# Reflexive monitoring and evaluation as a tool to stimulate peer-to-peer learning and impact during on-farm demonstrations

Laure Triste[a], Boelie Elzen[b], Hanne Cooreman[a], Anda Adamsone-Fiskovica[c], Frank Wijnands[d], Fleur Marchand[a], Herman Schoorlemmer[b]

[a] *Research institute for agriculture, fisheries and food (ILVO),* laure.triste@ilvo.vlaanderen.be, hanne.cooreman@ilvo.vlaanderen.be, fleur.marchand@ilvo.vlaanderen.be

[b] *Wageningen Plant Research,* boelie.elzen@wur.nl, herman.schoorlemmer@wur.nl

[c] *Baltic Studies Centre,* anda@lza.lv

[d] *Sadly, Frank past away during the writing on this paper.*

Abstract: Over the past two years, two AgriDemo-F2F and PLAID (H2020) have investigated peer learning in the context of on-farm demonstrations across Europe. Both projects will produce their results in the coming months. Early 2018, NEFERTITI (H2020) started with the objective to boost peer learning through the development of 10 supra-regional networks of on-farm demonstrations in Europe. The NEFERTITI demonstration networks will build on the results from AgriDemo-F2F and PLAID.

The assessment of P2P learning during farm demonstration events is key in the three projects. The projects used a series of 'Monitoring and Evaluation' (M&E) methods and tools, such as observation templates, interview guides, participant surveys, journals, and group discussions. Whereas, in AgriDemo-F2F and PLAID the M&E methods and tools are aimed at M&E performed by and for actors external to the case demo events, NEFERTITI additionally also focusses on internal M&E performed by and for the organisers of the demonstrations events, thus creating tools for reflection.

An evaluation of the M&E tools of the three projects indicates that evaluating peer learning during demonstrations is far from easy. Shortcomings of the M&E tools can be

attributed both to the (lacking) scientific rigour of the data provided by the tools in specific fields and to the ease to use the tools. The contradiction between scientific rigour and the user friendliness is a specific challenge for NEFERTITI as they have to be used by both organisers of demonstrations who often lack experience in using M&E tools and by researchers who will analyse the monitored data to come up with overall conclusions and recommendations.

Keywords: On-farm demonstrations; monitoring & evaluation; peer-to-peer learning, professionalising demonstrators

## 1. Introduction

It has long been recognised that farmers enthusiastically engage with knowledge exchange through demonstration activities (Ingram et al., 2018), and that innovation and uptake of new farming technologies or practices result from interactive engagement with likeminded individuals or groups (Burton et al., 2017; Koutsouris et al. 2017). This is often referred to as peer-to-peer (P2P) learning (Koutsouris et al. 2017).

Over the past two years, two on-going H2020 projects, notably AgriDemo-F2F and PLAID, have investigated this peer-to-peer learning in the specific context of on-farm demonstrations in a variety of cases across Europe. Both projects will produce their final results by mid-2019. Early 2018, another four-year project has started by the name of NEFERTITI with the objective to boost peer-to-peer learning through the development of 10 supra-regional thematic networks of on-farm demonstrations across Europe. These NEFERTITI demonstration networks, each consisting of 4-5 regional hubs, will build on the results from AgriDemo-F2F and PLAID.

A key topic in all three projects deals with the assessment of P2P learning processes during on-farm demonstrations. To do so, the two initial projects used a series of 'Monitoring and Evaluation' (M&E) methods and tools, such as observation protocols, interviews,

participant surveys, monitoring journals, and group discussions. In NEFERTITI, more advanced methods for M&E are developed, referred to as 'reflexive M&E'. This form of M&E has two main objectives: (i) to assess if and to what extent peer-to-peer learning takes place at demonstrations; (ii) to stimulate the processes that are observed, in this case the peer-to-peer learning.

Realising this dual ambition appears to be quite a challenge. Firstly, the experiences from AgriDemo-F2F and PLAID indicate that measuring the extent to which peer-to-peer learning at demonstrations has been taking place is far from easy. Furthermore, M&E tool development in NEFERTITI shows that it is difficult to balance 'scientific rigour' and 'usefulness in practice'. On the one hand, these tools have to be used by organisers of demonstrations who often lack experience in using this type of tools and therefore require simple ones. On the other hand, the researchers who investigate the M&E results to come up with overall conclusions and recommendations, need comprehensive data that would require more advanced tools.

The aim of this paper is to reflect on the M&E tools used in AgriDemo-F2F, PLAID and NEFERTITI to measure peer-to-peer learning, and to discuss their shortcomings, successes and use. This reflection will be used as an input for a workshop at the ESEE 2019 conference. In this workshop we seek to make a further step in developing effective M&E approaches for demonstration projects. During this workshop (i) the reflection on the M&E methods and tools used in AgriDemo-F2F, PLAID and NEFERTITI will be presented, and (ii) ideas to improve M&E tools to evaluate and stimulate P2P learning during demonstrations will be harvested and discussed with the workshop participants.

To set the scene, this paper first describes the context of farm demonstrations, the role of M&E in demonstration events and what to take into account to monitor and evaluate them. Secondly, we go into detail on the M&E tools used/developed in AgriDemo-F2F, PLAID and

NEFERTITI. Thirdly, we assess the tools of AgriDemo-F2F and PLAID with regards to their effectiveness in evaluating P2P learning and we discuss the main shortcomings and merits of the M&E tools of AgriDemo-F2F and PLAID.

## 2. M&E in farm demonstrations

### *2.1 The context of farm demonstrations*

There are various pressures on farmers to innovate. Some of these are internal to the agro-food system, e.g. decreasing soil health due to mono-cultures or increased competetion between farmers. Others are induced by societal or political pressure due to 'side effects' of farming systems on the wider environment, e.g., pollution of surface waters from nutrients or herbicides.

Demonstration events (further referred to as *demo* events) aim to show innovation within a working farm context or within a local setting to increase the awareness and understanding of the innovation. Demo events can have various formats, but they all are based on knowledge exchange between farmers (peer learning) and between farmers and other innovation actors (advisers, researchers, input providers…). These exchanges can have multiple forms and often a combination of types of knowledge exchange occurs during a demo event depending on the objectives, e.g. dissemination of knowledge, provision of advice and solutions, co-design of tools or conducting research (source: Guidelines Hub Campaign Plan Hub M&E Journal, NEFERTITI, 2019).

AgriDemo-F2F and PLAID showed that a variety of activities can constitute a demo event. Variation can be attributed to the type of demo activity (e.g., indoor presentations or field walks), the type of demonstrator (e.g., farmers, farming advisors, researchers), the type of knowledge exchanged (e.g., know-why, know-what, and know-how), the type of

participants (e.g., farmers, advisors, businesses, teachers and students, policy officials, general public), or the type of facilitation of knowledge exchange.

Demo events are often not a standalone activity, but occur in a broader range of activities with the overall ambition to stimulate innovation to make agriculture more sustainable. In the context of demonstration, 'an innovation' should be interpreted as 'something new to the demo event participant'. It may even refer to the re-introduction (often after adaptation) of century old practices or crop varieties in present-day agriculture. Innovations can have a big impact on different aspects of the farm system, and the adoption of innovations is dependent on a lot of other factors (e.g. legislation, market, income, neighbours, …).

### *2.2 The role of M&E for demo events*

Monitoring and evaluation are often used together and seem different to separate (World Bank, 2000). Monitoring refers to continuously collecting information by observing, measuring, documenting and verifying what is going on in relation to planned activities and objectives. Evaluation refers to analyzing information and making ex-post assessments based on qualitative or quantitative interpretation of activities, behaviors or outcomes against criteria or standards, in reference to pre-established objectives and expectations (World bank, 2000). Evaluation usually aims at determining effectiveness ("doing the right things", completing activities and achieving goals) and efficiency ("doing the things right", optimal use of resources). According to the World Bank (2000, p. 3), "*monitoring (in the sense of continuous collection of data on a project) is required both for implementation and ex-post assessment, whereas evaluation (in the sense of making judgments and decisions about a project) is also a continuous process during project implementation and after completion".* Both monitoring and evaluation can be formal and/or informal. Criteria can be explicit or

implicit and M&E can be carried out continuously and/or at specifically agreed time intervals (also ex-ante, mid-term, ex-post). Both monitoring and evaluation can focus on different levels of activity, using a range of supporting tools and techniques (Elzen et al., 2019).

