Routledge
Taylor & Francis Group

The Journal of Agricultural Education and Extension

Competence for Rural Innovation and Transformation

ISSN: 1389-224X (Print) 1750-8622 (Online) Journal homepage: http://www.tandfonline.com/loi/raee20

Evaluation of farmer participatory extension programmes

Jorie Knook, Vera Eory, Matthew Brander & Dominic Moran

To cite this article: Jorie Knook, Vera Eory, Matthew Brander & Dominic Moran (2018) Evaluation of farmer participatory extension programmes, The Journal of Agricultural Education and Extension, 24:4, 309-325, DOI: 10.1080/1389224X.2018.1466717

To link to this article: https://doi.org/10.1080/1389224X.2018.1466717

JOURNAL OF AGRICULTURAL EDUCATION AND EXTENSION
2018, VOL. 24, NO. 4, 309–325
https://doi.org/10.1080/1389224X.2018.1466717

Evaluation of farmer participatory extension programmes

Jorie Knook [a], Vera Eory[b], Matthew Brander[a] and Dominic Moran[c]

[a]University of Edinburgh Business School, Edinburgh, UK; [b]Land Economy, Environment and Society, Scotland's Rural College, Edinburgh, UK; [c]The Royal (Dick) School of Veterinary Studies, The Roslin Institute, University of Edinburgh, Edinburgh, UK

ABSTRACT

Purpose: Participatory extension programmes are widely used to promote change in the agricultural sector, and an important question is how best to measure the effectiveness of such programmes after implementation. This study seeks to understand the current state of practice through a review of *ex post* evaluations of participatory extension programmes.

Design/methodology/approach: A systematic literature review of the peer-reviewed literature was undertaken to analyse the evaluations based on: (i) year of publication; (ii) location of the study; (iii) programme delivery; (iv) evaluation methods; (v) outcome variables; and (vi) inclusion of evaluation in initial programme design.

Findings: The review finds that almost all studies use an experimental or quasi-experimental research design (i.e. using a control group or counterfactual), but some studies do not account for endogeneity or selection bias. Furthermore, only a small number of the evaluations were planned as part of the original programme design, which causes difficulties in obtaining robust counterfactuals. The review also finds that relatively few evaluations, approximately 20%, measure the programme impact on environmental outcomes and only 15% of the evaluations have been undertaken for programmes in developed countries.

Practical implication: Limitations with current evaluation practice are identified, and recommendations are provided for improving practice, including better treatment of endogeneity, and the complementary use of qualitative data.

Theoretical implication: The review provides a contribution to the debate about the use of quantitative versus qualitative evaluation methods, by addressing the use of both quantitative and qualitative evaluation methods in a complementary way.

Originality/value: Despite their widespread implementation, this is the first systematic literature review for published evaluations of participatory extension programmes in the agricultural sector.

ARTICLE HISTORY
Received 11 September 2017
Accepted 16 April 2018

KEYWORDS
Impact evaluation; voluntary uptake; extension programme; agriculture; ex post evaluation; discussion groups

Introduction

Extension activities are widely applied to stimulate change in the agricultural sector (Black 2000). For many years, extension was based on the linear top-down transfer of technology,

CONTACT Jorie Knook jorie.knook@sruc.ac.uk University of Edinburgh Business School, 29 Buccleuch Place, Edinburgh EH8 9JS, UK; Land Economy, Environment and Society, Scotland's Rural College, West Mains Road, Edinburgh EH9 3JG, UK

Supplemental data for this article can be accessed at https://doi.org/10.1080/1389224X.2018.1466717

in which technology was developed and validated by researchers, communicated by extension agents and adopted by farmers (Black 2000). However, since the 1980s, this approach has been subjected to various critiques, such as failing to account for the context and complexity of the agricultural sector (Pretty and Chambers 2003), which in turn decreases the adoption of technology. Therefore, an alternative extension approach has been developed in which farmers play a more central or 'participatory' role in the acquisition of knowledge and change of practice (Scoones and Thompson 2009; Cristóvão, Koutsouris, and Kügler 2012) In these 'participatory' extension programmes (PEPs), researchers and extension agents fulfil a facilitating role, while farmers actively set the agenda and engage with their peers (Black 2000).

Given the public investment in PEPs, and the increasing requirement for accountability by policy-makers and funding bodies, it is important that these programmes are reliably evaluated (Faure, Desjeux, and Gasselin 2012; Klerkx, Landini, and Santoyo-Cortés 2016). However, evaluating PEPs might present new challenges, as PEPs may require a different evaluation approach than the evaluation of top-down extension programmes. The evaluation of top-down programmes is mostly focused on programme outcomes, but it is questioned whether this approach sufficiently addresses the main aim of PEPs, which is to include farmers in agenda setting and collective learning (Murray 2000). Therefore, it is interesting to identify the current state of evaluation practice and identify recommendations for improvement.

To identify the current state of practice, this study provides a systematic review of peer-reviewed evaluations for PEPs. Although previous reviews overlap this topic, e.g. reviews focused on Farmer Field Schools (FFS) (Van den Berg 2004; Davis 2006; Van Den Berg and Jiggins 2007), or evaluations of all types of extension programmes (IEG, 2011), this review is the first – to the authors' knowledge – to focus specifically on PEP evaluations. The contribution lies in the identification of best practice for *ex post* evaluation methods, derived from the peer-reviewed literature for agricultural PEPs. Ultimately, the review identifies limitations within the currently applied evaluation methods and provides recommendations for future evaluations. The remainder of this paper proceeds as follows: Section two describes the systematic review method; Section three presents and discusses the findings from the review, and the paper concludes with a number of recommendations.

Methods

Definition and scope

PEPs are programmes in which farmers interact with peers and experts, where experts fulfil a facilitating role and farmers actively participate in goal and agenda setting. Programme meetings take place over a period of time and create knowledge by participatory learning methods, such as group or one-on-one meetings, training sessions and (experimental) demonstrations (Black 2000). The intended outcomes from PEPs include changing farm practices, enhancing social learning, increasing resilience to challenges and uncertainties, and sharpening farmers' management skills and decision-making abilities (Cristóvão, Koutsouris, and Kügler 2012).

This systematic review focuses on peer-reviewed studies that measure the effect of PEPs *ex post*, that is, after the implementation. Although there are evaluation studies reported in the grey literature, i.e. the sources of literature outside of traditional academic publications,

such as theses, reports from governments or organisations and working papers, these are not included in the review, because the main aim is to identify best practice for evaluation from a scientific perspective.

Sources of information

An initial inventory of peer-reviewed publications on the evaluation of PEPs was conducted including studies until August 2017, by using the electronic databases of ISI Web of Knowledge (www.isiknowledge.com) and Google Scholar (scholar.google.com). We used the following key words, either alone or in combination: 'agriculture', 'evaluation', 'participatory extension programme', 'voluntary advisory programme', 'policy', intervention', 'impact' and 'assessment'. This initial search resulted in 45 publications. To provide a more exhaustive list of evaluations, we conducted a second search in the previously mentioned electronic databases including studies until December 2017, by using additional search terms often associated with PEPs (Black 2000; Braun 2006; Cristóvão, Koutsouris, and Kügler 2012): 'participatory learning and action', 'participatory technology development', 'facilitation of local processes', 'local development', 'agroecological extension', 'farmer field schools', 'farmer first', 'farmer-led extension', 'farmer networks', 'study circles', 'farming systems research and extension', 'farmer study groups', 'rural resource center', 'farmer to farmer training', 'master farmer training', 'local learning groups', 'participatory advisory programme and discussion groups'. This yielded an additional 26 studies, bringing the total to 71. This expanded set of key words may still not provide an exhaustive list, but address the main studies in this field.

