

IDS Bulletin

Transforming Development Knowledge

Volume 49 | Number 4 | September 2018

THE MILLENNIUM VILLAGES: LESSONS ON EVALUATING INTEGRATED RURAL DEVELOPMENT

Editor **Chris Barnett**



| | |
|--|-----|
| Notes on Contributors | iii |
| Foreword Richard Longhurst | vii |
| Introduction: Lessons from the Millennium Villages Evaluation; Where Next for Integrated Development? Chris Barnett | 1 |
| Integrated Development, Past and Present Edoardo Masset | 17 |
| The Cost-Effectiveness of Complex Projects: A Systematic Review of Methodologies Edoardo Masset, Giulia Mascagni, Arnab Acharya, Eva-Maria Egger and Amrita Saha | 33 |
| Assessing Value for Money in Integrated Development Programmes – The Case of a Millennium Villages Project in Northern Ghana Arnab Acharya and Tom Hilton | 53 |
| <u>Abductive Reasoning to Explain Integrated Development: Lessons from the Multi-Method Evaluation of the Millennium Villages Project</u> Dee Jupp and Chris Barnett | 67 |
| Can Immersion Research Add Value in Understanding Integrated Programme Interventions? Dee Jupp, David Korboe and Tony Dogbe | 83 |
| Learning About Integrated Development Using Longitudinal Mixed Methods Programme Evaluation Emily Namey, Lisa C. Laumann and Annette N. Brown | 97 |
| Applying Factorial Designs to Disentangle the Effects of Integrated Development Holly M. Burke, Mario Chen and Annette N. Brown | 115 |
| Glossary | 129 |

Abductive Reasoning to Explain Integrated Development: Lessons from the Multi-Method Evaluation of the Millennium Villages Project*

Dee Jupp¹ and Chris Barnett²

Abstract The use of mixed methods in impact evaluation often focuses on the triangulation of findings or using qualitative data to explain statistical assessments of impact (net effect). In this article, we argue for the benefits of combining and contesting *ex ante* etic-informed theory with *ex post* emic-informed grounded theory. Using the case of the Millennium Villages Project (MVP) evaluation in northern Ghana, we highlight examples of where emergent theory led to a deeper understanding that valued how local people view change. In the context of an independent evaluation that set out to assess the MVP model – and with little scope for the evaluation to influence or adapt implementation – the theories of change created abductively helped fill a theoretical void. We argue that this is particularly applicable for integrated programmes where *ex ante* theory may struggle to capture the complex interaction of different sector-based activities, overlook unintended effects, and undervalue local perceptions of change.

Keywords: integrated development, Millennium Villages, integrated rural development, multisector, theory of change, impact evaluation, complexity, local perspectives, participatory approaches, participant observation.

1 Introduction

Whilst many acknowledge the virtues of mixed methods research as ‘the best of both worlds’, the focus for mixing in international development evaluations³ has often been on the triangulation of findings through the careful integration, complementation, and sequencing of methods.⁴ This article describes the insights gained – within the confines of an external evaluation – by comparing *ex ante* theory with *ex post* theory and the combination of inductive and deductive modes of reasoning. This leads to the use of abductive (pragmatist, explanation-driven) reasoning to explain impact. The

© 2018 The Authors. *IDS Bulletin* © Institute of Development Studies | DOI: 10.19088/1968-2018.162



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non Commercial No Derivatives 4.0 International licence (CC BY-NC-ND), which permits use and distribution in any medium, provided the original authors and source are credited, the work is not used for commercial purposes, and no modifications or adaptations are made. <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

The *IDS Bulletin* is published by Institute of Development Studies, Library Road, Brighton BN1 9RE, UK
This article is part of *IDS Bulletin* Vol. 49 No. 4 September 2018 ‘The Millennium Villages: Lessons on Evaluating Integrated Rural Development’; the Introduction is also recommended reading.

article uses a real-life example of a longitudinal evaluation of the Millennium Villages Project (MVP) in northern Ghana and highlights the challenge of developing specific theories for integrated rural development projects – a point also highlighted by Masset (this *IDS Bulletin*). The MVP evaluation team combined theory-based impact evaluation (TBIE)⁵ (such as that advocated by White 2009; Jimenez and Puri 2017) with naturalistic, contextual studies which generated grounded theory. We argue that combining *ex ante* theory-based approaches with *ex post* grounded (emergent) theory acknowledges the value of both, goes beyond simple triangulation of findings, and generates more complex-aware explanations of change processes on which to base future development interventions. The article also highlights the essentiality for evaluators to understand, value, and embrace different ways of generating knowledge and being willing to do so – no easy task for multidisciplinary teams.

2 Background to the independent evaluation

The impact evaluation of the MVP in northern Ghana was commissioned by the UK Department for International Development (DFID) as a multi-year external evaluation of the Millennium Village (MV) model. For DFID, an independent external evaluation was a condition of providing funding for the project, which started in 2012 and ran for five years. At that time, there had been little robust evidence to support the claims of the MV model as a means of achieving the Millennium Development Goals (MDGs) at scale, with much resultant controversy in social media and the press,⁶ as well as between leading academics and policymakers (e.g. Clemens and Demombynes 2013).

As such, the evaluation was required to be independent, and operationally separate, in order to permit the MV model to be implemented as intended by the project team and using the project's intervention logic as the basis to verify results. This is in contrast to many complex-aware evaluations currently carried out that emphasise the continuous development and adaptation of theories of change in response to learning and feedback (Patton 2011; Andrews *et al.* 2015). Rather, the evaluation of the MVP was a TBIE along the lines advocated by 3ie and others, which has become a dominant mode of evaluation to measure and attribute the impact of interventions (White 2009, 2011). Such evaluations depend on the development of a robust theory of change which articulates the sequence of results of an intervention, including how the causal chain is expected to work and the contextual (external) factors that help determine change. Good theories of change are constructed through careful processes of incorporating stakeholder views and existing evidence, and benefit from the contributions of both quantitative and qualitative exploratory research. TBIEs may draw on different approaches such as theory-based evaluation (Weiss 1997), theory-driven evaluation (Chen 2012), or contribution analysis (Mayne 2001), but all subscribe to the notion that the theory is primarily developed *ex ante* and the evaluation is intended to verify the extent to which the theory matches what is observed.

By contrast, we present the case where the *ex ante* theory of change is critically compared with *ex post* grounded theory in order to enrich analysis of the change process. Drawing on the different traditions of research practice – and fusing inductive and deductive approaches – helps evaluators from different disciplines to interrogate findings and provide better explanations of change. This we explore in terms of abductive theories of change that examine outcomes, and seek to find the best possible explanation for these outcomes whilst retaining the proviso that the contributions and attributions *may* be true. As Reichertz explains:

It is precisely in this quality of being a ‘means-of-inferencing’ that we find the secret charm of abduction. On the one hand it is a logical inference (and thereby reasonable and scientific), and on the other hand it extends into the realm of profound insight (and therefore generates new knowledge) (2014: 300).

We argue that, as such, the process of abduction is particularly important for integrated programmes where theory-driven approaches alone may struggle to capture the complexity of how different sector-based interventions interact, overlook unintended or unforeseen effects, and undervalue how change is perceived by the intended beneficiaries of the intervention.

3 The challenge of ‘theoretical eclecticism’

The initial challenge for the evaluation team was to understand the theory of the MV model in order to design the TBIE. The *Terms of Reference* (ToR) for the UK DFID-commissioned evaluation of the MVP noted:

The MV model provides an integrated package of interventions to lift a rural community out of poverty. Its central hypothesis is that a local ‘big push’ addressing the most immediate capital deficiencies in communities and households is a necessary condition for reaching a threshold required to move towards local resilience and self-sustaining economic growth.⁷

The ‘big push’ was designed to break the poverty trap⁸ and this was the theoretical basis provided by the MV pioneers (UN Millennium Project 2005: 39). It is, in essence, the idea that massive scaling-up of resources to the poor will lead to ‘take-off’ which in turn will lead to a virtuous circle of self-sustaining growth. Many observers found this to be an insufficient theoretical basis and offered more robust theories, including that the MV model followed a ‘market-oriented approach’ (Broad and Cavanagh 2006), Sen’s capability approach (Diepeveen 2008), and modernisation theory (Carr 2008). The variety of possibilities led Carr to describe this as ‘theoretical eclecticism’ and he wrote:

Thus, the MVP, in the absence of overt theoretical considerations, plucks ‘useful’ aspects of theory from their contexts, and amalgamates them into a single, hybrid approach to development that contains points of access and purchase for academics and policy makers of all theoretical persuasions: in other words, the MVP has something for everyone (2008: 338).

Whilst links can be made to aspects of all these theories, the reality was that the MVP had itself articulated no clear overarching mid-range theory,⁹ except loosely something based on the grand theory of the 'poverty trap'. In practice, the MV model was a collection of interventions designed to demonstrate how the MDGs could be achieved at a local level. Carr (2008) noted that 'Sachs and McArthur (2005: 347) stated clearly that the MVP was built upon the "core truth" that there were "known packages of effective and generally low-cost interventions" that can and should be applied to the challenges of extreme poverty'.

4 The hunt for theory

The external evaluation team searched for other 'theory' on which to base the evaluation, recognising that these had to be promulgated and owned by the project – not purely an external interpretation. Apart from the grand theory of the 'poverty trap' mentioned above, a couple of others emerged. First, a set of generic 'intervention logics' provided by the project showed the anticipated interconnections from inputs-to-outputs-to-outcomes and then to MDG impacts. These were not, however, specific to the northern Ghana MVP, made no reference to planned interventions and connections that would really occur, and were largely silent on the sequencing and synergies between different sectoral activities. Second, there was a logical framework (logframe) required of the project funders, DFID. This attempted to fit the complexity of the MVP into a simple, linear hierarchy from activities through to impact. While this was specific to the northern Ghana MVP, in essence it listed the intervention packages and MDG indicators by which their achievements would be assessed. This pared-down representation of the project did not do justice to its integrated nature and intended synergy – with no real detail on the assumptions and multidirectional interconnections. In summary, none of these options provided a basis against which to evaluate the MVP.¹⁰

The evaluation team decided to take a pragmatist view of the evaluation, combining a suite of evaluation methods to best answer the broad question of whether the integrated approach worked (or otherwise).¹¹ The decision to take a pragmatist approach did not diminish the challenge posed by the fuzziness around the programme's theoretical basis – one that reoccurred in dialogue with the evaluation's peer review group (PRG) throughout the evaluation's five years. Clearly a theory-based impact evaluation would require inferring a theory. But what would be the best way? And how to ensure that this remained true to the intentions of the MVP?

5 Combining deductive and inductive approaches

The most common way in which theories of change are developed are essentially deductively. This is where the team started, with *ex ante* theories of change developed through a deductive process informed by information from the project, and implicitly the team's mental models. By midline, for example, it was possible to create a series of

causal chains from a more detailed and nuanced examination of the data, as well as from the MVP's reporting. Plus, prior to the endline in 2016, another set of causal chains was developed in a workshop, drawing on actual interventions identified from the MVP progress reports, with some verification in the field, including with project staff. The limitations of this approach were nonetheless acknowledged by the team, noting that the theories of change have been 'split sectorally for clarity and ease of interpretation' and 'this approach has limitations and does not fully take into account the complexity of an integrated programme such as MVP and the added value [synergy] of multi-sectoral implementation' (Nelson *et al.* 2017: 3).

It was in the latter part of the evaluation that the team deliberately pitted the *ex ante* and *ex post* theories against each other. The Reality Check Approach (RCA)¹² – unlike all the other quantitative and qualitative streams within the evaluation – was not framed by *ex ante* theory, and used themes that emerged from conversations, observations, and experience during immersion in villages to construct grounded theory (in this case, *ex post*). This was an emic view of change experienced and perceived by people living in the MV communities themselves. As evaluators, it was important to undertake this comparison. This was not only because there had been a lack of clarity on the theoretical underpinnings of the project, but also because of the constraints imposed by being independent from the project and providing an accountability function with limited influence.

To illustrate the advantages of bringing deductive theory development and inductive theory development together, we discuss just two examples from the MVP evaluation: agricultural production, and income/consumption dynamics. The explanations for change developed for each of these phenomena benefited from exploring both *ex ante* (deductive) theoretical perspectives and *ex post* (grounded) theory perspectives and resulted in a more balanced interpretation of results.

6 Example 1: agricultural production

Before the start of the project, the MVP had identified maize production as a potential profitable crop for this area of northern Ghana. The project directly supported this by provision of subsidised improved seeds, a fertiliser advance scheme, tractor hire, and training on agronomic practice through a lead-farmer approach, which worked through farmer organisations (enabled by the provision of motorbikes and fuel, provision of additional agricultural extension posts, and salary 'top-ups' for these and the original extension officers) and indirectly through improved roads and telecommunications. The *ex ante* theory of change developed from the project intentions indicated that this (along with a focus on a few other selected potential crops) would lead to increased income, food security, and nutrition.

Based on the statistical (difference-in-difference) analysis of MVs and control villages (CVs), the evaluation finds that:

[Overall,] the MVP has had a sizeable impact on agricultural production... with an approximately 38% rise over the period of implementation. This can be explained mostly by input increases (fertilisers, seeds, land, tractor rents and other animals/machinery for hire), with 74% of the productivity increase explained in this way... The land area dedicated to maize and beans¹³ has increased considerably in the MV areas (Barnett *et al.* 2018: 166).

and

the MVP had a large impact on food security, with a reduction in the number of households reporting not having enough food in the previous 12 months (*ibid.*: 126).

This seemed to validate the *ex ante* theory of change with the project inputs and activities being directly and wholly responsible for the improvements. But, was this the complete picture of what happened? The qualitative studies suggested a slightly different scenario. In 2013, there had been a major problem with the provision of fertiliser in time for maize production, tensions over repayment arrangements, and eventually an abandonment of this form of support. There was some evidence that tractors, seeds, and training may have been preferentially provided to larger landowners. But nevertheless, there was a palpable 'buzz' amongst less well-off farmers included in participatory rural appraisal (PRA) focus groups and those met during the RCA that people were getting good prices for maize (market price increased between 2012 and 2015)¹⁴ and prices for the traditional crop of millet were decreasing. Previously, farmers in the area had grown millet to make their staple dish of *tuo zaafi* (TZ) and to brew alcohol. It was becoming common knowledge that maize could replace millet for both traditional purposes (TZ and alcohol). Millet has a longer growing season and is not harvested until late October/November and farmers soon heard that maize could be cultivated in a matter of three months (late April to early August), thereby freeing up the land for a dry season crop. The crop of choice was cowpeas, as it is sown soon after the maize is harvested (in July/August) and matures in about two months. Cowpeas are relatively drought-tolerant and the harvested dried beans are easy to store. By rotating with maize, farmers also benefit from the nitrogen fixation attributes of this legume.

Interestingly, the MVP had not articulated this potential in their bundle of intervention strategies and had not promoted cowpea production. As a consequence, the evaluation team did also not include this in their articulation of causal chains. The understanding of this phenomenon grew over time. By 2015, the midline RCA study noted that there was a clear distinction emerging between families who were able to grow cowpeas and those that did not. By 2017, the RCA study noted 'a slight improvement in the availability of cooked food each day in homes and more willingness to serve beans with "*tuo zaafi*" than previous years' (Jupp 2018: 48). Also, the PRA focus groups:

Box 1 The extraordinary rise of cowpea cultivation

Intrigued by so many people telling us how key cowpea cultivation had become in their lives and how many described a change in their outlook since they started to grow it, we tried to establish the drivers of this phenomenal uptake. Five years ago, cowpea was grown in riverine areas (e.g. Builsa North) but generally for consumption only as yields were not particularly good (less than three bags per acre) and no insecticides or weedkillers were used. People told us they began to hear about a private farmer in Yagba who was ‘getting 15 bags from an acre’ and wanted to find out how he did this and found that he achieved these yields by using chemicals bought from a single supplier in Bolga. Dealers quickly saw that demand for these chemicals was increasing and opened a number of outlets, permanent and at weekly markets, especially in Fumbisi. People copied the practice during the 2012–14 seasons and were excited by the yields they got and the fact that it filled an otherwise relatively unproductive farming period (October–December). Soon buyers were sending trucks from Kumasi, Tamale, and Techiman, and farmers shared with us that there is a ‘big demand down south’. Large landowners also became interested in cultivating cowpeas, providing sharecropping opportunities too. People described the market as commercial and competitive and shared the advantages of storage if they can afford it.

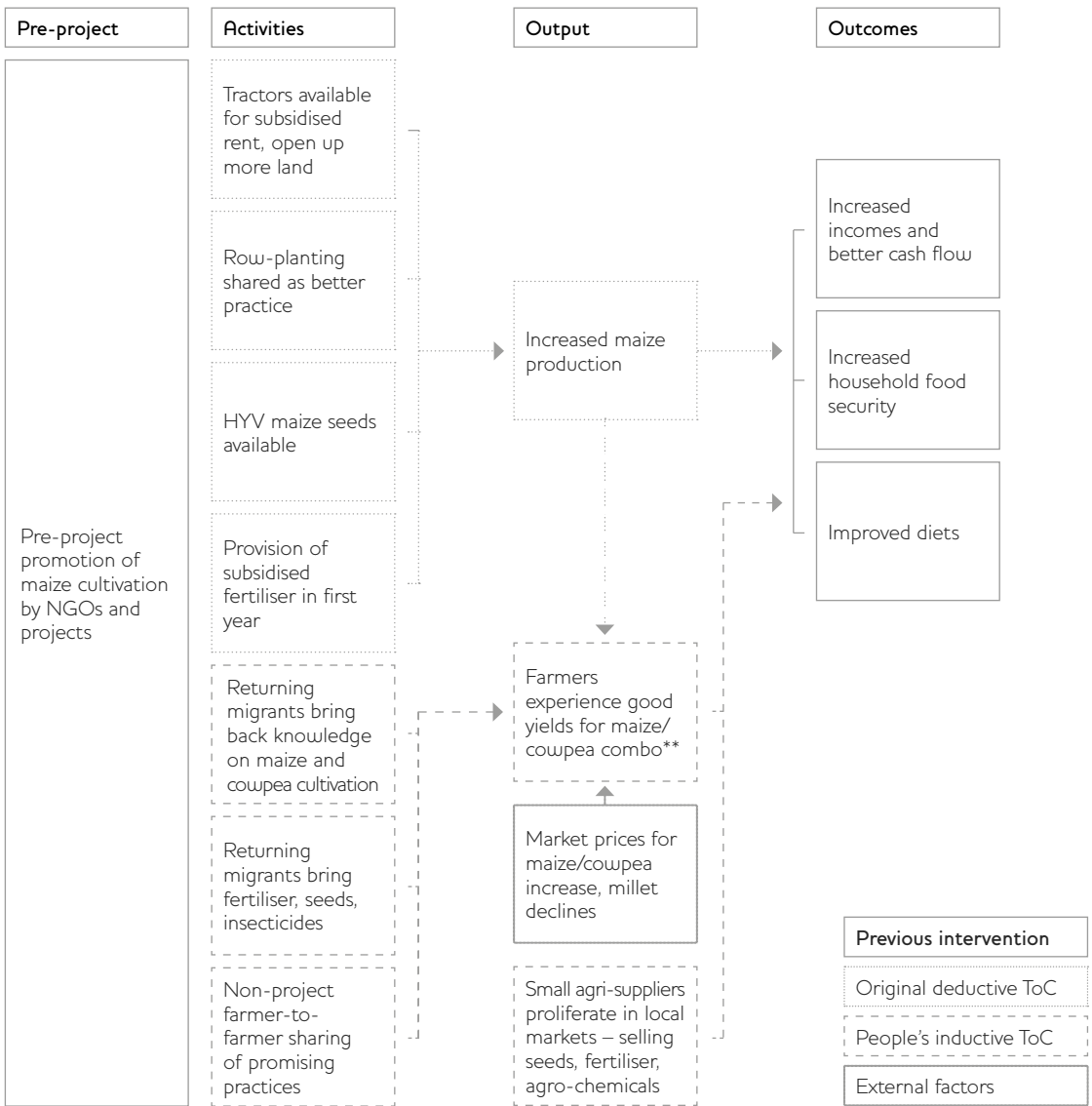
Source: Barnett *et al.* (2018: 119–20).

indicated that MV farmers transferred skills they were taught about soybean cultivation to farming the higher-return cowpea crop. In contrast, farmers in those CVs [control villages] where cowpea has become a key crop often said they had picked up the requisite knowhow during their migration stints to the fertile so-called ‘overseas’ areas and along the White Volta, where they often hired themselves out as farm labourers on cowpea farms (*ibid.*: 99).