Further, literature on M&E distinguishes between internal versus external evaluation (Conley-Tyler, 2005). Internal evaluation refers to evaluation performed by members of the organisation. In the case of the demo projects, for example, this can be the organisers of demo events. External evaluation is performed by non-members of the organisation. For the demo projects, for example, this can be the project partners.

Besides the distinction in who performs the evaluation, one should also take into account the purpose of the M&E, or in other words, the audience for who the evaluation serves. For example, the purpose or audience might be "internal", when the aim of the evaluation is to evaluate and to improve a specific programme of demo events or to improve the performance of a team organizing demo events. For such evaluations, the involvement of organisation members in the M&E process seems logic. On the other hand, the M&E purpose might be "external" when the monitored information is analysed and evaluated to contribute to research insights or protocols who serve a wider audience than the individual organisation (Conley-Tyler, 2005). Based on these distinctions, Conley-Tyler (2005) distinguishes 4 types of arrangements for M&E depending on who the evaluator and who the audience is: (i) Type 1: insiders for insiders, e.g. the organisers of demo events evaluate them to improve their future demo events, (ii) Type 2: insiders for outsiders, e.g. organisers of demo events evaluate their demo events to prove their succes to their funders or stakeholder, (iii) Type 3: outsiders for insiders, e.g., non-members of the demo organisation evaluate a demo event to provide lessons learned to the organisers of the demo event, (iv) Type 4: outsiders for outsiders, e.g., non-members of the demo organisation evaluate demo events to come to recommendations

and best practices for a wider public of demo organisations. We will use this distinction throughout this paper.

## 3. Monitoring and Evaluation tools developed and used in the FarmDemo projects

### *3.1 M&E tools used in PLAID and AgriDemo-F2F*

Many organisers of the demonstration events studied during PLAID and AgriDemo-F2F already evaluated demonstration events afterwards. Especially, organisations that frequently organise demonstration events thus try to learn on how to improve the set-up of a demo event (referring to the M&E type: insiders for insiders). However, only few of these evaluations happened in a structured way and were usually carried out on the basis of what the organisers had observed themselves and what their 'impressions' were.

In AgriDemo-F2F and PLAID by contrast, a systematic approach was proposed to monitor and evaluate the demo case studies. However, this approach was developed with the aim to be implemented by project partners (so not members of the organisation of the demo event) and with the aim to retrieve information on best practices for the organisation of demo events (thus also for an audience external to the specific demo event). Based on the M&E typology presented in 2.2, the AgriDemo-F2F and PLAID M&E approaches, can be categorized as Type 4, outsiders for outsiders.

In both projects, a variety of methods and tools were developed. To evaluate the preparation of the demo event, interview guides for interviews with the organisers, host farmer, demonstrators were drafted. To monitor the set-up of the demo event, the participants and the knowledge exchange activities, observation protocols were developed. To evaluate the effectiveness and impact of the demo events, the projects designed questionnaires for on-site participant surveys, (telephone) interview guides for follow-up interviews with organisers and

demonstrators, (telephone) post-event surveys to interview participants in the months after the event, and guidelines for focus group discussions.

The M&E performed within the AgriDemo-F2F project has the aim to evaluate how different structural and functional aspects of demo events contribute to the effectiveness of learning during the demo events (Cooreman et al., 2018). Therefore, they took into account the extent of the learning (amount of participants stating they have learned, and the extent of change in practice) and the nature of learning. To define the nature of learning, they focussed on the single, double and triple loop outcomes of the learning processes induced by the demo events (Argyris & Schön, 1996). Single loop learning (SLL) refers to acquiring factual knowledge and developing skills in order to manage problems on a daily basis (e.g. knowing how to apply an irrigation scheme/technology or pesticide). Building on SLL, double loop learning (DLL) explores the underlying values and assumptions, and requires critical reflection. This refers to a deeper level of learning, requiring metacognitive skills to develop an awareness of the own thinking (e.g. getting insights in the question: "Why is my farming system the way it is and should I change my farming system?"). Triple-loop learning (TLL) is conceived as reflecting on how we organize ourselves to make decisions, what kinds of principles we apply to decide whether something is right or better, and whether such principles need to be changed (Romme & Van Witteloostuijn, 1999). DLL and TLL learning help people understand which strategy works better to achieve a goal, leading to better mid- and long-term solutions (Hummelbrunner, 2015).