To focus on recently conducted studies which evaluate the effect of PEPs, we only included publications which: (i) focus on the effect after implementation of the PEPs, also referred to as 'ex post' evaluations; (ii) present the effect of the PEP using empirical findings; (iii) focus on PEPs within the agricultural sector; and (iv) have been published in or after the year 2000. The search resulted in 71 studies, which we further analysed based on six aspects. Firstly, the year of publication was used to identify a trend in the number of evaluations over time. Secondly, the location of the programme was identified to analyse the spatial distribution of the PEPs evaluated. Thirdly, the type of delivery was analysed, to find similarities in programme design. There are, for instance, a number of studies that apply the farmer field school approach, which is a uniform programme design applied in many developing countries. Fourthly, evaluation methods were categorised to identify the types of method and their frequency of use. Fifthly, the outcome variables used in the evaluation studies were identified, e.g. economic outcome variables, environmental outcome variables, etc. Finally, we identified whether the evaluation studies were built into the initial programme design. In the absence of any explicit mention of an evaluation in initial programme design, the presence of a baseline survey was taken as an indicator of evaluation planning.

Findings and discussion

General findings

A total of 71 published evaluation studies for PEPs were identified, from 42 different journals. A general finding in relation to terminology is that 'impact assessment', 'effect' or

'effectiveness' are used interchangeably to indicate some form of quantitative evaluation. Sixty-eight out of 71 studies found a positive difference after the intervention, the exceptions being Feder, Murgai, and Quizon (2004a, 2004b) and Rejesus et al. (2012). It is possible that there are additional, unpublished, evaluations that do not show a positive effect on the participants, but these have not been published in the scientific literature due to publication bias, i.e. editors, funders, reviewers and researchers have a preference for studies that show a statistically significant effect (Duflo et al. 2007).

The Appendix provides an overview of the collected studies that will be discussed in terms of the six aspects previously mentioned (see supplementary data section).

Year, location and type of delivery

Only eight publications were found that conducted an evaluation between 2000 and 2006, indicating an increase in evaluations over the last decade (Figure 1).

An overview of the studies categorised per continent is depicted in Figure 2. Analysis of the location of the studies shows that the majority of the studies have been conducted in countries in Africa and Asia. Further analysis shows that 62 of the 71 studies were conducted in developing countries (as classified by the United Nations (2018)), which can be explained by the fact that the majority of PEPs are implemented in the developing world (Anderson and Feder 2004). A popular type of PEP in developing countries are FFS, which use education to strengthen farmers' capacity to what can be considered as 'best practices'. Typically, FFSs consist of 20–25 farmers who, under the guidance of a trained facilitator, meet on a weekly basis for a predefined period to discuss environmental topics, such as soil fertility and pest management, but also other topics, such as the development of marketing skills (FAO 2017). Forty-eight out of the 62 developing country studies focused on the evaluation of these FFSs. The other 14 evaluations were applied to a wide range of PEPs. For instance, Pamuk, Bulte, and Adekunle (2014) and Pamuk

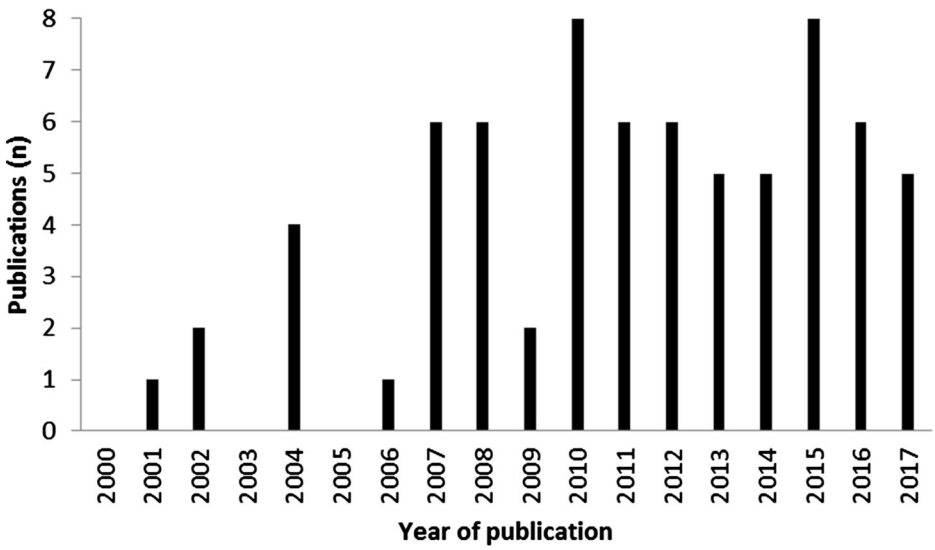

Figure 1. Evaluations of the participatory extension programmes by the year of publication.

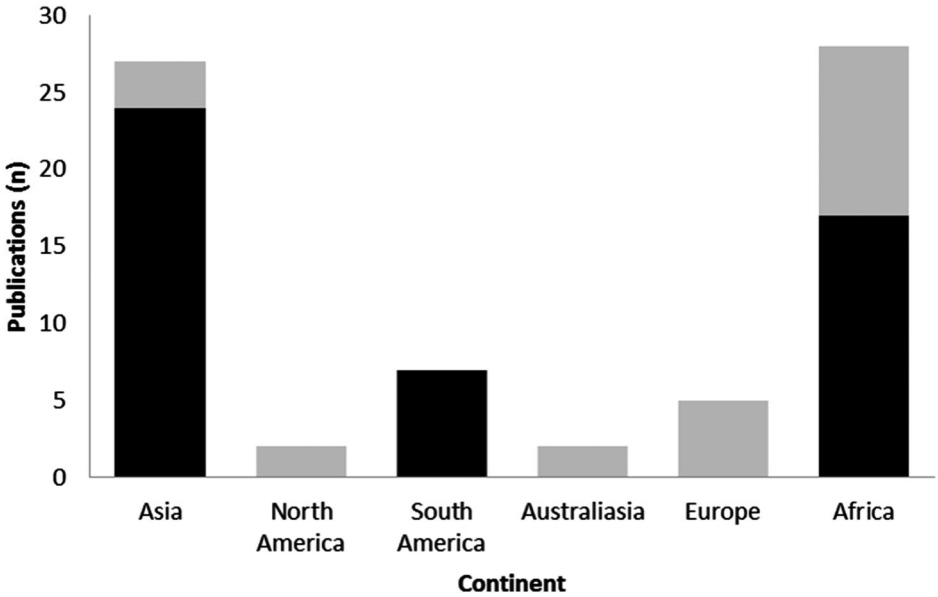

Figure 2. Categorisation by continent. In black, the share of studies evaluating an FFS is indicated and in grey the share of studies in which a different PEP is evaluated.

et al. (2015) studied 'participatory innovation platforms', in which local stakeholders meet and collectively identify problems and solutions; Kraaijvanger, Veldkamp, and Almekinders (2016) studied 'participatory experimentation programmes', in which farmer groups participate in learning cycles consisting of experience, design, experimentation and reflection; and Schreinemachers et al. (2016) looked at 'farmer training', which refers to participatory training of farmers during a two-day workshop, followed by regular farm visits by experts and peers.

FFSs tend not to be applied in developed countries, although the PEPs used in developed countries show similarities in programme delivery, such as the inclusion of education and group activities. A total of nine evaluation studies were conducted in developed countries: Bruges and Smith examined the effect of farmer participatory research groups regarding the adoption of sustainable practices in New Zealand; Hill, Bradley, and Williams (2017) focused on the programme 'Farming Connect' in Wales, which aims to promote knowledge transfer, advice and training for farms and forestry holdings; Hennessy and Heanue (2012), Läpple and Hennessy (2015) and Läpple, Hennessy, and Newman (2013) looked at the effectiveness of discussion groups in the dairy sector in Ireland; Prager and Creaney (2017) evaluated how discussion groups in Ireland and monitor farms in Scotland work and which factors influence their success; King, Gaffiiey, and Gunton (2008) considered whether participatory action learning, a participatory extension approach for farmer groups, increases learning compared to more traditional extension approaches in Australia; Roche et al. (2015) evaluated a participatory-based experimental learning programme in which experts work with focus farms to change dairy producer behaviour to control Johne's disease; and Tamini (2011) evaluated the uptake of best management practices after participation in farmer advisory clubs in Canada.

Evaluation methods

The different evaluation methods found within the published studies of PEPs are analysed using the categories presented in Figure 3. This categorisation first divides the evaluation methods according to whether they are quantitative, qualitative or mixed methods, with the quantitative methods further subdivided according to their treatment of endogeneity.

Quantitative methods

Sixty-four of the 71 identified studies evaluated the PEP by (mainly) applying a quantitative method. These studies were further categorised (according to the categorisation in

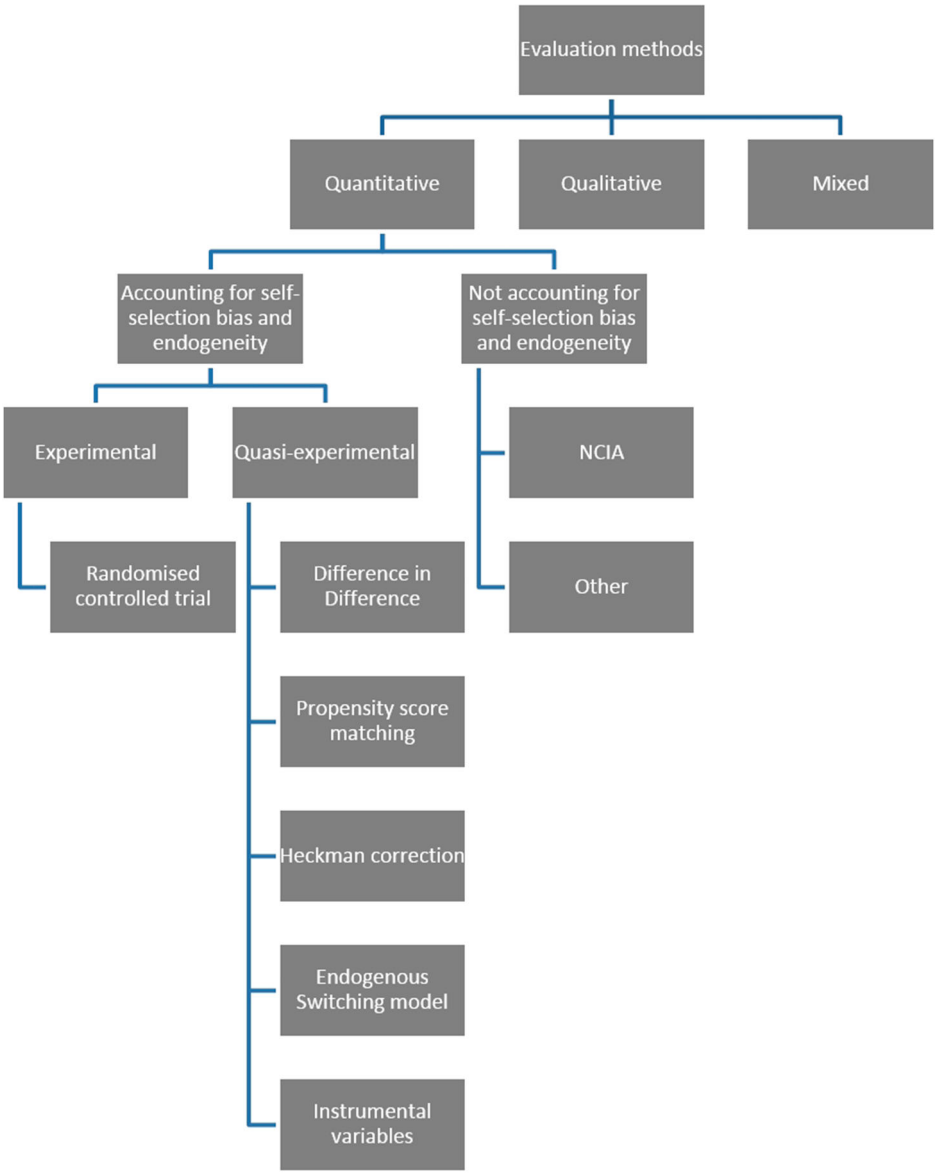

Figure 3. A categorisation of the evaluation methods used to evaluate PEPs.

Figure 3) depending on their use of (1) a method accounting for self-selection bias and endogeneity, including experimental, quasi-experimental and other approaches, or (2) a method not accounting for self-selection bias and endogeneity. An overview of the quantitative methods applied is provided in Figure 4.

Accounting for self-selection bias and endogeneity. Thirty-two studies used an experimental or quasi-experimental research design to conduct a quantitative evaluation, hereafter also referred to as 'impact evaluation', which is a widely used term in literature and addresses the effectiveness of a PEP by comparing it to the situation in the absence of the PEP (Gertler et al. 2016) and only one study used a different approach.

Experimental research design. Experimental research designs allocate participants randomly to a treatment or control group to prevent selection bias, which arises when participants and non-participants differ in characteristics that are related to participation in the programme and to the outcome (Duflo et al. 2007). Random allocation of participants is assumed to correct for any imbalance in characteristics, and the groups only differ in the presence or absence of the treatment. The effect or impact of the treatment can therefore be estimated as the difference between the control group and treatment group (Duflo and Kremer 2003; Duflo et al. 2007).

Within the evaluation methods identified, the randomised controlled trial (RCT) is the only purely experimental method, and was only used in one of the studies: Guo et al. (2015), which randomly selected treatment villages for participation in the programme in question. Although RCT optimally accounts for selection bias, the application is complicated and this is most likely the reason for the limited use of the method. The methodological challenges include: the need to plan the evaluation during the initial stages of PEP implementation; overcoming ethical restrictions which may arise when non-participants

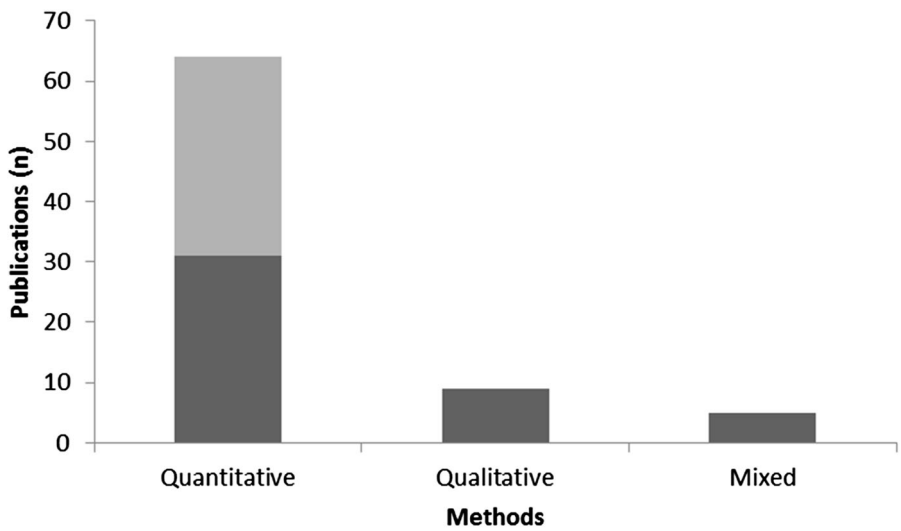

Figure 4. Categorisation of the 71 evaluation studies by quantitative, qualitative and mixed methods. The number of studies depicted is higher than the number of evaluation studies, because some studies applied multiple methods and therefore could be classified in more than one category.

are disadvantaged because of exclusion from the treatment group; accounting for spill-over effects that can occur when participants exchange information with non-participants (Duflo et al. 2007). As an illustration of the difficulties with this method, Guo et al. (2015) found that the treatment villages in their study did not show a high level of comparability with the control villages, and to avoid imbalance between the treatment and control groups they applied matching techniques (see below) to account for the differences.

Quasi-experimental design. In contrast to experimental research design, quasi-experimental approaches allow non-randomised selection or self-selection of the treatment group, e.g. prospective participants can volunteer for the treatment group. Any endogeneity and self-selection bias can then be accounted for using one of several techniques: (i) difference-in-difference (DiD); (ii) propensity score matching (PSM); (iii) the Heckman correction (HC); or (iv) endogenous switching model (ESM).

The review found that difference-in-difference (DiD) was the most commonly used method, with 14 studies applying this approach. This method compares the before and after changes of a treatment group with the before and after changes of a control group, thereby controlling for differences in unobservable characteristics. The approach quantifies the difference between the groups in two steps: (i) it quantifies the average difference in outcome for the treatment and control group; and (ii) it calculates the average difference between the average changes for the treatment and control group (Bertrand, Duflo, and Sendhil 2004). An example of a study which applied DiD is Mancini et al. (2008), which measured the before and after effect of FFSs in India. An RCT was not possible because although the programme targeted specific villages, farmers' participation was on a voluntary basis. Therefore, DiD was used to account for seasonal or systematic effects other than the training effect, such as socio-economic factors, which might have favoured the participation of more progressive, wealthier and educated farmers. Togbé et al. (2014) also applied DiD to account for the non-randomised selection of farmers participating in the FFS. In Rejesus et al. (2012), the DiD approach was used to account for differences in village characteristics, because villages were selected to participate in the FFS based on access to the village, and the presence of active farmer groups.

Five studies applied PSM, which addresses endogeneity and self-selection bias by creating a propensity score for each participant based on socio-economic and other relevant characteristics, and then matches scores between members of the treatment and control groups to create groups that are as closely matched as possible (Stuart 2010). This method is useful when baseline data and longitudinal observations are lacking (and therefore the DiD method cannot be applied), but abundant cross-sectional data for participants are available. However, a crucial assumption and potential weakness of the PSM method is that there are no unobservable characteristics that may cause a difference in results between the treatment and control group. Examples of the application of PSM can be found in Godtland et al. (2004), which applies PSM to evaluate the effect of FFSs (on knowledge of integrated pest management) with cross-sectional data, and in Schreinemachers et al. (2016), which uses PSM to quantify the effect of farmer training.

Five studies applied both PSM and DiD (Rodriguez, Rejesus, and Aragon 2007; Davis et al. 2012; e.g. Todo and Takahashi 2013) to account for systematic differences between participant and non-participant outcomes, which may continue to exist even after matching observable characteristics (Heckman, Ichimura, and Todd 1997). These differences may, for example, occur due to programme selection based on unmeasured characteristics,

or because the treatment and control groups reside in different regions. Although combining PSM and DiD has the advantage of controlling for both observable and unobservable characteristics, it requires an extensive dataset.

One study applied the HC to account for endogenous effects (Rejesus et al. 2009); one applied an ESM (Läpple, Hennessy, and Newman 2013); and five studies applied instrumental variables (IV) (Tamini 2011; e.g. Wafula et al. 2016). All three methods require an instrument or exogenous variable, i.e. a variable that is not included in the equation of interest and via which the counterfactual can be established. However, in many situations, there is no obvious or measurable exogenous variable, which complicates the application of this method.

Not accounting for self-selection bias and endogeneity. In addition to the studies above, 32 studies calculated the effect of PEPs by conducting an impact evaluation, but did not account for endogeneity or self-selection bias either explicitly or correctly. We have categorised these studies into 'No correct impact evaluation (NCIA)' and 'Other methods'.

NCIA. Although the NCIA studies show positive results, the reliability of the findings is questionable. For instance, Tin et al. (2010) conducted a baseline and an *ex post* survey to measure change over time. However, they did not include a control group to properly account for unobservable characteristics. Furthermore, Läpple, Hennessy, and Newman (2013) provide a critique of Hennessy and Heanue's (2012) PEP evaluation, noting that a major limitation is the neglect of self-selection bias and endogeneity. This causes the under- or over-estimation of programme effects, weakening the policy relevance of this work (Läpple, Hennessy, and Newman 2013). Other studies such as Bentley et al. (2007) do not mention self-selection or endogeneity at all.

Other quantitative methods. One study applied a different quantitative methodology to assess the effect of a PEP. Bourne et al. (2017) assessed the performance of participatory advisory programmes by using social network analysis (SNA), which is the process of investigating social structures through the use of networks, as a tool to examine farmer networks. The study used SNA to analyse whether farmer networks change due to participation in an advisory programme. More specifically, it assessed the contribution towards joint decision-making, cooperation in the implementation of innovations and management of collective activities. Although the study presents a new framework to assess the PEP on these outcomes, it does not address the change in information over time, or compares the findings with a control group. Hence, we argue that in the application of this method a valid counterfactual is lacking, which undermines the findings of this study.

The absence of a reliable method to account for endogeneity in 32 published studies is a striking finding from this review and should be noted by journal editors and reviewers, as well as researchers undertaking evaluation studies.

Qualitative evaluation methods

We identified seven qualitative studies focused on PEP evaluation. This number appears low and we suspect that some qualitative studies are potentially disguised behind atypical titles and therefore are more difficult to detect by keyword search. King, Gaffiiey, and Gunton (2008) is an example of a qualitative study, which applied the convergent interviewing approach, seeking to reveal participants' reported experience of effective learning.

The study observed a positive effect and argues 'soft' evaluation techniques such as convergent interviewing are a successful tool when faced with 'difficult to measure' PEP effects. Furthermore, Prager and Creaney (2017) combine qualitative interviews, participant observation and document analysis to draw conclusions about how participatory groups work and what influences their success.

Mixed evaluation methods

Five studies mentioned the application of both a quantitative and qualitative method. All these studies applied a qualitative method in addition to an impact assessment to measure the effect of a PEP and are thus partially already discussed in the previous sections. For instance, David and Asamoah (2011) conducted focus groups prior to the impact assessment to increase the understanding of farmers' perceptions of the impact of the FFS. They specifically asked for the impact on knowledge, decision-making skills, experimentation and knowledge diffusion, which helped in identifying suitable indicators for the impact assessment. Dolly (2009) aimed to assess 14 FFSs in Trinidad & Tobago in relation to six key extension challenges. Besides conducting interviews with individuals from the treatment and control groups, they also attended FFS meetings and included the observations during the meetings in the interpretation of the interview findings.

Not all studies explain the use of additional qualitative methods. For instance, the methodology section of Davis et al. (2012) refers to qualitative data obtained from document analysis and semi-structured interviews, but omits a transparent description of how the qualitative data are included in the study, and the results section only presents an analysis of the quantitative data. Similarly, Lund et al. (2013) undertook interviews to gain insight into the views of participants on the programme and how knowledge was acquired through programme participation. However, again only results from the quantitative data analysis are presented.

Hill, Bradley, and Williams (2017) also applied both quantitative and qualitative methods. The study included qualitative interviews with farmers to gather data on the farmers' own perceptions of the effectiveness of the PEP in question. The findings were then compared to the results from a quasi-experimental impact assessment. This showed that the qualitative approach finds a more positive outcome than the quantitative approach, which may be explained by interviewer bias and overly positive reporting in the qualitative interviews. Arguably, the use of a qualitative method for impact assessment and the subsequent comparison with a quantitative method is not a fair approach as the appropriate use of qualitative methods should be to provide a more in-depth and nuanced understanding of participant motivations and perceptions (rather than being an alternative to quantitative impact evaluation). Kraaijvanger, Veldkamp, and Almekinders (2016) used quantitative and qualitative methods to complement each other. To gain insight into which changes farmers made and whether the programme was responsible for these changes, data were collected via interviews and observations, which in turn provided detailed insight into the functioning of the programme. Overall, several studies argue that in the evaluation of participatory programmes, qualitative and quantitative methods should be used to complement each other (Murray 2000; Munro 2014).

Outcome variables

As mentioned in the 'Methods' section, PEPs aim to change farm practices, enhance social learning, increase resilience to challenges and uncertainties and sharpen farmers' management skills and decision-making abilities. We firstly found that although there is no reason to assume one aim is more important than another, the majority of the studies, with exception of Duveskog, Friis-Hansen, and Taylor (2011) and Jones, Glenna, and Weltzien (2014), include outcome variables related to the first aim: the change in farm practices. Across the 69 studies that included outcome variables related to practice change, 23 different evaluation outcome variables were identified, as shown in Appendix (see supplementary data section). The most common variable was 'knowledge acquisition', followed by 'financial performance'. In a sense, the 'knowledge acquisition' and also the 'knowledge diffusion' variables are of a different order to the other outcome variables, as they may subsume or include any of the other categories, i.e. the knowledge acquired may relate to financial management, productivity, food security, etc. Furthermore, in some evaluations, e.g. Tin et al. (2010), only knowledge acquisition is used as an indicator, because in this study, it is assumed that increased knowledge translates into a change of farming practice. Although David and Asamoah (2011) also use knowledge acquisition as a single indicator, they recognise that practice change does not only depend on knowledge, but other factors as well, such as economic conditions. This point, i.e. that knowledge acquisition does not entail impact, is widely recognised within the literature on agricultural innovation (Rogers 2003; Meijer et al. 2015). In order to address this issue, the majority of studies using knowledge acquisition as an indicator combine it with indicators measuring the actual change in practice (Godtland et al. 2004; Erbaugh et al. 2010; e.g. Mutandwa and Mpangwa 2004).

A second notable finding is that few evaluation studies focused on environmental outcomes, with only 1 considering ecological footprints, 10 considering pesticide use and 1 focusing on practice change in relation to climate change. Although this is likely to largely reflect the focus of the PEPs themselves, it nevertheless indicates that there is relatively limited research experience in evaluating the effectiveness of PEPs on environmental outcome indicators.

Thirdly, it should be noted that although most PEPs evaluated are FFSs, and FFSs have a largely uniform programme design, there is no standard set of indicators applied to their evaluation. Simpson and Owens (2002) address this issue by highlighting six key issues around FFSs in Africa: relevance and response to local concerns, knowledge acquisition, knowledge diffusion, local institutionalisation and organisational development, impact on relationships and FFS integration into existing programmes. They argue that in addition to outcome variables extra attention should be paid to these six aspects to evaluate the effectiveness of FFS programmes. Only one evaluation assessed an FFS on all these six aspects (Dolly 2009), but in 22 FFS studies knowledge diffusion and acquisition are used as outcome variables, which indicates the partial use of the six indicators proposed by Simpson and Owens (2002).

Inclusion of evaluation in initial programme design

The final aspect of the systematic review identified whether the evaluation was built into the initial PEP design, i.e. whether data collection and the evaluation method were planned

prior to programme implementation. The reason for including this aspect in the review is that such planning is a key determinant of the type and robustness of the *ex post* evaluation that can be subsequently undertaken (Baslé 2006).

The studies that applied either an RCT or DiD method were usually planned as part of the PEP design, because both methods require data before and during the programme. However, although Larsen and Lilleør (2014) applied the DiD method, they mention the absence of a detailed evaluation plan at the beginning of their data collection. They evaluate a programme that was phased-in at different villages, and so although only cross-sectional data were available for the first phase, it was possible to gather baseline data for the second phase. In addition, although the evaluation commenced after the start of the PEP, the authors sought to avoid *ex post* bias, the cherry-picking of suitable indicators later, by basing the evaluation on previously stated aims. Davis et al. (2012) and Feder, Murgai, and Quizon (2004a) also provide a potentially useful approach for undertaking a DiD method in the absence of complete baseline data. Their baseline survey did not contain all the data required to compute the impact of the PEP, and therefore they used recall data from farmers to fill the data gaps for the situation before the implementation of the programme. Moreover, Jørs et al. (2016) did not have access to a complete longitudinal dataset either, because longitudinal data were only available for FFS and exposed farmers, but not for a control group. Therefore, DiD was only applied to make a comparison between FFS and neighbouring farmers and cross-sectional data were used to assess the programme compared to the control group.

The studies that used a cross-sectional dataset did not have the evaluation built in. For instance, Godtland et al. (2004) and Läpple and Hennessy (2015) explicitly mention the limitation in a choice of evaluation methods due to the lack of baseline data.

