2008.09.007

Lundvall, B.-Ä., & Johnson, B. (1994). The learning economy. Journal of Industry Studies 1(2): 23-42.

MacMillan, T., & Benton, T. (2014). Engage farmers in research. Nature 509: 25-27.

Marsden, T., Flynn, A., & Harrison, M. (1999). Consuming Interests: The Social Provision of Food. London: UCL Press.

Mills, J., Gaskell, P., Ingram, J., Dwyer, J., Reed, M., Short, C., & Pike, T. (2016). Engaging farmers in environmental management through a better understanding of social psychological factors. Agriculture and Human Values (In press).

Mills, J., Gibbon, D., Ingram, J., Reed, M., Short, C., & Dwyer, J. (2011). Organising Collective Action for Effective Environmental Management and Social Learning in Wales. The Journal of Agricultural Education and Extension 17(1): 69-83. doi:10.1080/1389224x.2011.536356

Moran-Ellis, J., Alexander, V., Cronin, A., Dickinson, M., Fielding, J., Sleney, J., & Thomas, H. (2006). Triangulation and integration: processes, claims and implications. Qualitative Research 6(1): 45-59.

Morgan, K., & Murdoch, J. (2000). Organic vs. conventional agriculture: knowledge, power and innovation in the food chain. Geoforum 31(2): 159-173.

OECD. (2015). Fostering Green Growth in Agriculture. The Role of Training, Advisory Services and Extension Initiatives. Retrieved from Paris:

Percy, R. (2005). The contribution of transformative learning theory to the practice of participatory research and extension: theoretical reflections. Agriculture and Human Values 22(2): 127-136. doi:10.1007/s10460-004-8273-1

Reed, M. (2008). For Whom? The Governance of the British Organic Movement. Food Policy (In press).

Reed, M. (2010). Rebels for the Soil. The Rise of the Global Organic Food and Farming Movement. London: Earthscan.

Sewell, A., Gray, D., Blair, H., Kemp, P., Kenyon, P., Morris, S., & Wood, B. (2014). Hatching new ideas about herb pastures: learning together in a community of New Zealand farmers and agricultural scientists. Agricultural Systems 125: 63-73.

Sutherland, L.A., Mills, J., Ingram, J., Burton, R.J., Dwyer, J., & Blackstock, K. (2013). Considering the source: commercialisation and trust in agri-environmental information and advisory services in England. Journal of Environmental Management 118: 96-105. doi:10.1016/j.jenvman.2012.12.020

Van den Berg, H., & Jiggins, J. (2007). Investing in farmers—the impacts of farmer field schools in relation to integrated pest management. World Development 35(4): 663-686.

Waddington, H., & White, H. (2014). Farmer Field Schools. From Agricultural Extension to Adult Education. 3rd Systematic Review Summary 1. Retrieved from London: