



**Evidence in
Practice**

Contents

Stakeholder Map
Timeline
Process Diagram
Appendices
Print

Full Report Summary + Findings

Case Studies

View Alphabetically
View Geographically
View by Discipline

GHANA | ECONOMIC DEVELOPMENT

Graduating the Ultra Poor

A pilot and randomized controlled trial of an innovative poverty alleviation program illustrates the challenges of balancing research and implementation priorities.

By Tory Grieves





**Graduating the Ultra Poor
Table of Contents**

Evidence in Practice Introduction	03
Part I: The Story of GUP	04
The Graduation Approach	05
The CGAP-Ford Foundation Graduation Program	06
Overview of Graduating the Ultra Poor	07
Choosing Ghana	08
Launching GUP	09
Making it Work:Balancing Flexibility and Rigor	10
Sharing Knowledge	11
The End of GUP	13
Results from/Reactions to GUP	14
Escaping Poverty Program, 2016-2019	15
Lost in Translation?	16
Part II: Evidence in Practice: Key Themes + Insights	18
Alternative Definitions of Evidence	18
Operationalizing Evidence for Integration	18
Incentives Structures	19
Definition of Integrators	21
Conclusion	22
Part III: Resources	
Stakeholder Map	23
Timeline	24
Process Diagram	25
Appendices	26

Image Credits

Cover: Susan Quinn/USAID
Page 9: Kasia McCormick/USAID
Page 12: Kasia McCormick/USAID
Page 13: Yooku Ata-Bedu, USAID/Ghana
Page 15: Susan Quinn/USAID
Page 16: USAID/Ghana
Page 20: USAID/A. Kauffeld
Page 22: USAID/A. Kauffeld



The **Evidence in Practice research project at the Yale School of Management**, funded by the William and Flora Hewlett Foundation, was conducted from January 2016 to January 2018 in order to better understand the conditions under which rigorous evidence can be effectively integrated into public policies and non-governmental organization (NGO) practices in the field of international development.

The Evidence in Practice project followed a rigorous methodology comprised of three broad elements: First we conducted an initial round of expert interviews with individuals who have spent a significant portion of their professional lives attempting, researching, or promoting the integration of evidence into development practice, including academics, government officials, foundation program officers, NGO practitioners, and think-tank

directors. Second, we conducted a matched comparison of eight cases of development programs or interventions where rigorous evidence was integrated with varying degrees of effectiveness. This case study is one of the eight produced by the project. The third component, conducted in parallel to the eight case studies, consisted of interviews with prototypical representatives of each of the stakeholder groups, or individuals who could clearly describe the typical experience of enacting a particular stakeholder role. Our synthesis analysis is presented in the accompanying report.¹

Stakeholder Characterization

Based on our research, we have found it useful to think of the flow of evidence into policy and practice as an “ecosystem” in which a set of archetypical stakeholder groups interact. This set of stakeholder categories was described and reinforced by our interviewees throughout the project. While this is not a perfect description (e.g., some organizations fall within more than one stakeholder group and individuals often shift across stakeholder groups or play roles that effectively span categories), it can help frame the conversation to identify the critical roles, incentives, and relationships that animate

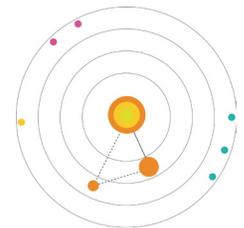
the complex relationship between “evidence” and “practice.” These representative stakeholder groups are: **Researchers, Funders, Influencers², Intermediaries, Policymakers, Implementers, and Beneficiaries³**.

Each of the cases thus contains a map of the specific organizations (and individuals) that defined its evolution, their structural affiliation to a stakeholder category (in some cases, organizations played more than one formal, structural role), the informal roles that certain individual actors played, as well as the key relationships between these individuals and organizations.

¹ Please see the appendix for a detailed description on Data and Methods.

² While some of our interviewees identified “Influencers” (such as the media, the general public, lobbyists, and influential individuals) as playing an important role in the evidence-to-practice eco-system, this group did not play an explicit role in the narrative of any of the case studies. So we have included the category here, though it does not appear in the stakeholder maps of the individual case studies.

³ We use the term “beneficiaries” to indicate those whom a specific policy or program is intended to help. Different analytic frameworks use various terms to describe this group, including clients, users, recipients, etc.



See the relationship between the stakeholders in this project in the stakeholder map.



Part I: **The Story of GUP**

The Graduating the Ultra Poor (GUP) project was a randomized controlled trial (RCT)⁴ and pilot program conducted in northern Ghana between 2010 and 2013. The program sought to further test the “Graduation” approach originally developed by BRAC in Bangladesh through its “Challenging the Frontiers of Poverty Reduction—Targeting the Ultra Poor,” or CFPR-TUP program. The GUP program in Ghana was part of a larger set of ten pilots (and eight associated RCTs) conducted in eight countries between 2006 and 2014 by the Consultative Group to Assist the Poor (CGAP)-Ford Foundation Graduation Program.

In Ghana, the program was designed by Innovations for Poverty Action (IPA), funded by the Ford Foundation and the International Initiative for Impact Research (3ie), and implemented by IPA and Presbyterian Agricultural Services (PAS), a local NGO. IPA has been working in Ghana since 2008, where it operates one of its largest offices worldwide.

A 2015 paper published in *Science*⁵ analyzed six of the RCTs from the CGAP/Ford pilots; Ghana is described as having produced statistically significant, positive outcomes across ten variables, as well as a positive cost-benefit ratio. Success was defined in the paper as the achievement of “the Graduation program’s primary goal, to substantially increase consumption of the very poor...by the

conclusion of the program and [maintain it] 1 year later.”⁶

GUP was effective in terms of learning from, and contributing to, a global body of evidence on poverty alleviation among the ultra-poor, which collectively has led to further adaptation and adoption of the Graduation approach in about 100 countries worldwide. Less clear is whether this “spread” of the Graduation approach is the only appropriate measure of success for the stand-alone pilot in Ghana. GUP was based on a novel economic development program and was designed to generate further, rigorous evidence on its effectiveness, as reflected in the *Science* article. This was the intended purpose of the program and it was handily

achieved through careful implementation and evaluation. But, according to many Ghanaians involved in the program, GUP could have achieved much more in terms of developing capacity building or scaling of the approach on the ground. It is important to note that this was never a priority of IPA in the design or implementation of GUP. Rather, IPA and its partners sought to understand whether the Graduation approach—which was novel and unproven at the time—was an effective approach to poverty alleviation and livelihood development across contexts. There are reasons to believe, however, that achieving local, persistent effects of GUP could have been defined as a priority, which would have

BRAC: Known formerly as the Bangladesh Rehabilitation Assistance Committee, then as the Bangladesh Rural Advancement Committee, and later as Building Resources Across Communities.

⁴ A randomized controlled trial (RCT) is a research design that randomly assigns participants to either a treatment or a control group. The goal is to determine whether the treatment group achieves the hypothesized outcomes of the intervention, using the control group as the counter-factual. “Randomized Control Trials,” Innovations for Poverty Action. [Access here.](#)

⁵ A. Banerjee, E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Pariente, J. Shapiro, B. Thuysbaert, and C. Udry. “Development economics. A multifaceted program causes lasting progress for the very poor: Evidence from six countries,” *Science* 348, no. 6236: 772-790. [Access here.](#)

⁶ *Ibid.*



resulted in a scaling up the program—and its demonstrated impacts—within Ghana.

In other contexts, programs simultaneously generate scientific evidence while creating long-standing results on the ground. As reflected in its original design, IPA's explicit goal was to generate rigorous evidence of a complex intervention. Building institutional capacity and understanding in Ghana for further local implementation of the Graduation approach was never a priority. Yet, for other stakeholders—PAS in particular—the initial expectation certainly was that the program would yield changes in policy and programs on the ground. A shift toward a Graduation approach to poverty alleviation might yet occur in Ghana, as some participants hypothesized that GUP may have been ahead of the government's capacity and priorities at the time.

This case brings into relief how different types of evidence are appropriate for different types of questions. A series of RCTs, for example, is an excellent way to determine if a particular intervention is effective across geographies and cultures, or if it is highly context-dependent. Yet RCTs, on their own, do not integrate other critical evidence about on-the-ground, pragmatic considerations necessary to create a robust and sustainable

“Researchers look for evidence, implementers look for results.”

IMPLEMENTER

program—all of which are elements essential to realizing full translation of evidence into policy and practice.

In the case of the GUP RCT, consideration and generation of quantitative evidence dominated other forms of evidence, partly because of how differences in stakeholder incentives informed the project's design. The case also demonstrates that when this occurs, tensions among stakeholders—who are trying to translate different types of evidence in different ways and for different reasons—are likely to emerge.

Finally, GUP illustrates the mechanisms that hinder or facilitate the integration of relevant evidence into practice. Formal channels were certainly critical, as IPA made a point

of incorporating knowledge transfer through exchange trips and other learning events into the program's design. However, any lasting impacts of GUP in Ghana today were driven by the day-to-day translation conducted informally by self-appointed “champions” of the program—those who advocated and went above-and-beyond for its cause. This finding has significant implications for program design, which must allow and account for the ways in which individuals function as conduits for translation. GUP, therefore, also helps illuminate how champions are created, and the costs of failing to identify, cultivate, and engage certain types of champions at critical junctures.

The Graduation Approach

As noted above, the Graduation approach was pioneered by BRAC in Bangladesh, through its CFPR-TUP program, and its predecessor “Income Generation for the Vulnerable Groups Development Program” (IGVGD). In general terms, the approach blends practices from the fields of social protection, livelihood development, and financial services, to fulfill the multifaceted needs of those in extreme poverty (those living on less than US\$1.25 per day) and graduate them to sustainable livelihoods. While the upfront costs of this methodology tend to be higher than those of other poverty reduction programs, the

According to *The State of the World's Children, 2014*, 29% of Ghanaians live on less than \$1.25 a day.



underlying logic is that it provides the impetus to generate a persistent impact and thus is likely to reduce the longer-term cost of providing services to the ultra-poor.

Thus, the Graduation approach employs a “big push” over a limited period of time (typically 24 to 36 months) to disrupt the poverty cycle. The program provides beneficiaries with a set of services in a defined sequence, as follows:⁷

1. Identify the most vulnerable households within a community;
2. Provide regular and time-bound consumption support to enable them to meet basic needs;
3. Help families plan their livelihoods and then transfer appropriate productive assets to them;
4. Develop their ability to save money as a tool to build resilience;
5. Enhance their technical and entrepreneurial skills through livelihood training;
6. Provide close mentorship to participants throughout the process in a way that develops their capacity and self-confidence.

The CGAP-Ford Foundation Graduation Program

Inspired by the BRAC program’s positive results, CGAP and the Ford Foundation established the CGAP-Ford Foundation Graduation Program and, in 2006, began a series of ten pilots to test the approach’s potential to be adapted and implemented across geographies. Integrated into the pilots, in addition to qualitative research at each site, IPA conducted ten pilot programs and eight RCTs, as seen in Table 1.⁸

As noted above, the Graduation approach has since spread to approximately 100 countries and has been adopted by a number of NGOs and, in approximately 30 countries, by governments. It has also had a significant effect on the social protection programs of governments that were already operating at scale, as is the case in Peru.⁹

⁷ “The Graduation Approach,” UNHCR. [Access here.](#)

⁸ “Graduation Pilots Overview,” CGAP-Ford Foundation Program. [Access here.](#)

⁹ “Graduation Implementations” Microfinance Gateway. [Access here.](#)

Table 1. Randomized Control Trials

2006	Haiti Pilot and qualitative assessment
2007	India Bandhan’s “Targeting the Hardcore Poor Program” – with an RCT Trickle Up’s “Ultra Poor Program” SKS’ “Ultra Poor Program” – with an RCT
	Pakistan Pilot and RCT
2009	Honduras Pilot and RCT
2010	Ethiopia Pilot and RCT
	Peru Pilot and RCT
	Yemen Pilot and RCT
2011	Ghana Pilot and RCT



Overview of Graduating the Ultra Poor

As one of the last Graduation pilots to be implemented, GUP was tailored to test the relative efficacy of various Graduation program elements, as informed by previous pilots.¹⁰ It was comprised of three treatment arms: the full Graduation from Ultra Poor (GUP) arm, the Savings Out of Ultra Poverty (SOUP) arm, and the Asset Only (AO) arm. At the outset of GUP, as with other Graduation pilots, IPA and PAS conducted a participatory wealth ranking exercise, through which villagers ranked the relative wealth of those in their community, establishing a bottom-up selection process for program beneficiaries.¹¹ In general, program beneficiaries were women over the age of 18, who were either the household head, wife of the household head, or daughter of the household head.¹²

The RCT was designed to randomize at the village level, and within treatment villages, eligible households were randomly assigned to treatment or control groups. Of the households that did receive the program, households were randomly assigned to receive either the full GUP treatment (though half of these did not receive the program in order to measure “spillover” effects from non-participating households living nearby), SOUP treatment, or AO treatment.¹³

The full GUP treatment arm was comprised of the following six components:¹⁴

- 1. Consumption Support:** During the lean season (14 out of 24 months), households received weekly cash transfers of US\$2.74 to 4.11, depending on household size.
- 2. Health:** Households were enrolled in the National Health Insurance Scheme and received health and nutrition education.
- 3. Productive Asset Transfer:** A one-time transfer of a productive asset valued at about US\$205.48 was made to each participating household. Forty-four percent of participants chose goats and hens, roughly a quarter picked goats and maize inputs, and a small number picked shea nuts and hens.
- 4. Technical Skills Training:** Training on running a business and managing their chosen livelihood was provided. For example, households who selected livestock were taught how to rear the livestock, including vaccinations, feed and treatment of diseases.
- 5. Savings Account:** Households opened savings accounts and PAS staff collected deposits on a weekly basis.
- 6. Household Visits:** Through the program, PAS staff made weekly visits to each household to provide monitoring, coaching, and encouragement.

¹⁰ The principal investigators were: Abhijit Banerjee (MIT), Dean Karlan (Yale University), Robert Osei (University of Ghana), Bram Thuysbaert (Ghent University), and Christopher Udry (Yale University).

¹¹ “Graduating the Ultra Poor in Ghana,” *Innovations for Poverty Action*. [Access here](#).

¹² A. Banerjee, D. Karlan, R. Osei, B. Thuysbaert, C. Udry, *Graduation from ultra poverty in Ghana, 3ie Grantee Final Report* (New Delhi: International Initiative for Impact Evaluation, 2017). [Access here](#).

¹³ Further description of each treatment arm can be [found in Appendix I](#).

¹⁴ “Graduating the Ultra Poor in Ghana,” *Innovations for Poverty Action*. [Access here](#).



“How do you sell something that we don’t have evidence of? ... We want to be sure that we are not telling people this works, but at the same time you want them to know it exists.”

RESEARCHER

The SOUP arm only received the savings portion of the GUP treatment; a part of this group also received a 50% match of their savings to test whether this would further incentivize saving. The AO arm only received the productive asset transfer portion of the program.

Under the full GUP treatment, average household consumption increased by 11%, while household savings were three times larger relative to the control group. Moreover, despite the high cost of the program, IPA calculated that the program had generated a 133 percent return on investment.¹⁵ Equally helpful, GUP demonstrated that the impacts were far stronger with the entire combination of components. Results from the AO group indicated that these treatment arm alone were not sufficient to reliably increase consumption,

income, or asset holdings.¹⁶ (Results from the SOUP group are not yet available.)

Since the conclusion of GUP, IPA Ghana has embarked on Phase II (2016-2019)—known as “Escaping Poverty,” implemented by Heifer International and funded by the Ford Foundation. This second stage will assess, in a more granular fashion, various aspects of the GUP model to identify which combinations of elements have the greatest impact, allowing for more practical and cost-effective implementation, particularly by governments.

¹⁵ *Ibid.*

¹⁶ A. Banerjee, D. Karlan, R. Osei, B. Thuysbaert, C. Udry, *Graduation from ultra poverty in Ghana, 3ie Grantee Final Report (New Delhi: International Initiative for Impact Evaluation, 2017)*. [Access here.](#)

Choosing Ghana

The decision to implement the Graduation approach in Ghana was primarily based on IPA’s existing presence in the country and the desire, by IPA and its funders, to conduct an RCT in West Africa in order to have a second sub-Saharan African site in addition to Ethiopia. IPA had multiple offices in Ghana, including one in the country’s poorest region, which implied an advantage in program oversight, implementation, and evaluation. However, implementing the Graduation pilot in Ghana also carried challenges. The program lacked both substantial buy-in from the government (in contrast to Ethiopia) and a large NGO with nationwide presence to serve as implementing partner (as in Honduras or Peru) to complement IPA, which would be conducting the RCT evaluation. Moreover, according to several sources, serving as implementer was not a natural role for IPA, which is predominantly a research organization. One IPA staff member commented,

“[Implementing] has never been our first best solution, so we don’t see ourselves as implementers. I guess the way I’d sum up [choosing Ghana] is ‘no regrets,’ because there wasn’t an option that presented itself that was a more convenient idea...that said if at any time



there would be someone with more implementation experience than us who would be willing to deliver the program within the context of having external or rigorous evaluation... that's the best case scenario."

Ultimately, IPA decided to co-implement GUP alongside a relatively small, regional NGO: Presbyterian Agricultural Services. Ghana is the only pilot program for which IPA served a dual role as implementer and evaluator.

Launching GUP

IPA's selection of Presbyterian Agricultural Services as local implementing partner was based on its regional expertise in northern Ghana working with communities that had not previously benefited from NGO programs. Established in 1967, PAS was one of the few organizations working specifically with poor, smallholder farmers of neglected communities in northern Ghana. While its primary expertise was in providing technical agricultural training, PAS also provided financial education and health services. In addition, both IPA's regional office and PAS' operating center were located in Tamale, the capital city of the Northern region.

IPA decided to implement GUP in northern Ghana (in two of three administrative regions



in the north: the Upper East and Northern Regions) for several reasons. Over recent decades in Ghana, various governmental administrations had worked to bring economic stability to the poor through programs such as Economic Recovery Programs (1980s-1990s)¹⁷ and the Ghana Poverty Reduction Strategies of the early 2000s^{18, 19}. More recently, the government had designed programs to target the ultra-poor, such as a cash transfer program known as Livelihood Empowerment Against Poverty (LEAP) and the Local Enterprises & Skills Development Program (LESDP).²⁰ While such efforts helped Ghana successfully achieve

According to the World Bank, Ghana's economic performance improved in the first half of 2017, after substantial fiscal slippage in 2016. The fiscal deficit for the first half of 2017 was 2.7% of GDP—on track to meet its target of 3.5% of GDP.

¹⁷ In 1983, Ghana launched an economic recovery program (ERP) in order to reverse a period of serious economic decline characterized by lax financial management and high inflation rates over 100 percent.

¹⁸ The Ghana Poverty Reduction Strategy (GPRS I, 2003-2005) and the Growth and Poverty Reduction Strategy (GPRS II, 2006-2009) were initially developed as a condition for development assistance under an IMF-World Bank debt relief initiative in 2002. The Strategy sought to achieve macroeconomic stability and reduce poverty in the country.

¹⁹ A. Banerjee, D. Karlan, R. Osei, B. Thuysbaert, C. Udry, *Graduation from ultra poverty in Ghana*, 31e Grantee Final Report (New Delhi: International Initiative for Impact Evaluation, 2017). [Access here](#).

²⁰ *Ibid*.



some of the Millennium Development Goals,²¹ the three northern administrative regions still “harbor the poorest of the poor.”²² Indeed, 53% of the households included in the GUP study were living on less than \$1.25 per day at the outset of GUP, compared with 29% in the country as a whole.²³

The program was officially launched in Tamale through a series of events attended by the program’s implementers including IPA Ghana and PAS staff; district representatives of the Ministry of Gender, Children and Social Protection; the Ministry of Agriculture; and the Chief Executive of the Northern district.

Making it Work: Balancing Flexibility and Rigor

IPA’s decision to evaluate *and* co-implement the project in Ghana was one of the single most influential decisions of the GUP project. IPA was responsible for the project design and technical oversight, as well as budget administration. The majority of operational implementation responsibility was delegated to PAS, for procuring assets and delivery of services. An implementation coordinator from IPA was responsible for guiding PAS on technical components of the project.

PAS staff, known as team leaders, worked closely with IPA supervisors to carry out daily

“[Those outside the project] got to know that the ultra poor, all is not lost for them. If you’re there to guide them they will be able to do something beneficial for themselves.”

IMPLEMENTER

program activities through the program’s 25 field agents, who were responsible for visiting each beneficiary community on a weekly basis.

IPA’s role as both implementer and evaluator of GUP became particularly challenging on a day-to-day basis. As an implementer, IPA’s close work with PAS rendered it sympathetic and, to a large extent, responsible for responding to unforeseen issues that arose during implementation, which often entailed diverging from the principal investigators’ original program design. But as evaluator, IPA was committed to ensuring that GUP strictly

adhered to the project design so that it would fulfill the goal of the RCT: to generate rigorous evidence to contribute to a global body of knowledge. The tension between these two roles was evident from the onset, as IPA saw the need to actively intervene to support the limited capacity of its relatively small implementing partner. In the second year of the project, IPA brought on another internal monitoring agent because it felt that too many deviations from the project design were occurring.

²¹ The eight Millennium Development Goals (MDGs) were established at the Millennium Summit of the United Nations (UN) in 2000 with the goal of being achieved by 2015. The eight goals are: 1) to eradicate extreme poverty and hunger, 2) to achieve universal primary education, 3) to promote gender equality and empower women, 4) to reduce child mortality, 5) to improve maternal health, 6) to combat HIV/AIDS, malaria and other diseases, 7) to ensure environmental sustainability, and 8) to develop a global partnership for development.

²² A. Banerjee, D. Karlan, R. Osei, B. Thuysbaert, C. Udry, *Graduation from ultra poverty in Ghana*, 3ie Grantee Final Report (New Delhi: International Initiative for Impact Evaluation, 2017). [Access here.](#)

²³ “*Graduating the Ultra Poor in Ghana*,” Innovations for Poverty Action. [Access here.](#)



“GUP has and will continue to have an impact, because you realize that with some policymakers, even though you’ll not see translation into policy immediately, they incorporate it in their plans.”

IMPLEMENTER

Flexibility

Despite the program’s necessarily rigid design as an RCT, IPA made several decisions throughout the project to adapt it to conditions on the ground. For example, it delegated the responsibility of hiring field agents to PAS, rather than selecting the field agents itself—as defined by the research protocol. This was significant not only because it demonstrated an appreciation of PAS’s local expertise, but also because it helped ensure that project staff working most directly with PAS fit within the culture of the NGO, which has a faith-based history and vision. IPA also shifted the responsibility of assessing livelihood options from a consultant

it had hired to PAS, which had been wary that the consultant lacked an understanding of the local context.

Finally, PAS stressed the importance and influence of village chiefs, so IPA agreed to engage them early on to explain the characteristics, broader benefits, and potential of the program to garner buy-in—particularly for the participatory wealth ranking system. Harnessing PAS’ knowledge of the local context in this way, IPA avoided potential conflicts such as asset theft that might otherwise have occurred.

Rigor

Though a frequent source of friction between the co-implementers, the intense interaction between IPA and PAS also created a positive force for the project to work. IPA staff visited the PAS office every week to oversee implementation and help troubleshoot emerging issues. The researchers also visited field sites in Ghana each quarter to better understand the context on the ground. This intense engagement by IPA staff with implementation in the field, formalized through weekly meetings between IPA and PAS, was a key strength of the program.

One the other hand, the rigorous protocols, designed to ensure that the RCT would yield its intended evidence, constrained the program’s flexibility to adapt to local conditions. Ultimately, the priority was to generate rigorous evidence, and not for the GUP pilot to translate into a sustained program on the ground.

Sharing Knowledge

From the global Graduation pilot program’s inception, CGAP and the Ford Foundation defined knowledge sharing and outreach key tenets. In addition to the quantitative and qualitative research, CGAP and the Ford Foundation created a Community of Practice to bring together implementers across the

A randomized controlled trial (RCT) is a research design that randomly assigns participants to either a treatment or a control group. The goal is to determine whether the treatment group achieves the hypothesized outcomes of the intervention, using the control group as the counter-factual. [More here.](#)



ten pilot sites to share experiences and lessons among themselves and with other stakeholders. The Community of Practice also offered exchange visits, both to BRAC and among pilot sites, as well as ongoing technical assistance, as well as an annual convening of implementing organizations with international NGOs, donors, policymakers, and academic partners interested in the Graduation approach.

At first, the annual convenings were mainly a forum for participants to share their experiences implementing the Graduation pilots in various contexts. Over time, as CGAP and Ford Foundation expanded their focus to scaling the Graduation approach, the convenings included more policymakers, donors, international NGOs, and other stakeholders that were critical to the spread and scale-up of Graduation programs around the world.

In line with this, GUP included funding for its partners to attend the annual knowledge-sharing events, visit pilots in other countries to learn of their experiences, and for quarterly trips to Ghana by the principal investigators (in addition to monthly, detailed reports sent to them from the field).

Learning visits led to a stronger buy-in among PAS staff that, coupled with a greater depth of understanding, turned several of them into program champions. Two such champions participated in a learning visit to Bangladesh at the outset of the GUP project. As one noted,

“The study tour to Bangladesh to learn, before coming to really be leading a big project like that, was good. Then secondly, the [participatory wealth ranking], that is where we saw it, Bangladesh. It was a good idea there. Now we brought it and implemented, and it was super. How to select the eligible beneficiaries is one of the biggest parts of the projects. That was done and done successfully, and it was also as a result of the trip, but that is where we learnt it.”

Learning visits also allowed implementers to learn more about the Graduation approach and the results emerging from other Graduation pilots. In particular, visiting other pilots often validated IPA's rigorous approach to evidence collection: “There is a lot of difference in what people do and what they consider as impact. It was beneficial for my work and also for me in person, and reassuring that maybe IPA is doing a thorough job.”



Sharing actionable, implementation-centered knowledge about the Graduation approach was a priority for CGAP to facilitate the spread and scale-up of the approach. For example, in 2011 CGAP published an instructive guide on how to target ultra-poor communities entitled “Reaching the Poorest: Lessons from the Graduation Model.”²⁴ In 2014, CGAP and Ford Foundation published a technical guide to implementing Graduation programs in various contexts, entitled “From Extreme Poverty to Sustainable Livelihoods,”²⁵ which is “intended primarily for those with direct oversight responsibility for running Graduation Programs” and to “be useful for policymakers, technical assistance providers, researchers, and others interested in approaches to address extreme poverty.”²⁶

The End of GUP

GUP was conceptualized and communicated as a time-bound intervention. As the end of the GUP program neared in 2013, the pilot had served its intended purpose—to contribute to the global body of knowledge on the Graduation approach. However, others who had been involved—particularly PAS and program beneficiaries—were reluctant to see it come to a close. Overall, beneficiaries of GUP were better off after participating in the pilot program, as reported in the 2015 *Science* article. Implementers who had witnessed progress on the ground throughout



the two-year implementation thus wondered how the results could be built upon and why the positive evidence generated by the pilot was not sufficient to lead to scaling up the program in the country.

In line with the program’s commitment to sharing results, IPA held workshops in both Tamale and Accra, where many attendees were particularly surprised by, and excited

about, the savings portion of the GUP pilot. The program had shown that even the ultra-poor were capable of saving significant amounts of money, if provided access to services along with appropriate guidance.

This motivated some stakeholders to continue elements of the program after it concluded. The Presbyterian Cooperative Credit Union (PCCU), which had participated in the implementation of GUP by creating bank accounts where it deposited beneficiaries’ savings, discovered that small individual savings summed up to a relatively large market opportunity. When the program ended, PCCU continued to send officers to collect community savings in some beneficiary villages. In other communities that were not served by banks, PAS helped to establish a savings program for the community, especially women and the ultra-poor.

CGAP: The Consultative Group to Assist the Poor is a global partnership of more than 30 leading organizations that seek to advance financial inclusion.

²⁴ S. Hashemi and A. de Montesquiou, “Reaching the Poorest: Lessons from the Graduation Model,” CGAP, March 2011. [Access here.](#)

²⁵ A. de Montesquiou and T. Sheldon, “From Extreme Poverty to Sustainable Livelihoods: A Technical Guide to the Graduation Approach,” CGAP, September 2014. [Access here.](#)

²⁶ The graduation work has since been transferred from CGAP to the Partnership for Economic Inclusion (PEI) at the World Bank.



“In randomized control trials it’s all beautiful. You do this. You do that. Five years later, you have a draft paper. It’s way outside the timeline of the government and it was going to cost way more than the government would be able to afford, because all these treatments, you need sufficient sample size ...”

FUNDER

Results from/Reactions to GUP

GUP yielded an impressive 133% percent return on investment.²⁷ Total monthly consumption increased by 11%, households experienced a 91% increase in non-farm income, and women experienced significant gains in empowerment, sustained one year after the program’s conclusion.²⁸

One of the greatest champions of GUP, particularly toward the end of the program, was PAS, who felt that the Graduation

approach had definitively helped the ultra-poor and was convinced that the results had indeed been very positive. As a PAS implementer noted,

“...when you are doing something and you know that you are doing the right thing, that what you are doing is actually changing lives, then you’ll surely know, and I know what we did was actually what people actually wanted.”

Despite these positive results, GUP did not garner significant buy-in among policymakers, donors, and NGOs within Ghana in part because IPA had not cultivated strong relationships with them throughout program design and implementation. As one funder noted, “Unlike other pilots where CGAP-Ford played a more leading role, the emphasis in Ghana on policy advocacy was absent.” On top of this, at the results-sharing workshop in Accra, some participants reported that the results from GUP were not presented in as compelling a way as they might have been (for example, by comparing GUP to other Ghanaian government programs).

At the event, IPA emphasized—as it had from the program’s outset—that GUP was part of a larger initiative consisting of several pilots. Thus, IPA presented GUP’s results on a series of graphs that included those from the other pilots. For example, even though a 133% return had been unheard of in the development sector in Ghana, the result was overshadowed by other programs that achieved higher returns: 433% in India, 260% in Ethiopia, 179% in Pakistan, 146% in Peru,

²⁷ “Graduating the Ultra Poor in Ghana,” Innovations for Poverty Action. [Access here.](#)

²⁸ *Ibid.*



(and -198% in Honduras).²⁹ IPA wanted to communicate the consistency of impressively good results. In reality, despite its success, GUP appeared as the “second worst” pilot in all measures. As one funder recounted, “You can see the bar chart in your head, right?... In every GUP outcome, [my reaction] was like...‘huh.’”

Partly as a result of this contrast, donors and policymakers at the workshops, who were impressed by GUP’s results and accepted their validity, nonetheless questioned whether such results were worth the “relatively high cost” of the program relative to existing interventions such as LEAP. Others expressed doubt regarding the project’s design, postulating that adjustments to certain program elements—like offering petty trade as an alternative to goat rearing as a livelihood option—would have generated results closer to those reached in other countries. It could be argued that GUP’s results and its positive benefit-cost ratio alone should have motivated key stakeholders in Ghana to scale up the program. Yet, failure to engage



these stakeholders from the beginning as co-designers, as important voices throughout project implementation, or at least as an explicit intended audience with specific interests and needs when communicating results, affected their ultimate reactions to the program’s results, and unnecessarily constrained GUP’s impact to its contribution of evidence to the broader Graduation Program.

Escaping Poverty Program, 2016-2019

Following the GUP pilot, IPA has embarked on a second, modified pilot of the Graduation model in order to test various elements of the Graduation approach in more detail—known as “Escaping Poverty.” This program will be implemented through 2019, with Heifer International as the lead implementing partner.

The design of this second program reflects IPA’s recognition of factors that could have more positively contributed to the translation of the evidence generated by GUP into policy and practice within Ghana. In particular, Escaping Poverty has given higher priority to policymaker engagement and, most significantly, to testing a high-impact, lower-cost approach to Graduation:

“It was only towards the end, as we were trying to share our findings and gauge interest from donors and the government that we realized, ‘Oh, this is very expensive’ and yeah, we were trying to integrate it with a government program and we had a lot of discussions with some of the donors and government.”

Through Escaping Poverty, IPA is testing a more active approach to policymaker engagement. The program is directly involving district assemblies and various government

²⁹ A. Banerjee, E Duflo, N. Goldberg, D. Karlan, R. Osei, W. Pariente, J. Shapiro, B. Thuysbaert, and C. Udry. “Development economics. A multifaceted program causes lasting progress for the very poor: Evidence from six countries,” *Science* 348, no. 6236: 772-790. [Access here.](#)



departments as thought partners in the design of the program. For example, IPA is incorporating policymaker perspectives on community selection and livelihood options. Simultaneously, the pilot program will also serve as a means to test Heifer International's community model (a theory of change centered on the community as a driver of positive outcomes) against the Graduation model's household-based approach to poverty alleviation.

The design of this second pilot illustrates the extent to which IPA recognized the opportunity for evidence translation that was missed during the first phase of GUP because policymakers were not more deeply engaged in program design.

Lost in Translation?

As part of a multi-country, multi-year pilot program and set of RCTs, GUP contributed to a global body of evidence that lends credence to the Graduation approach and has fostered its spread to myriad countries. At the same time, it did not influence practices *on the ground* in Ghana through replication or scaling up of the program by government entities or other development actors. It may be more accurate to say that this has not *yet* occurred, as some believe that GUP might have been ahead of its time for the government of Ghana.

This raises questions about the temporal scale of integrating evidence into practice. GUP was conceptualized as a pilot to test the efficacy of certain elements of the Graduation approach. IPA thus never explicitly sought to pursue large-scale implementation of GUP in Ghana. When GUP was initiated, it was not clear that the program would be effective (that was the purpose of the multi-country RCTs). Planning for scale-up at the program's outset did not necessarily make sense.

At the same time, CGAP, the Ford Foundation, and implementing parties in other countries invested much more time and energy to engage policymakers before, during, and after pilot implementation in order to lay



the groundwork for in-country scale up. In Ethiopia, for example, where the pilot and RCT were implemented from 2010-2012, CGAP and the Ford Foundation began discussing—and, in one implementer’s words, “negotiating”—with policymakers the ways in which they could institutionalize the Graduation approach in government programs from 2003—seven years before the pilot was launched. Even so, it has taken several more years since the conclusion of the pilot program for the government to include Graduation as part of its Productive Safety Net Programme (PSNP), illustrating how much time and planning can be required, even among motivated and engaged policymakers, to integrate evidence into practice. From 2011-2013, approximately 11% of PNSP’s nearly 8 million beneficiaries have been included in activities modeled after the Graduation pilot.³⁰

Despite this lack of early engagement with key stakeholders in Ghana, one of the lasting legacies of GUP has been a greater awareness of the broader concept of productive inclusion among decision makers in the country:

“I think the Ministry still felt like at that time ... this was way further down the field than they were, but what it did do is I think it planted the word ‘Graduation’ in their mind.”

“These are my people. I identify with them. I know their pain. You can come here with all the religious intent to help, but I would help them because they are my kind. I will work with them and I believe that by the time I would finish with this project, they will not be the same.”

IMPLEMENTER

³⁰ Matt Hobson, “Pathways to Graduation: A Work in Progress in Ethiopia,” World Bank, December 2014. [Access here.](#)



Part II: Key Themes + Insights

This section discusses the Evidence in Practice themes as they pertain to GUP and summarizes key insights and implications for thinking about the translation of evidence to policy and practice more generally.

Alternative Definitions of Evidence

There are varying definitions and understandings of what constitutes “evidence,” dependent especially on the perspectives of each stakeholder group. For example, the framing, language, and limited accessibility of academic evidence can render it less useful to other stakeholders. These diverging views of evidence create barriers across stakeholder groups, as what constitutes valid evidence for each exists in different realms and in different forms that are challenging to reconcile.

Dispute over the definition of what should be considered ‘evidence’ is central to the story of GUP. Implementers and researchers held different views of the validity of visible or preliminary program results versus data collected per RCT protocols. For example, field agents responsible for visiting the communities each week saw that beneficiaries were in fact saving and that their livelihoods were more sustainable as a result of GUP. Based on this experiential evidence, PAS wanted to scale up the program in Ghana

rather than continue piloting. From IPA’s perspective, in contrast, the two-year pilot was meant to provide scientific evidence to a global body of knowledge on the Graduation approach, so any observations that emerged within that two-year period could not be considered as reliable evidence.

Throughout GUP, PAS’s and IPA’s understandings of evidence differed based on each organization’s goals and purpose. PAS sought to reduce local poverty, IPA sought to gather evidence on the Graduation approach in a second African country. Because IPA was the ultimate decision-maker during GUP, it prioritized *scientific* evidence. Even though IPA was heavily involved in program implementation, it had no mandate to integrate the pilot into a scaled-up program on the ground.

In contrast, implementers *knew* that the project was effective based on positive results that they observed in real time: “These are my people. I identify with them. I know

their pain... I would help them because they are my kind. I will work with them and I believe that by the time I finish with this project, they will not be the same.” While PAS had the motivation and incentives for in-country scaling, it lacked the decision-making power, the mandate, or the resources to lay the necessary groundwork among policymakers, funders, and other key stakeholders.

Operationalizing Evidence for Integration

Even organizations with strong monitoring and evaluation departments often do not transform the operational data into formats that could be widely used within the organization, or beyond, to expand actors’ understanding about what has been learned from past or existing programs. Data is thus used to evaluate retrospective operations, but not to improve the prospective design of new initiatives. This inhibits the application of experiential evidence, which may be rigorous and convincing, to new contexts and often prevents evidence from reaching other key stakeholders after it is produced, as it remains linked internally only to a given initiative.

One of the key strengths of the GUP pilot was the extent to which its operations benefited from the global CGAP-Ford Foundation program



and the evidence and experience from other Graduation pilots. CGAP and Ford Foundation, as key funders of GUP, demonstrated how a funder—in understanding the value of evidence—can build a robust learning agenda into a program, including quantitative and qualitative evaluations, experiential evidence, and periodic mutual learning opportunities. Practices such as annual gatherings of the community of practice, weaved evidence into the fiber of GUP, yielding enormous benefits for implementers and researchers.

In contrast, IPA was so focused on its specific mandate of generating scientific evidence through a rigorous RCT that it missed the opportunity to help PAS and policymakers codify GUP's evidence for scale-up within Ghana. IPA was almost exclusively focused on cultivating evidence to support the spread of the Graduation approach to other contexts via the contribution of the GUP RCT to the growing body of quantitative evidence. Conversely, PAS, whose structure had been influenced toward evidence-based implementation as a result of working with IPA, was almost entirely focused on rigorously implementing the program within Ghana. By focusing on the different components of integration, IPA and PAS were focused on operationalizing evidence in different ways. Had IPA recognized the value of the second type of evidence earlier on, it could have more

“It planted the word graduation in their mind, because I felt like up until then the word graduation just didn’t come up, and I think for good reason.”

FUNDER

proactively supported and leveraged PAS' strengths, relationships, and legitimacy to add the focus on in-country scale-up as a second layer to the project.

Incentives Structures

Throughout the ecosystem, within and across stakeholder groups, formal and informal incentive structures are frequently not conducive – and are often in contradiction – to the integration of evidence into practice. Typically, organizational incentives are defined around an insular view of the organization (e.g., academics publish in academic journals, policymakers must exercise their budgets according to program and budgetary rules,

NGOs must operationalize their programs as stated in their budgets and proposals to funders). Usually, these organizational incentives have no mandate or room for the explicit search of external evidence, much less for the generation of internal evidence that would then lead to continuous adaptation of programs and policies as new learning emerges.