Conclusion and recommendations

Given the level of investment and expectation of positive outcomes from PEPs, it is important that these PEPs are properly evaluated. To identify and develop best practice, this study provides a systematic review of published evaluations in this area. Based on the findings from the review, we offer several recommendations for improving evaluation practice.

Firstly, we would like to address the large amount of studies basing the evaluation of PEPs on practice change. As mentioned in the 'Methods' section, PEPs aim to change farm practices, enhance social learning, increase resilience to challenges and uncertainties and sharpen farmers' management skills and decision-making abilities. We find that evaluation studies mainly address the first aim: change in farm practices. We recommend the inclusion of the other aims as well, to provide a more holistic evaluation of the PEP.

Secondly, when conducting a quantitative evaluation, practitioners should select methods that address endogeneity and selection bias, as failure to do so undermines the reliability of the evaluation results due to under- or over-estimation of programme effects. Equally, agencies commissioning evaluations, as well as journal editors and reviewers, should request such methods to be used.

Thirdly, although a number of existing studies used some form of qualitative method alongside a quantitative method, the use of qualitative data was not well integrated or was treated as an alternative to quantitative methods. We recommend that qualitative data

should be used to complement quantitative assessments, in order to provide additional insights into the perceptions and motivations of participants, the barriers they face, and the context in which programmes are implemented (Davies, Nutley, and Smith 2000; Montuschi 2014). It is particularly important to understand the social context of a programme when trying to extract lessons from a specific study.

Fourthly, we recommend that *ex post* evaluation should be considered in the initial design of any PEP, and the policy-maker or commissioning agency should take responsibility for ensuring that this is the case. When a quantitative evaluation is not planned prior to programme implementation, only cross-sectional data will be available for the evaluation, restricting the evaluation to one moment in time. This makes it difficult to account for unobservable characteristics. Hence, planned evaluations (prior to programme implementation) allow the establishment of robust counterfactuals and have a large influence on the quality of the impact assessment (Läpple, Hennessy, and Newman 2013).

Fifthly, we want to make a recommendation regarding the choice of indicators selected for evaluation. Although it is essential that *ex post* evaluation is considered in the initial design of any PEP, this does not mean that all the outcome indicators have to be determined by the evaluating party beforehand. In order to align with the ethos of a participatory approach, where collectively setting goals is one of the main aims, and to ensure that the evaluation findings are relevant to the on-going implementation of the programme, the participants themselves should be involved in the selection of some of the outcome indicators (Murray 2000; Bruges and Smith 2008).

Sixthly, we have observed multiple impact evaluations that only use 'knowledge acquisition' as an indicator to assess the effectiveness of a PEP. Although knowledge is recognised as an important factor in practice change, change is also highly dependent on other factors, such as economic performance. Therefore, to measure the effectiveness of a PEP, the indicator 'knowledge acquisition' should be used in combination with other indicators in order to draw conclusions on the actual change in practice.

A final observation is that relatively few evaluations of PEPs have been conducted within a developed country context, and few measure the impact of programmes on environmental outcome variables. Given the increasing emphasis on the voluntary uptake of environmental measures in the agricultural sector (e.g. The Scottish Government 2017), this suggests a gap in the literature that should be addressed by future evaluation studies.

Acknowledgements

We would like to thank the three anonymous reviewers whose suggestions have helped to improve the manuscript. Furthermore, we would like to thank our colleagues Robyn Dynes and Cecile de Klein for supporting Jorie Knook during her research.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Scotland's Rural College.

Notes on contributors

Jorie Knook is a doctoral student at the University of Edinburgh. Her research is conducted at the University of Edinburgh Business School, Scotland's Rural College (SRUC) and AgResearch in New Zealand. Her research focuses on the evaluation of participatory extension programmes focused on the uptake of environmentally friendly farming practices.

Vera Eory is a researcher at Scotland's Rural College in Edinburgh. Her research focuses on the economic and environmental implications of greenhouse gas reduction practices in agriculture, particularly on the cost-effectiveness of mitigation practices, environmental co-effects and the heterogeneity of GHG abatement at farm type level.

Matthew Brander is a Lecturer in Carbon Accounting at the University of Edinburgh Business School's Centre for Business and Climate Change. His research focuses on the development of greenhouse gas accounting methods for organisations, products, projects, policies and countries.

Dominic Moran is Professor of Agricultural and Resource Economics at The Global Academy of Agriculture and Food Security at the University Edinburgh. His research focuses on the management of external impacts from agriculture and food supply chains.

ORCID

Jorie Knook http://orcid.org/0000-0002-2356-8008

References

Anderson, J. R., and G. Feder. 2004. "Good Intentions Extension : Agricultural and Hard Realities." *The World Bank Research Observer* 19: 41–60. doi:10.1093/wbro/lkhO13.

Baslé, M. 2006. "Strengths and Weaknesses of European Union Policy Evaluation Methods: Ex-post Evaluation of Objective 2, 1994–99." *Regional Studies* 40: 225–235. doi:10.1080/00343400600600595.

Bentley, J. W., O. Barea, S. Priou, H. Equise, and G. Thiele. 2007. "Comparing Farmer Field Schools, Community Workshops, and Radio: Teaching Bolivian Farmers About Bacterial Wilt of Potato." *J. Int. Agric. Ext. Educ.* 14: 85–95.

Bertrand, M., E. Duflo, and M. Sendhil. 2004. "How Much Should We Trust Differences-in-differences Estimates?" *The Quarterly Journal of Economics* 119: 249–275.

Black, A. 2000. "Extension Theory and Practice: A Review." *Australian Journal of Experimental Agriculture* 40: 493–502. doi:10.1071/EA97144.

Bourne, M., A. Gassner, P. Makui, A. Muller, and J. Muriuki. 2017. "A Network Perspective Filling a Gap in Assessment of Agricultural Advisory System Performance." *Journal of Rural Studies* 50: 30–44. doi:10.1016/j.jrurstud.2016.12.008.

Braun, A. 2006. *Inventory of Farmer Study Group Initiatives in Africa*. Wageningen: Endelea.

Bruges, M., and W. Smith. 2008. "Participatory Approaches for Sustainable Agriculture: A Contradiction in Terms?" *Agriculture and Human Values* 25: 13–23. doi:10.1007/s10460-007-9058-0.

Cristóvão, A., A. Koutsouris, and M. Kügler. 2012. "Extension Systems and Change Facilitation for Agricultural and Rural Development." In *Farming Systems Research Into the 21st Century: The New Dynamic*, edited by I. Darnhofer, D. Gibbon, and B. Dedieu, 201–227. Dordrecht: Springer.

David, S., and C. Asamoah. 2011. "The Impact of Farmer Field Schools on Human and Social Capital: A Case Study From Ghana." *The Journal of Agricultural Education and Extension* 17: 239–252. doi:10.1080/1389224X.2011.559076.

Davies, H. T. O., S. M. Nutley, and P. C. Smith. 2000. *What Works?: Evidence-based Policy and Practice in Public Services*. Bristol: Policy Press.

Davis, K. 2006. "Farmer Field Schools: A Boon or Bust for Extension in Africa?" *The Journal of Agricultural Education and Extension* 13: 91–97.