The RCA team were interested by the surprising rapid introduction and success with cowpeas over the five years that they had intermittently stayed with families and sought to understand how this had occurred. Box 1 describes what farmers shared.

The PRA added to this understanding by establishing that families were providing improved diets for their children mostly because they were adding cowpeas and eggs. Soybean takes longer than cowpeas to grow, takes longer to cook, and the taste is liked less than that of soybean, so the (limited) attempts to encourage its production by MVP were largely unsuccessful. Cowpeas are often now referred to as the ‘hungry season crop’.

Figure 1 Simplified schematic illustrating the development of the abductive theory



Notes NGO – non-governmental organisation; HYV – high-yielding variety; ToC – theory of change.

**Maize/coupea – two quick-growing, relatively drought-resistant crops that reduce risk from climate change, and increase productivity.

Source Authors' own.

Combining these insights and constructing a grounded theory of change alongside the *ex ante* theories of change, the evaluation team was able to abductively create a new theory of change (illustrated in Figure 1). The evaluation team concluded that the abductive best fit explanation was that the introduction of maize to substitute for millet was the key driver of change (for which the MVP can take some credit), but importantly it was the subsequent take-up of a second crop of

cowpeas in the dry season (due to factors beyond the project) that made an important contribution to income and food security.

7 Example 2: income and consumption dynamics

Our next example of *ex ante* and *ex post theories* combining to provide alternative explanations of phenomena concerns income and consumption data. Based on an analysis of the household survey, the midline evaluation report explained:

We calculated monetary poverty rates using the methodology and the poverty lines used by the Ghanaian Statistical Service. Monetary measures of poverty are based on consumption-expenditure data at household level which include purchases as well as own-produced items and in-kind gifts from other households. Consumption is an overall measure of well-being as it includes the consumption of food, education items, health, non-food items and an evaluation of the value of the use of durable goods and housing. Households that fail to consume a minimum basket of essential goods are classified as poor (Masset *et al.* 2016: 45).

It went on to suggest that the project had had no impact on overall poverty but did have a significant effect on food poverty.

The qualitative strands (PRA and RCA) which complemented the survey showed that households employed more diverse, locally relevant descriptors of poverty/wellbeing in analysing changes in their conditions. Having cash in what had hitherto been a largely cashless society was highlighted in both PRA and RCA as an important element of ‘not feeling poor’, as was having family members who could migrate south for cash-earning opportunities in the dry season (*ibid.*: 48). The RCA study indicated that indebtedness was of grave concern amongst some poorer families with repayment difficulties leading to concomitant stress, concern for the future, and exclusion from further formal and informal loans (*ibid.*: 49). The PRA study facilitated people’s self-generated descriptors of wellbeing from focus groups which included the multidimensional aspects of dignity, food security, personal health, educational attainment, tangible assets, and the ability to recover after financial shocks.

The endline report indicated that the project had had some effect on income, but expenditure remained unchanged. The endline report noted that:

the fact that incomes increase while expenditure remains unchanged is counterintuitive and deserves some additional explanation. We would expect an extremely deprived population to increase their consumption as their income increases, but this is not what is observed from the survey data. One hypothesis is that the households in our study appear to have perceived the increase in incomes brought about by the intervention as temporary and have not adjusted their consumption upward (Barnett *et al.* 2018: 69).

Box 2 Funerals: a significant but underestimated expenditure?

Despite the MVP reportedly having made efforts to educate citizens on the impoverishing impact of funerals... and the need to control such expenditures, there has been little obvious impact on this deeply entrenched behaviour. As highlighted in earlier reports, people shared that surplus cash is often earmarked for funerals, which may be delayed for years if there are insufficient resources to '*honour the dead properly*'.... Arguably though, funerals can be viewed as an investment both to ensure the dead join the spirits of the ancestors and concomitantly can provide guidance and protection of the living as typified by this quote (from a PRA focus group discussion): '*[I]f these funerals are not performed, households will not progress and their ancestors will [cease to] help them in their farming and other endeavours*'. Both the RCA and PRA studies find that people say that funerals take up significant shares of surpluses in most households. During the midline RCA, people helped to construct the average cost for the two parts of the funeral ceremony (*Kumka* and *Juka*) and calculated a figure of around GHS1,200.

The PRA study found that especially in the Bul'k (Builsa) project area, funerals threaten to reverse much of the gains from improved agricultural productivity. The overwhelming majority of those interviewed felt that funeral expenditures are still rising. Large quantities of millet and maize are used to feed the throngs of mourners who congregate from near and far, but also in brewing the alcoholic beverages served for refreshment. Livestock expenditures are equally high, with households yielding to longstanding cultural prescriptions obliging them to make animal sacrifices to their ancestors. Furthermore, both the PRA and RCA studies find that funeral spending has increased as contemporary expectations have increased. The RCA and PRA researchers experienced over the years a trend to play modern music on (hired) hi-fi systems and to serve guests with bottled beverages rather than the traditional *pito* (a local beer brewed from millet and served lukewarm). A ban by the Bul'k paramouncy on serving (cheap) local gin at funerals (because many people end up intoxicated) has contributed to the shift to (more expensive) bottled beverages. Increasingly, the actual food served also comes in takeaway packs, which is considered a mark of sophistication. A participant at MV4 [a village under the MVP] noted [that] '*we are now integrating the Ashanti culture into ours*.' Others expressed the opinion that the rich have set new standards which they feel '*obliged to meet, to avoid disgrace*'.

Source: Barnett *et al.* (2018: 69–70).

Initially, this seemed counter to the RCA and PRA findings that suggest that some people are consuming more (e.g. marriages are bigger with larger sound systems, funerals are bigger with more gifts for guests, etc.), there is more alcohol consumption, more eating out, and more expenditure on TVs, renting DVDs, more motorbikes, and fuel costs. Box 2 describes the increase in funeral expenditure over the period of the MVP.

The explanations emerging from the qualitative streams had support in the research literature. For example, a study with the Kasena people who live in northern Ghana reported:

Funeral rites take place sometime after burial of the deceased, the time lapse depending on the family circumstances – consensus among the immediate kin and elders about when to hold the rites and more importantly, the availability of quantities of foodstuff necessary for the rites (Awedoba 2014).

They describe these as the most important social events and ‘the occasion for conspicuous consumption’ which can be postponed for years.

Looking at the evidence abductively left the team puzzled. People had been keen to share that cowpeas had increased their cash incomes and that they had bought assets – notably TVs and motorbikes, and made improvements to their houses – but were these small investments (often made through credit repayments) too small to record a difference between MVs and CVs? These were all first-order purchases for those with cash profits from cowpeas. Second-order expenditure typically included that set aside for delayed social obligations (funerals and marriages) which, given the increasing lavishness of such celebrations, required savings from more than one crop season.

Eventually, the team settled on another explanation based on distinguishing between two types of expenditure: firstly, the consumption of current goods, which is employed in official poverty statistics (and which is not increasing in this case); and secondly, the purchase of durable goods and extraordinary expenditure, which is not included in official poverty statistics. This second type of expenditure (either as extraordinary goods or savings) *is* increasing – as supported by evidence from both the statistical analysis and the qualitative studies. Therefore, while households in the study area are perceiving the increased incomes resulting from the MVP as temporary, they are also altering their consumption patterns in ways that are underestimated by not being included in the expenditure used in official poverty statistics (such as gifts at funerals, which can be sizeable in the local culture) (Barnett *et al.* 2018: 69–70).

8 Concluding remarks

Of course, combining the perspectives of evaluator or project staff with a ‘bottom-up’ perspective of theories of change is not new, but it is also something that is not routinely done. If it is, it is often done with an

intention to improve *ex ante* theories of change. As Valters (2014) points out, the very process of developing theories of change

can also create an illusion of serious reflection by being a superficial process of critical thought, where people who engage with the theories (donors as well as implementers) do not actually reflect sufficiently on how power dynamics change in practice and how local people see change happen.

In this way, the drive for theory-based evaluations – where evidence is primarily collected against the theory – can limit the ability to see beyond mostly confirmatory explanations of how change occurs.

While the advantages of mixed methods are often extolled purely in terms of improved triangulation and complementarity in evaluation, the team recognised value in conflict and contestation from sharing findings from different methodological perspectives. Greene, Caracelli and Graham describe this process in terms of ‘the discovery of paradox and contradiction, new perspectives of frameworks, the re-casting of questions or results from one method with questions and results from the other method’ (1989: 259). This requires contesting different theoretical lenses to the data with the intention of producing more thoughtful analysis, and leading to abductive (best fit) explanations;¹⁵ which in the end would be more useful to understanding the theory of change behind the impact of the MVP.

This can generate important new insights. For instance, populist journalist Nina Munk followed the progress of a group of Millennium Villages for several years, and in her 2013 book *The Idealist: Jeffrey Sachs and the Quest to End Poverty*, she said a basic flaw of the MVP is that it is developed by academics living far away from the subject areas and with a poor understanding of local cultures, who do things such as promoting growing maize amongst people who have not historically eaten it or building a short-lived livestock market when there was no local demand. This is not wholly true and underestimates the contribution of the project: in the case of the northern Ghana MVP, maize *did* provide a breakthrough but the way in which it did was different to what was expected – i.e. mostly because it provided a cultivation window for the easy-to-grow, profitable, highly nutritious, and much liked cowpea. This insight was only possible from looking at a people’s perspective of change and inductively improving on the deductive *ex ante* theory of change.

The examples provided in this article are just a ‘taster’ of many others the evaluation team encountered through purposely combining *ex ante* deductive and *ex post* inductive theoretical perspectives. The resultant abductive explanations require that all evidence is considered together and provides a better basis for making conclusions about what should be replicated or scaled up. Most ToRs for evaluation typically require the elaboration of theory as a pre-requisite, but we argue that they should also include a requirement to look at the emerging data *ex post*

through the experience of people who were intended beneficiaries of the intervention. The differences can be profound. Schindler, Graef and König, for example, studied causal chain explanations for agricultural upgrading strategies (UPS) and concluded:

We observed that farmers and scientists have considerably different views on the positive and negative impacts of proposed agricultural UPS. While scientists focus mostly on direct causal impact chains of the UPS, the farmers consider indirect linkages that take their complex livelihoods into account (2016: 42).