In M&E performed in PLAID is aimed at assessing the extent to which individual elements of a demo contribute to the success of a demo event. Theses individual elements were classified along the following categories: (i) inputs (infrastructure, finances, human resources); (ii) access (geographic, social, economic) (incl. gender, age, income, stakeholder groups); (iii) demonstration process (methods, content, interaction form). The assessment of

success of demonstration is defined as the contribution to five key impact domains related to the farm/farmer (with knowledge and learning as underlying all of those – economic,social, and environmental – dimensions). Those are (Burton et al. 2017, p. 15): *"(i) Productivity & profitability (efficiency of production; increased output per unit of input; increased ability to produce a return on an investment, make profit); (ii) Resilience (improved capacity to adapt to changes); (iii) Environmental sustainability (responsible use of and attitude towards environmental resources); (iv) Quality of life (improved material/working/health/safety/leisure conditions); (v) Empowerment (enhanced self-reliance, skills, social capital)"*. Further, they take into account the differentiation between immediate/direct and longer-term/broader effects, as follows (Burton et al. 2017, p. 15): " *(i) Outputs – results achieved immediately after implementing a demonstration activity (e.g., the number of trained farmers); (ii) Outcomes – later (medium-term) changes that have occurred as a result of a demonstration activity (e.g., application of the gained knowledge by farmers on their own farms); (iii) Impacts – broader (long-term) changes affecting direct beneficiaries of a demonstration activity or a wider community/institutions/environment that become evident several years after the activity has taken place (e.g., increased annual productivity levels of local farms."*

The tools developed in these projects were used by the project partners of each project to evaluate a couple of demo events in their region. The feedback was afterwards shared with the organisers of the demo events, which they indicated as useful to plan future demo events. All data were collected and analysed to define good practices for farm demonstration within the two projects.

Table 1. Overview of the aspects of a demo event that can be monitored (M)/ evaluated (E) with the M&E tools of PLAID and AgriDemo-F2F. For both projects, these tools were based

on a 'case study methodology' that specified the research questions for the various demonstration aspects.

| **Aspects monitored/evaluated** | **Tools in PLAID** | **Tools in AgriDemo-F2F** |
|---|---|---|
| Demo organisation process (e.g., actors involved, decision on the topic and set-up) | Interviews with organisers (M)<br>Participatory observation in preparatory meetings with demo organisers (M)<br>Evaluation meeting with demo organisers (E) | Interview guides for interviews with organisers, host farmer and demonstrator (M)<br>Participant survey (M) |
| Host farm facilities during the demo event | Observation tool (M)<br>Participant survey (M) | Observation tool (M)<br>Participant survey (M) |
| Demo set-up (demo topic/object of demonstration, demonstrator, activities for knowledge exchange) | Observation tool (M)<br>Participant survey (M)<br>Evaluation meeting with demo organisers (E) | Observation tool (M)<br>Interview guides for interviews with demonstrator (E)<br>Participant survey (E) |
| Participants motivations and characteristics | Participant survey (M) | Pre-survey before the demo event for participants (M) |

| **Aspects monitored/evaluated** | **Tools in PLAID** | **Tools in AgriDemo-F2F** |
|---|---|---|
| Learning outcomes (single, double, triple loop learning / know-what, know-why, know-how) | Participant survey (M) | Participant survey (E)<br>Telephone survey with participants (E) |
| Impact / learning effectiveness of the demo | Focus Groups with demo participants (M & E) | Participant survey (E)<br>Telephone survey with participants (E)<br>Interview guides for interviews with demonstrator (E)<br>Guidelines for focus group discussions. (E) |

### *3.2 M&E tools in the NEFERTITI project*

The M&E tools in PLAID and AgriDemo-F2F were primarily intended to analyse demo events and to provide recommendations on how to increase their effectiveness (Type 4 M&E: external for external). The lessons learned of these two projects are used in NEFERTITI, that aims at disseminating these lessons amongst organisers of demonstrations across Europe. In the 10 NEFERTITI demonstration networks, five demo events per year will

be organised in each regional hub (45 hubs across Europe) during three demo campaign years. In total this will then cover around 750 demonstrations in the years 2019-2021.

Also in NEFERTITI, a structured M&E approach is developed. This M&E approach will serve two general functions. First, the approach aims to increase the capacity for 'self-assessment' in the hubs and networks, to help them realise their objectives in an optimal way. So it allows to add reflexivity to the development process of a (series of) demo event(s), by incorporating continuous or periodic assessment activities. This related to M&E type 1: internal for internal, as mentioned in section 2.2. Second, the approach aims to gather and report on the functioning of ongoing activities, what is learned in the hubs, networks and the wider agricultural knowledge and innovation system. The monitoring of these activities are used for later analysis to provide recommendations on networking and farm demonstrations to improve future demo events (Elzen et al., 2019). This second aim is related to M&E type 2 and type 3 (section 2.2), as both people internal to the organisation of the demo event, as project partners (researchers) external to the demo event (to perform higher level analysis) will contribute to the M&E. So compaired to AgriDemo-F2F and PLAID, the NEFERTITI M&E approach has the additional characteristic that it will also be performed and used by the actual practitioners in the field of the farm demo events.

However, the combination of these two functions (internal reflection and external recommendations) set some general requirements for the development of the M&E tools. First, the tools should be simple, logical and practical to use, because local practitioners who will carry out the M&E in the hubs and networks are often not experienced in this. Second, the M&E reporting (from each of the hubs and networks) should be very structured and contain enough detail, because they will be analysed by researchers to extract recommendations on farm demonstrations. These two general requirements imply that the

M&E tools within NEFERTITI should be both scientifically rooted and easy to comprehend and use.

To realise this, a two-step approach for tool development has been applied. First, a 'general M&E approach' was developed that is rooted in scientific insights and a detailed assessment of the requirements of the NEFERTITI project. Second, based on the general M&E approach, a very practical set of 'M&E guidelines and tools' was developed, reviewed and tested by the monitors in the hubs and networks, and they were trained to use them. The development of the guidelines and tools is seen as a continuous process. At the end of the first demo campaign year, the appropriatness of the M&E approach for this purpose will be assessed and the M&E approach will be adjusted if needed.

The M&E guidelines for the networks and hubs consist of three sections: (i) M&E objectives, role of the monitor and main monitoring tasks, (ii) Step 1: Preparing M&E, (iii) Step 2: Carrying out M&E. The first part explains the importance of M&E as a tool to help hubs to make their activities more effective and what it generally entails. The second part (Step 1) provides guidance on how a hub can plan its M&E activities, also specifying who will take which responsibilities in M&E. One of the tools is a simple table that hub coaches fill in. It lists all the hub activities, as well as the objectives and expected outcomes for these activities. By later comparing the actual outcome with the expected outcomes, this becomes a tool to stimulate reflexivity in the hub as a basis for learning and improving. The third part (Step 2) describes how to actually perform M&E, including four specific tools: the hub M&E journal, a checklist, a participant survey and a method for team reflection.

The *Hub M&E Journal* is an Excel file containing separate tabs for monitoring and evaluation of the hub meetings, demo events, cross visits and the annual hub report. The *Checklist* describes important aspects and guiding questions for M&E of demo-events. The *Exit Poll* is a short survey for participants of a demo event. The *Method for team Reflection,*

guides the hub to organise of a group session to reflect on hub activities, and to draw lessons for improvement.

Ideally, all demo activities organised under the NEFERTITI project are monitored and evaluated thoroughly using these tools. However, this could cause some complaints of the network and hub members because of the effort. As a compromise, the M&E guidelines now require from hubs to select one so-called 'showcase demo' each year, for which they perform a more in-depth monitoring while for the other demos they would do this in a less refined way. Hopefully, they will experience the value and benefits of this in-depth monitoring in such a way that it motivates them to do this in-depth monitoring for other demos as well.

Further, to assist the hubs in the M&E process, a helpdesk is set up to support the hubs in carrying out M&E. This helpdesk takes a pro-active approach by monitoring the M&E process and taking the initiative to interact with the monitors to share experiences and address problems that are encountered. The purpose is to keep track of the development of the hubs, their harvesting of lessons learned and to highlight crosscutting themes and opportunities for learning between the hubs and networks. Regularly, the project team will contact the NEFERTITI hub teams (individually or in a group skype) to inquire their progress, potential challenges ahead and need for support and exchange.

## 4. Results and discussion: Evaluation of the M&E tools

### *4.1 Scientific rigor of data provided by the M&E tools*

The data delivered by the use of the M&E tools in AgriDemo-F2F and PLAID allowed to identify aspects of the specific demo events that could be improved. Examples are, to increase room for informal exchange, to show real novelties (much of what is demonstrated was not considered really new), to provide more background information, to better specify the

target audience (e.g. interesting for large farms and not for small farms), to give the floor to farmers who are actually using an innovation, to organise the event using a bottom up approach (taking into account farmers‘ needs).

Further, sharing these findings with the organisers of the demo events (e.g., in focus groups), was perceived as very valuable for the organisation of future demo events. It thus seems that monitoring a demo event is worth the effort and can provide considerable value in learning how to improve future demonstration events. The lessons learned by organisers related to gaining insights in the needs and motivation of participants to attend a demo event, the topics participants are interested in, the barriers for implementation of an innovation, planning follow-up activities, and collecting contact information of participants for follow-up activities .

The integration of the M&E data from all case studies thus provided the projects with a lot of information on good and bad practices for farm demonstrations. However, the M&E tools were not able to provide sufficient information to draw conclusions on the following aspects.

First, the M&E tools did not allow to capture the specific objectives of the demo event. Setting explicit objectives for the demo event is important for both the organisation and set-up of the demo event and to allow useful M&E during and after the demo event. A comparison of the actual outcome of the demo event with the objectives, allows to evaluate the successful and less successful aspects as well as to identify different side effects. This kind of evaluation can subsequently be used to improve future demonstration events. PLAID and AgriDemo- F2F case studies showed that only very few organisers explicitly define the demo objectives, but immediately start to list the activities that they will carry out. The crucial question for M&E is to what extent the variety of activities of a demo event actually helps to achieve the demonstration’s objectives. However, using the M&E tools fo AgriDemo-F2F

and PLAID, the questions asked on the objective were not specific enough to clearly make such an assessment.

Based on this finding, PLAID identified four key aspects of the demonstration objectives that should be thought through by the organisers of a demo event: why do we organise this demo event?, what will be demonstrated?, who do we target?, and which outcomes do we aim for?. The 'why'-aspect of a demo objective specifies the motivation or need for the demo. The 'what'- aspect of a demo objective specifies the object that is demonstrated, e.g. farming equipment, farming practice, crop varieties, etc. Based on the 'why' and 'what' aspects discussed above, the demo organisers should specify the who, i.e., which audience they target. The last aspect of the demo objective are the targeted outcomes. These can be divided into short-term goals, i.e., what the demo participants take home, and longer term goals, i.e., how participants implement what they learned during the demo event. These four aspects of the demo event are closely linked and partially define each other. As a starting point, it is useful to address them in the order of why, what, who and targeted outcomes but while doing so, it is likely that organisers will need to jump back and forth a bit to ensure that the four aspects lead to a coherent objective or set of objectives.

Second, the AgriDemo-F2F and PLAID tools did not sufficiently allow to evaluate how the composition and size of the participant group (gender, age, organic-non-organic, …) deviates from the targeted farmer group. This is also related to the lack of specific objectives, without clear information on the targeted farmer group. Often the viewpoint of organisers is "the more participants, the better", without taking into account the impact on the knowledge exchange processes. Indeed, analysis of the case studies in AgriDemo-F2F, revealed that participants from smaller demo events (<25 participants) were more likely to rate the event as 'effective', compared to their peers who attended medium and large events (≥ 25

participants). The idea is that when farmers are involved and can participate in a smaller group this can support interactive knowledge exchange and experiential learning.