Davis, K., E. Nkonya, E. Kato, D. A. Mekonnen, M. Odendo, R. Miiro, and J. Nkuba. 2012. "Impact of Farmer Field Schools on Agricultural Productivity and Poverty in East Africa." *World Development* 40: 402–413. doi:10.1016/j.worlddev.2011.05.019.

Dolly, D. 2009. "An Assessment of the Implementation and Outcomes of Recent Farmer Field Schools to Improve Vegetable Production in Trinidad and Tabago." *The Journal of Agricultural Education and Extension* 16: 7–19. doi:10.5191/jiaee.2009.16201.

Duflo, E., R. Glennerster, M. Kremer, and A. Latif Jameel. 2007. *Using Randomization in Development Economics Research: A Toolkit (No. 333)*.

Duflo, E., and M. Kremer. 2003. "Use of Randomization in the Evaluation of Development Effectiveness." World Bank Operations Evaluation Department (OED) Conference on Evaluation and Development Effectiveness, 37. doi: 10.1017/CBO9781107415324.004.

Duveskog, D., E. Friis-Hansen, and E. W. Taylor. 2011. "Farmer Field Schools in Rural Kenya: A Transformative Learning Experience." *Journal of Development Studies* 47: 1529–1544. doi:10.1080/00220388.2011.561328.

Erbaugh, J. M., J. Donnermeyer, M. Amujal, and M. Kidoido. 2010. "Assessing the Impact of Farmer Field School Participation on IPM Adoption in Uganda." *The Journal of Agricultural Education and Extension* 17: 5–17. doi:10.5191/jiaee.2010.17301.

FAO. 2017. "Farmer Field School Approach [www Document]". Accessed August 22, 2017. http://www.fao.org/agriculture/ippm/programme/ffs-approach/en/.

Faure, G., Y. Desjeux, and P. Gasselin. 2012. "New Challenges in Agricultural Advisory Services From a Research Perspective: A Literature Review, Synthesis and Research Agenda." *The Journal of Agricultural Education and Extension* 18: 461–492. doi:10.1080/1389224X.2012.707063.

Feder, G., R. Murgai, and J. B. Quizon. 2004a. "Sending Farmers Back to School: The Impact of Farmer Field Schools in Indonesia." *Review of Agricultural Economics* 26: 45–62. doi:10.1111/j.1467-9353.2003.00161.x.

Feder, G., R. Murgai, and J. B. Quizon. 2004b. "The Acquisition and Diffusion of Knowledge: The Case of Pest Management Training in Farmer Field Schools, Indonesia." *Journal of Agricultural Economics* 55: 221–243. doi:10.1111/j.1477-9552.2004.tb00094.x.

Gertler, P. J., S. Martinez, P. Premand, L. B. Rawlings, and C. M. J. Vermeersch. 2016. *Impact Evaluation in Practice*. 2nd ed. Washington, DC: The World Bank.

Godtland, E. M., Sadoulet, E., De Janvry, A., Murgai, R., and Ortiz, O. 2004. "The Impact of Farmer Field Schools on Knowledge and Productivity: A Study of Potato Farmers in the Peruvian Andes." *Economic Development and Cultural Change* 53: 63–92. doi:10.1086/423253.

Guo, M., X. Jia, J. Huang, K. B. Kumar, and N. E. Burger. 2015. "Farmer Field School and Farmer Knowledge Acquisition in Rice Production: Experimental Evaluation in China." *Agriculture, Ecosystems & Environment* 209: 100–107. doi:10.1016/j.agee.2015.02.011.

Heckman, J. J., H. Ichimura, and P. E. Todd. 1997. "Matching Evidence Job As an Econometric Estimator: Evidence From Evaluating a Job Training Programme." *The Review of Economic Studies* 64: 605–654. doi:10.1111/1467-937X.00044.

Hennessy, T., and K. Heanue. 2012. "Quantifying the Effect of Discussion Group Membership on Technology Adoption and Farm Profit on Dairy Farms." *The Journal of Agricultural Education and Extension* 18: 41–54. doi:10.1080/1389224X.2012.638784.

Hill, B., D. Bradley, and E. Williams. 2017. "Evaluation of Knowledge Transfer; Conceptual and Practical Problems of Impact Assessment of Farming Connect in Wales." *Journal of Rural Studies* 49: 41–49. doi:10.1016/j.jrurstud.2016.11.003.

IEG (Independent Evaluation Group). 2011. *Impact Evaluations in Agriculture: An Assessment of the Evidence*. Washington, DC: World Bank.

Jones, K., L. L. Glenna, and E. Weltzien. 2014. "Assessing Participatory Processes and Outcomes in Agricultural Research for Development From Participants" Perspectives." *Journal of Rural Studies* 35: 91–100. doi:10.1016/j.jrurstud.2014.04.010.

Jørs, E., F. Konradsen, O. Huici, R. C. Morant, J. Volk, and F. Lander. 2016. "Impact of Training Bolivian Farmers on Integrated Pest Management and Diffusion of Knowledge to

Neighboring Farmers." *Journal of Agromedicine* 21: 200–208. doi:10.1080/1059924X.2016.1143428.

King, C., J. Gaffiiey, and J. Gunton. 2008. "Does Participatory Action Learning Make a Difference ? Perspectives of Effective Learning Tools and Indicators From the Conservation Cropping Group in North Queensland, Australia." *The Journal of Agricultural Education and Extension* 7: 133–146. doi:10.1080/13892240108438821.

Klerkx, L., F. Landini, and H. Santoyo-Cortés. 2016. "Agricultural Extension in Latin America: Current Dynamics of Pluralistic Advisory Systems in Heterogeneous Contexts." *The Journal of Agricultural Education and Extension* 22: 389–397. doi:10.1080/1389224X.2016.1227044.

Kraaijvanger, R., T. Veldkamp, and C. Almekinders. 2016. "Considering Change: Evaluating Four Years of Participatory Experimentation with Farmers in Tigray (Ethiopia) Highlighting Both Functional and Human-social Aspects." *Agricultural Systems* 147: 38–50. doi:10.1016/j.agsy.2016.05.001.

Larsen, A. F., and H. B. Lilleør. 2014. "Beyond the Field: The Impact of Farmer Field Schools on Food Security and Poverty Alleviation." *World Development* 64: 843–859. doi:10.1016/j.worlddev.2014.07.003.

Läpple, D., and T. Hennessy. 2015. "Assessing the Impact of Financial Incentives in Extension Programmes Evidence From Ireland." *Journal of Agricultural Economics* 66: 781–795. doi:10.1111/1477-9552.12108.

Läpple, D., T. Hennessy, and C. Newman. 2013. "Quantifying the Economic Return to Participatory Extension Programmes in Ireland: An Endogenous Switching Regression Analysis." *Journal of Agricultural Economics* 64: 467–482. doi:10.1111/1477-9552.12000.

Lund, T., I. Nyborg, M. H. Rahman, and M. G. Sathre. 2013. "Social Impacts of IPM-FFS on Urban and Peri-Urban Vegetable Producers in Cotonou, Benin." *African Journal of Food, Agriculture, Nutrition and Development* 13: 7727–7745. doi:10.1017/S1742758410000020.

Mancini, F., A. J. Termorshuizen, J. L. S. Jiggins, and A. H. C. van Bruggen. 2008. "Increasing the Environmental and Social Sustainability of Cotton Farming Through Farmer Education in Andhra Pradesh, India." *Agricultural Systems* 96: 16–25. doi:10.1016/j.agsy.2007.05.001.