They go on to explain that,

while scientists assessed very high positive impacts on yield because of fertilizer application, the farmers also considered in their evaluation the high risk of lack of rain and chemical fertilizer application, which would increase yield failure even further during drought years compared to not applying fertilizer (*ibid.*).

However, this example is confined to the co-production of *ex ante* theory of change. In situations where external evaluation must be independent, the review of both *ex ante* (project-led) theory and *ex post* (people-led) theory fills an important gap.

It is in this way that we can get closer to providing better explanations of what happened and make better predictions of what might happen if aspects of the programme are reproduced. This is a particular challenge for those evaluating models of complex, integrated projects, where the multiplicity of interconnections are difficult to fully know from the start. It is even more crucial for evaluations that primarily have an accountability function, and where the main motivation is to ‘test’ and better understand whether the model of integration works – rather than continuously adapting the model. As Carr (2008: 336) notes in reference to MVP sites: “‘known packages of effective interventions’ may not capture the complex linkages between sectoral issues that result in the local challenges identified as problems by villagers’”. Applying abductive reasoning to theories of change is one way to better understand these linkages, as well as challenge the confirmatory bias of our own mental models.

Notes

* This issue of the *IDS Bulletin* was prepared as part of the impact evaluation of the Millennium Villages Project in northern Ghana, 2012–17, funded by the UK Department for International Development (DFID) (www.dfid.gov.uk). The evaluation was carried out by Itad (www.itad.com) in partnership with IDS (www.ids.ac.uk) and PDA-Ghana (www.pdaghana.com). The contents are the responsibility of the evaluation team and named authors, and do not necessarily reflect the views of DFID or the UK Government.

1 Technical Advisor, Empatika and External Research Associate, School of Development Studies, University of East Anglia.

- 2 Honorary Associate at IDS and Director of Technical Excellence, Itad, Hove, UK.
- 3 While the differences between research and evaluation are debatable and contested, this article focuses more narrowly on the application of different methods to the field of international evaluation. We distinguish 'evaluation' from the 'research' more broadly in terms of evaluation's focus on: (i) a more limited set of questions (essentially around understanding change due to a project or intervention); and, (ii) its express purpose of informing stakeholders (funders, implementers, etc.) in operational and policy decisions. By 'international development evaluations', we focus on those that take place primarily in low- or middle-income countries, with a strong focus on addressing different aspects of poverty and inequality.
- 4 Such as the sequential and concurrent designs for impact evaluation described in Bamberger (2012: 9–12).
- 5 TBIE combines an experimental or quasi-experimental design (measuring the net effect of an intervention) with a theory of change (using other methods to explain the link between project activities and the net effect, or impact).
- 6 Such as in *The Economist* (www.economist.com/feast-and-famine/2012/05/14/millennium-bugs), and correspondence in *The Lancet* ([www.thelancet.com/journals/lancet/article/PIIS0140-6736\(12\)60848-4/fulltext](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(12)60848-4/fulltext)).
- 7 *Terms of Reference*, 22 September 2011 (DFID 2012: 1).
- 8 Broadly, this is an economic theory that suggests that there is a self-reinforcing mechanism which forces people to remain poor, and this is so binding that it requires a significant injection of capital for people to escape it.
- 9 Mid-range theory is defined here as a theory of change with limited scope that is able to explain a specific set of phenomena, in contrast to grand theory that might explain a phenomenon at a more abstract societal level.
- 10 Discussion based on an unpublished internal paper to the Peer Review Group, 'Responding to the PRG Comment on: Informing the Theory of Change with the Baseline Findings', 11 December 2013.
- 11 The evaluation design comprised four main strands: (1) a difference-in-difference (DD) methodology used to estimate impact, based on the difference in the change over time in the average outcomes between the project and in the comparison groups; (2) an Institutional Assessment that captured institutional change, particularly at the district and community levels; (3) an adapted form of PRA that was used to obtain feedback and insights on preliminary statistical findings from the perspective of different wellbeing groups; and (4) a reality check approach (RCA) that used a condensed immersion approach to better understand the realities of households and how they viewed change.
- 12 Immersive research where a team of researchers live in the homes of people in the community and informally interact, observing and experiencing daily life without using a preconceived evaluation framework (www.reality-check-approach.com).

- 13 'Beans' refers to cowpeas throughout.
- 14 Tables 7.1–7.3, www.e-agriculture.gov.gh/index.php/2014-07-22-14-39-46/agric-facts-and-figures-2015.
- 15 The core evaluation team included a range of professionals who describe themselves as economists, evaluators, statisticians, social scientists, and anthropologists. In discussing emerging findings, internal workshops and virtual meetings often involved forms of contestation that spanned methodological, theoretical viewpoints and worldviews – but with the aim of generating new insight.