Third, the AgriDemo-F2F and PLAID tools did not sufficiently allow to evaluate the impact of demonstrations in terms of their contribution to the sustainable development of agriculture. This is related to the difficulty to capture the long-term effects of demonstrations and the fact that demo events should be seen in the whole range of activities organised to stimulate sustainable development in agriculture. Furthermore, demo events are often only focussed on a single topic/aspect that might have influence on other aspects of the farming or agricultural system, which makes it difficult to show an innovation's contribution to sustainable development. Additionally, demo events not always clearly state their (targeted) contributions to sustainable development and thus participants might not be aware that it is contributing to sustainable development. This might influence the results in the post-surveys and participants surveys after the demo event. It was, for example, not asked which personal definition organisers or participants attributed to sustainable agriculture. Yet, what is clear, is that demonstrations are an important tool to help farmers in their decision-making on the adoption of new farming practices and other innovations.

Fourth, the tools did not allow sufficiently enough to evaluate how different knowledge exchange activities contributed to learning outcomes in terms of double and triple loop learning and the diffusion of knowledge after the demo. Participant surveys and telephone surveys have been used to try and grasp these learning outcomes, besides single loop learning and adoption. However, answers on these surveys were often not enough in-depth to investigate the effect of knowledge exchange activities on learning outcomes.

### *4.2 Ease to use the M&E tools*

Feedback from the project partners on the use of the M&E tools provided some insights in their practicality and what should be taken into account when developing M&E tools for demo events.

A first point of concern is the difficulty to perform participant surveys after the demo event. These difficulties are related to the fact that people often run away quickly after a demo event and/or that it requires quite some people to take live participant surveys after the demo event to have a meaningful number of completed forms. A way to deal with this, is to include a time slot dedicated to the evaluation of the demo event by the participants at the end of the event. The participant survey can be handed out in paper to all participants and maybe a (short) interactive discussion session can follow this. For example, multiple partners from the AgriDemo-F2F project expressed concerns regarding making participants complete participant survey . When there's no specific time foreseen in the demonstration programme, it was often difficult to get most participants to fill in the post demonstration survey.

A second point of concern might be the misunderstanding of different concepts or questions used in the tools. For example, in the participant survey of AgriDemo-F2F, the participants were asked if they learned something about sustainable agriculture. However, they were not asked to provide their personal definition of sustainable agriculture, so there might have been different interpretations od these concepts. A way to deal with this, is the participatory development of the M&E tools together with the users, or to instruct the end users and let them test the tools, as was done in the NEFERTITI project.

A third point of concern is the lack of time to perform surveys in the period (a week, a month, several months) after the demo event to investigate the impact of a demo event. This is probably specifically relevant for organisers of demo events, for which the assessment of the

demo impact is not a specific assignment in their work description. On the contrary, for social researchers investigating the impact of demo event might be easier justified according to their work description.

Although in NEFERTITI the tools have not been used yet in practice, the tension between scientific rigor and practical use specifically came to the front in this project. The developers of the M&E guide felt that the proposed M&E guide was rather simplistic, whereas the practitioners who had to use it still found it rather complicated. This urged the developers to simplify it further, because they realised that a too complicated tool would not provide them with sensible data from the 45 hubs. As there will be three 'campaign years' in the project, it might be an opportunity to "complicate" the tools in later years, once the hubs have gained experience with the "easy" version of the M&E scheme. Maybe, the hubs will even come up with additional questions to better understand the processes going on during demo event, thus advancing the M&E tools themselves.

## 5. Conclusion

The evaluation of the M&E tools of AgriDemo-F2F and PLAID show that they allowed to capture very useful information on demo events, but also highlight some gaps. For example, the tools did not sufficiently allow to evaluate the specific objectives of the demo event, how the composition and size of the participant group (gender, age, organic-non-organic, …) deviates from the targeted farmer group, the impact of demonstrations in terms of their contribution to the sustainable development of agriculture, how different knowledge exchange activities contributed to learning outcomes in terms of double and triple loop learning and the diffusion of knowledge after the demo. Although it might sound very obvious, one of the most important lessons learned is the importance for organisers to set clear objectives at the start of the organisation of a demo event, to be able to evaluate it in a

meaningful way. This allows the organisers to compare what they intended to achieve with what they have actually achieved. This will add reflexivity to the process of organising a demo event and help the organisers to learn in a more structured way on how to best organise demo events.