Meijer, S. S., D. Catacutan, O. C. Ajayi, G. W. Sileshi, and M. Nieuwenhuis. 2015. "The Role of Knowledge, Attitudes and Perceptions in the Uptake of Agricultural and Agroforestry Innovations among Smallholder Farmers in Sub-Saharan Africa." *International Journal of Agricultural Sustainability* 13: 40–54. doi:10.1080/14735903.2014.912493.

Montuschi, E. 2014. *Scientific Objectivity, in: Philosophy of Social Science: A New Introduction*. Oxford: Oxford University Press. 123–144.

Munro, E. 2014. "Evidence-based Policy." In *Philosophy of Social Science: A New Introduction*, edited by N. Cartwright and E. Montuschi, 48–67. Oxford: Oxford University Press.

Murray, P. 2000. "Evaluating Participatory Extension Programs: Challenges and Problems." *Australian Journal of Experimental Agriculture* 40: 519–526. doi:10.1071/EA99079.

Mutandwa, E., and J. Mpangwa. 2004. "An Assessment of the Impact of Farmer Field Schools on Integrated Pest Management Dissemination and Use : Evidence From Smallholder Cotton Farmers in the Lowveld Area of Zimbabwe." *African Journal of Sustainable Development* 6: 24–31.

Pamuk, H., E. Bulte, and A. A. Adekunle. 2014. "Do Decentralized Innovation Systems Promote Agricultural Technology Adoption? Experimental Evidence From Africa." *Food Policy* 44: 227–236. doi:10.1016/j.foodpol.2013.09.015.

Pamuk, H., E. Bulte, A. Adekunle, and A. Diagne. 2015. "Decentralised Innovation Systems and Poverty Reduction: Experimental Evidence from Central Africa." *European Review of Agricultural Economics* 42: 99–127. doi:10.1093/erae/jbu007.

Prager, K., and R. Creaney. 2017. "Achieving On-farm Practice Change Through Facilitated Group Learning: Evaluating the Effectiveness of Monitor Farms and Discussion Groups." *Journal of Rural Studies* 56: 1–11. doi:10.1016/j.jrurstud.2017.09.002.

Pretty, J. N., and R. Chambers. 2003. "Toward a Learning Paradigm: New Professionalism and Institutions for Agriculture." In *Rethink. Sustain. Power, Knowledge, Institutions (Evolving Values a Capital World)*, 189.

Rejesus, R. M., M. E. M. Mutuc, M. Yasar, A. V. Lapitan, F. G. Palis, and T. T. N. Chi. 2012. "Sending Vietnamese Rice Farmers Back to School: Further Evidence on the Impacts of Farmer Field Schools." *Canadian Journal of Agricultural Economics/Revue Canadienne D'agroeconomie* 60: 407–426. doi:10.1111/j.1744-7976.2011.01242.x.

Rejesus, R. M., F. G. Palis, A. V. Lapitan, T. T. N. Chi, and M. Hossain. 2009. "The Impact of Integrated Pest Management Information Dissemination Methods on Insecticide Use and Efficiency: Evidence from Rice Producers in South Vietnam." *Review of Agricultural Economics* 31, doi:10.1111/j.1467-9353.2009.01468.x.

Roche, S. M., A. Jones-Bitton, M. Meehan, M. Von Massow, and D. F. Kelton. 2015. "Evaluating the Effect of Focus Farms on Ontario Dairy Producers" Knowledge, Attitudes, and Behavior Toward Control of Johne's Disease." *Journal of Dairy Science* 98: 5222–5240. doi:10.3168/jds.2014-8765.

Rodriguez, D. G. P., R. M. Rejesus, and C. T. Aragon. 2007. "Impacts of an Agricultural Development Program for Poor Coconut Producers in the Philippines: An Approach Using Panel Data and Propensity Score Matching Techniques." *Journal of Agricultural and Resource Economics* 32: 534–557. doi:10.2307/40982695.

Rogers, E. M. 2003. *Diffusion of Innovations*. New York: Free Press.

Schreinemachers, P., M. H. Wu, M. N. Uddin, S. Ahmad, and P. Hanson. 2016. "Farmer Training in Off-season Vegetables: Effects on Income and Pesticide Use in Bangladesh." *Food Policy* 61: 132–140. doi:10.1016/j.foodpol.2016.03.002.

Scoones, I., and J. Thompson. 2009. *Farmer First Revisited: Innovation for Agricultural Research and Development*. 1st ed. Rugby: Practical Action.

The Scottish Government. 2017. *The Draft Third Report on Policies and Proposals 2017-2032*. Edinburgh: The Scottish Government.

Simpson, B. M., and M. Owensb56. 2002. "Farmer Field Schools and the Future of Agricultural Extension in Africa 3rd Runner-up Outstanding." Paper presented at the 18th annual association for International Agricultural and Extension Education Conference, Durban, May 26–30, 9: 29–36. doi:10.5191/jiaee.2002.09204.

Stuart, E. A. 2010. "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science* 25: 1–21. doi:10.1214/09-STS313.

Tamini, L. D. 2011. "A Nonparametric Analysis of the Impact of Agri-environmental Advisory Activities on Best Management Practice Adoption: A Case Study of Quebec." *Ecological Economics* 70: 1363–1374. doi:10.1016/j.ecolecon.2011.02.012.

Tin, H. Q., P. C. Struik, L. L. Price, N. P. Tuyen, N. P. Hoan, and H. Bos. 2010. "Increase of Farmers' Knowledge Through Farmer Seed Production Schools in Vietnam as Assessed on the Basis of Ex-ante and Ex-post Tests." *The Journal of Agricultural Education and Extension* 16: 229–247. doi:10.1080/1389224X.2010.489766.

Todo, Y., and R. Takahashi. 2013. "Impact of Farmer Field School on Agricultural Impact and Skills: Evidence from an Aid-funded Project in Rural Ethiopia." *Journal of International Development*, 362–381. doi:10.1002/jid.

Togbé, C. E., R. Haagsma, A. K. N. Aoudji, and S. D. Vodouhê. 2014. "Effect of Participatory Research on Farmers, Knowledge and Practice of IPM: The Case of Cotton in Benin." *The Journal of Agricultural Education and Extension* 8622: 1–20. doi:10.1080/1389224X.2014.971829.

United Nations. 2018. *World Economic Situation and Prospects 2018*. New York: United Nations.

Van den Berg, H. 2004. *IPM Farmer Field Schools : A Synthesis of 25 Impact Evaluations*. Office 53.

Van Den Berg, H., and J. Jiggins. 2007. Investing in Farmers – The Impacts of Farmer Field Schools in Relation to Integrated Pest Management. *World Development* 35: 663–686. doi:10.1016/j.worlddev.2006.05.004

Wafula, L., J. Oduol, W. Oluoch-Kosura, J. Muriuki, J. Okello, and J. Mowo. 2016. "Does Strengthening Technical Capacity of Smallholder Farmers Enhance Adoption of Conservation Practices? The Case of Conservation Agriculture with Trees in Kenya." *Agroforestry Systems* 90: 1045–1059. doi:10.1007/s10457-015-9882-y.