References

- Andrews, M.; Pritchett, L.; Samji, S. and Woolcock, M. (2015) 'Building Capability by Delivering Results: Putting Problem-Driven Iterative Adaptation (PDIA) Principles into Practice', in A. Whaites, E. Gonzalez, S. Fyson and G. Teskey (eds), *A Governance Practitioner's Notebook: Alternative Ideas and Approaches*, Paris: Organisation for Economic Co-operation and Development (OECD)
- Awedoba, A.K. (2014) 'Wealth, Consumption and Migration in a West African Society. New Lifestyles and New Obligations among the Kasena, Northern Ghana', *Anthropos: International Review of Anthropology and Linguistics* 109: 45–55
- Bamberger, M. (2012) *Introduction to Mixed Methods in Impact Evaluation*, Impact Evaluation Note 3, Washington DC: InterAction and The Rockefeller Foundation
- Barnett, C. et al. (2018) *Impact Evaluation of the SADA Millennium Villages Project in Northern Ghana: Endline Summary Report*, commissioned by the UK Department for International Development, Hove: Itad
- Broad, R. and Cavanagh, J. (2006) 'The Hijacking of the Development Debate: How Friedman and Sachs Got It Wrong', *World Policy Journal* 23.2: 21–30
- Carr, E.R. (2008) 'The Millennium Village Project and African Development: Problems and Potentials', *Progress in Development Studies* 8.4: 333–44
- Chen, H.T. (2012) 'Theory-Driven Evaluation: Conceptual Framework, Application and Advancement', in R. Strobl, O. Lobermeier and W. Heitmeyer (eds), *Evaluation von Programmen und Projekten für Eine Demokratische Kultur*, Wiesbaden: Springer VS
- Clemens, M. and Demombynes, G. (2013) *The New Transparency in Development Economics: Lessons from the Millennium Villages Controversy*, CGD Working Paper 342, Washington DC: Center for Global Development
- DFID (2012) *PO 5603 MV-Evaluation: Terms of Reference, Impact Evaluation of a New Millennium Village in Northern Ghana*, ITT Volume 3, London: Department for International Development
- Diepeveen, S. (2008) 'Putting Empowerment into Practice: Evaluating the Potential for "Development as Freedom" in the Millennium Villages Project', *Undercurrent: The Canadian Undergraduate Journal of Development Studies* 5.5: 6–13

- Greene, J.; Caracelli, V. and Graham, W. (1989) 'Toward a Conceptual Framework for Mixed-Method Evaluation Designs', *Educational Evaluation and Policy Analysis* 11.3: 255–74
- Jimenez, E. and Puri, J. (2017) 'The Wicked Cases of Education and Climate Change – The Promise and Challenge of Theory-Based Impact Evaluations', in R.D. van den Berg, I. Naidoo and S.D. Tamondong (eds), *Evaluation for Agenda 2030: Providing Evidence on Progress and Sustainability*, Exeter: International Development Evaluation Association (IDEAS)
- Jupp, D. (2018) *Endline Reality Check Approach for the Millennium Villages Evaluation*, Hove: Itad
- Masset, E. et al. (2016) *Millennium Villages Evaluation: Midterm Summary Report*, commissioned by the UK Department for International Development, Hove: Itad
- Mayne, J. (2001) 'Addressing Attribution through Contribution Analysis: Using Performance Measures Sensibly', *The Canadian Journal of Program Evaluation* 16.1: 1–24
- Munk, N. (2013) *The Idealist: Jeffrey Sachs and the Quest to End Poverty*, New York NY: Doubleday
- Nelson, K. et al. (2017) 'Millennium Villages Evaluation: Sectoral Causal Chains', revised May 2017, unpublished, Itad, Hove
- Patton, M.Q. (2011) *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, New York NY: Guilford Press
- Reichertz, J. (2014) 'Induction, Deduction, Abduction', in U. Flick (ed.), *The SAGE Handbook of Qualitative Data Analysis*, London: SAGE Publications
- Sachs, J.D. and McArthur, J.W. (2005) 'The Millennium Project: A Plan for Meeting the Millennium Development Goals', *The Lancet* 365.9456: 347–53
- Schindler, J.; Graef, F. and König, H.J. (2016) 'Participatory Impact Assessment: Bridging the Gap Between Scientists' Theory and Farmers' Practice', *Agricultural Systems* 148 (October): 38–43
- UN Millennium Project (2005) *Investing in Development: A Practical Plan to Achieve the Millennium Development Goals*, London: Earthscan Publications
- Valters, C. (2014) *Six Key Findings on the Use of Theories of Change in International Development*, Justice and Security Research Programme blog, London School of Economics, <http://blogs.lse.ac.uk/jsrp/2014/08/18/six-key-findings-on-the-use-of-theories-of-change-in-international-development/> (accessed 7 August 2018)
- Weiss, C.H. (1997) 'Theory-Based Evaluation: Past, Present, and Future', *New Directions for Evaluation* 76: 41–55
- White, H. (2011) 'Achieving High-Quality Impact Evaluation Design through Mixed Methods: The Case of Infrastructure', *Journal of Development Effectiveness* 3.1: 131–44, <https://doi.org/10.1080/19439342.2010.547588> (accessed 7 August 2018)
- White, H. (2009) *Theory-Based Impact Evaluation: Principles and Practice*, 3ie Working Paper 3, New Delhi: International Initiative for Impact Evaluation (3ie)