Further, the tension between scientific rigour and use proved a challenge in all three projects, even though it referred to different types of M&E (both for internal and external audience). This should be kept in mind by developers of M&E tools. This issue can be addressed by using participatory development of M&E tools, or to instruct the end users and let them test the tools.

## 6. Acknowledgements

This research has been funded by the H2020 project NEFERTITI granted by the European s Horizon 2020 research and innovation programme under grant agreement No. 772705.

## References

Argyris, C. and Schön,D. (1996) Organizational learning II: Theory, Method, and Practice. Massachusetts: Addison-Wesley.

Burton, R., Elzen, B, Tisenkopf, T., Ādamsone-Fiskoviča, A., Grivins, M. (2017) Deliverable D2.1 PLAID: a practice-based conceptual framework and typology. PLAID. Retrieved from (on 20/05/2019): https://www.plaid-h2020.eu/sites/www.plaid-h2020.eu/files/Formatted%20Conceptual%20framework.pdf

Cooreman, H. Debruyne, L, Ingram, J., Koutsouris, A., Marchand, F. (2018) D5.1 State-of-the-art report on determining effectiveness of learning approaches. AgriDemo-F2F. Retrieved from (on 20/05/2019): https://agridemo-h2020.eu/report-on-determining-effectiveness-of-learning-approaches/.

Conley-Tyler, M. (2005). A fundamental choice: Internal or external evaluation?. Evaluation Journal of Australasia, 4(1-2), 3-11.

Elzen, B., Dockes, Anne-Charlotte, Giagnocavo, C, Marchand, F., Schoorlemmer, H., Triste, L. (2019). D5.1 Monitoring and Evaluation Approach for NEFERTITI Hubs and Networks. NEFERTITI.

Hummelbrunner, R. (2015). Learning, systems concepts and values in evaluation: proposal for an exploratory framework to improve coherence. IDS Bulletin, 46(1), 17–29.

Ingram, J., Chiswell, H., Mills, J., Debruyne, L., Cooreman, H., Koutsouris, A., Pappa, E., Marchand, F. (2018) Identifying functional characteristics that enable learning in demonstrations: a discussion paper. Proceedings of the 13th European IFSA Symposium held from 1-5 July 2018 in Chania, Crete - Greece. Retrieved from (on 20/05/2019): http://ifsa.boku.ac.at/cms/fileadmin/Proceeding2018/1_Ingram.pdf

Koutsouris, A., Papa, E., Chiswell, H., Cooreman, H., Debruyne, L., Ingram, J., Marchand, F. (2017) D2.1 The analytical framework. Demonstration farms as multi-purpose structures, provinding multi-functional processes to enhance peer-to-peer learning in the context of innovation for sustainable agriculture. AgriDemo-F2F. Retrieved from (on 20/05/2019): https://agridemo-h2020.eu/docs/D2.1_Rapport_AGRIDEMO_analytical%20framework.pdf

Romme, A., & Van Witteloostuijn, A. (1999). Circular organizing and triple loop learning. Journal of Organizational Change Management, 12(5), 439–454.

Van Mierlo, B.C., Regeer, B., Van Amstel, M., Arkesteijn, M.C.M., Beekman, V., Bunders, J.F.G., De Cock Buning, T., Elzen, B., Hoes, A.C. and Leeuwis, C. (2010). Reflexive Monitoring in Action. A guide for monitoring system innovation projects. Oisterwijk, Uitgeverij Boxpress, Netherlands.

World Bank (2000). Monitoring and Evaluation for AKIS Projects. The World Bank Rural Development Family Agricultural Knowledge & Information Systems (AKIS). Work in progress for public discussion. Prepared by Gary Alex and Derek Byerlee with inputs from the AKIS Thematic Team Framework and Options.
Retrieved from (on 20/05/2019): http://siteresources.worldbank.org/INTARD/825826-1111400636162/20431913/monitoringandeval.pdf