While IPA was focused on a rigorous evaluation of the Graduation model, PAS was dedicated to helping communities in northern Ghana and continuing to build its reputation as a committed NGO. These differing perspectives are not necessarily contradictory, but they nonetheless inhibited the integration of GUP into policy and practice in Ghana. In particular, not engaging policymakers upfront in a deep or sustained way unnecessarily boxed GUP into remaining as an isolated pilot project with no focus on its potential for scaling up:

“I think to add the whole idea of academic versus governance...there was a big issue of you were going to do this just for a study and then what? We know for sure we won't be able to do it right after. It's almost like you tease these people, introduce them to something that they can never have and then you just drop them. That was also something that I



think operationally didn't make sense for the role of the ministry.”

The challenge for IPA was that, because policymakers were not engaged as partners in problem definition or program design, “selling” GUP before final results were observed was unacceptable, as initial or intermediate assessments were not reliable enough:

“How do you sell something that we don't have evidence of? ... [at] the midline [assessment]...we want to be sure that we are not telling people this works, but at the same time you want them to know it exists. For the policymaker, like if you don't have the evidence, he has no time for you.”

IPA's commitment to rigorous research means that it is appropriately conservative about discussing the efficacy of a program until results are analyzed. Similarly, policymakers are wary of investing political capital in an approach or program that does not have institutional legitimacy or strong evidence of success. It does not help that different stakeholders operate on radically different timelines and budgets considerations:



“In an RCT it's all beautiful. You do this. You do that. Five years later, you have a draft paper. It's way outside the timeline of the government and it was going to cost way more than the government would be able to afford.”

One funder expressed disbelief that the paper published in *Science* had not immediately influenced policymaking within the Ministry

of Gender, Children, and Social Inclusion, which reveals the expectation that academic evidence would be translated directly into policy *by policymakers themselves*. Naturally, the policymakers interviewed for this case study instead mentioned looking to other programs implemented within Ghana and other countries as key forms of evidence, and none mentioned a single academic paper as a point of reference. Furthermore,



with government staff turnover occurring frequently in Ghana, policymakers that may have been interested in the program in 2010 were likely no longer in decision-making positions in 2015. As described by a funder, “It takes quite a bit of time to get [policymakers] excited.”

These dynamics are pervasive and present real barriers. But they also constitute a false dichotomy. Had GUP engaged policymakers from the beginning, there would be no need to “sell” results, as the overall question and the interest in its potential would be shared throughout the process. GUP thus demonstrates the importance of establishing relationships and co-creating managerial and financial structures with partners that have the potential to carry forward activities into the future; otherwise, as GUP illustrates, a program may become unnecessarily rigid, inhibiting key stakeholder involvement and, in turn, the integration of evidence.

Definition of Integrators

GUP illustrates the ways in which champions—or individuals with particular enthusiasm for the project/initiative—can play a major role in the *integration* of evidence into practice. Throughout GUP, champions emerged among implementers and funders, such as an implementer who took the initiative to

“I was expecting that result, because when you are doing something and you know, yes, you are doing the right thing, what you are doing is actually changing lives ... ”

IMPLEMENTER

personally visit local government offices to explore potential intersections between local development programs and GUP.

Individuals who straddle stakeholder categories and thereafter create bridges between them can also become natural champions for integration of evidence and practice. When a PAS field agent became employed by the Presbyterian Cooperative Credit Union, for example, he helped design the financial product that enabled savings to continue after the formal end of GUP. As he described it, “We now don’t even see them as GUP communities, we see them as our members.” A blending of perspectives, goals, and incentives allowed this champion to adopt a dual perspective through which

broader project outcomes were prioritized over more actor-specific results. Another project implementer, for example, was particularly convinced that the savings portion of GUP was instrumental for villagers to develop more sustainable livelihoods. After GUP, she adapted its savings model to establish village saving loan associations (VSLA) for women in several villages. She even used the word ‘graduated’ to describe the progress achieved by an estimated 350 women who were part of these associations in 2015.



Conclusion

GUP is a particularly interesting case to examine through the lens of the translation of evidence into practice because of its complexities; while evidence on the Graduation approach both fed into and grew out of GUP, translation in terms of what would have been needed for seeding in-country scaling was not fully realized.

Rightfully, the Graduation model has been heralded as an important contribution to exploring effective approaches to poverty alleviation around the world. Graduation pilots like GUP have made critical contributions to a global body of evidence about the Graduation approach, which has been integrated into policies and practices on the ground by dozens of NGOs and governments. This successful adaptation of the model across the world has been facilitated by several core tenets of the graduation approach, most notably weaving knowledge-sharing into the fiber of the Graduation Program. Through a funding structure that provided ample opportunities for knowledge sharing events, exposure visits, and rigorous quantitative and qualitative evaluation, GUP and its sibling pilots were designed to promote true integration of evidence across the program. At the same time, IPA has recognized the tradeoffs implicit in the design and



implementation of GUP and is accordingly testing new approaches through Escaping Poverty The design of this second phase reflects awareness of how to facilitate government engagement and in-country scale-up of the Graduation approach. The most critical way in which IPA, CGAP, and Ford Foundation did not design GUP

for greater translation and integration of evidence into practice was by not engaging policymakers as official partners and co-creators early on in the program's design and throughout implementation. This would have allowed policymakers greater understanding of the program's design and cost structure, its ambitious goals, and most importantly, could have incorporated elements that would have rendered the program more tailored to scale-up in Ghana. By engaging the government as a co-creator in the program, IPA could also have strategically allocated "credit" for some elements of program implementation and successes to the government in order to satisfy policymakers' incentives while still achieving its goals.

The design and ongoing implementation of Escaping Poverty, focused on developing a more cost-efficient Graduation program and engaging policymakers early on, is a positive sign that simultaneous achievement of local and global integration of evidence may yet be achieved in Ghana.

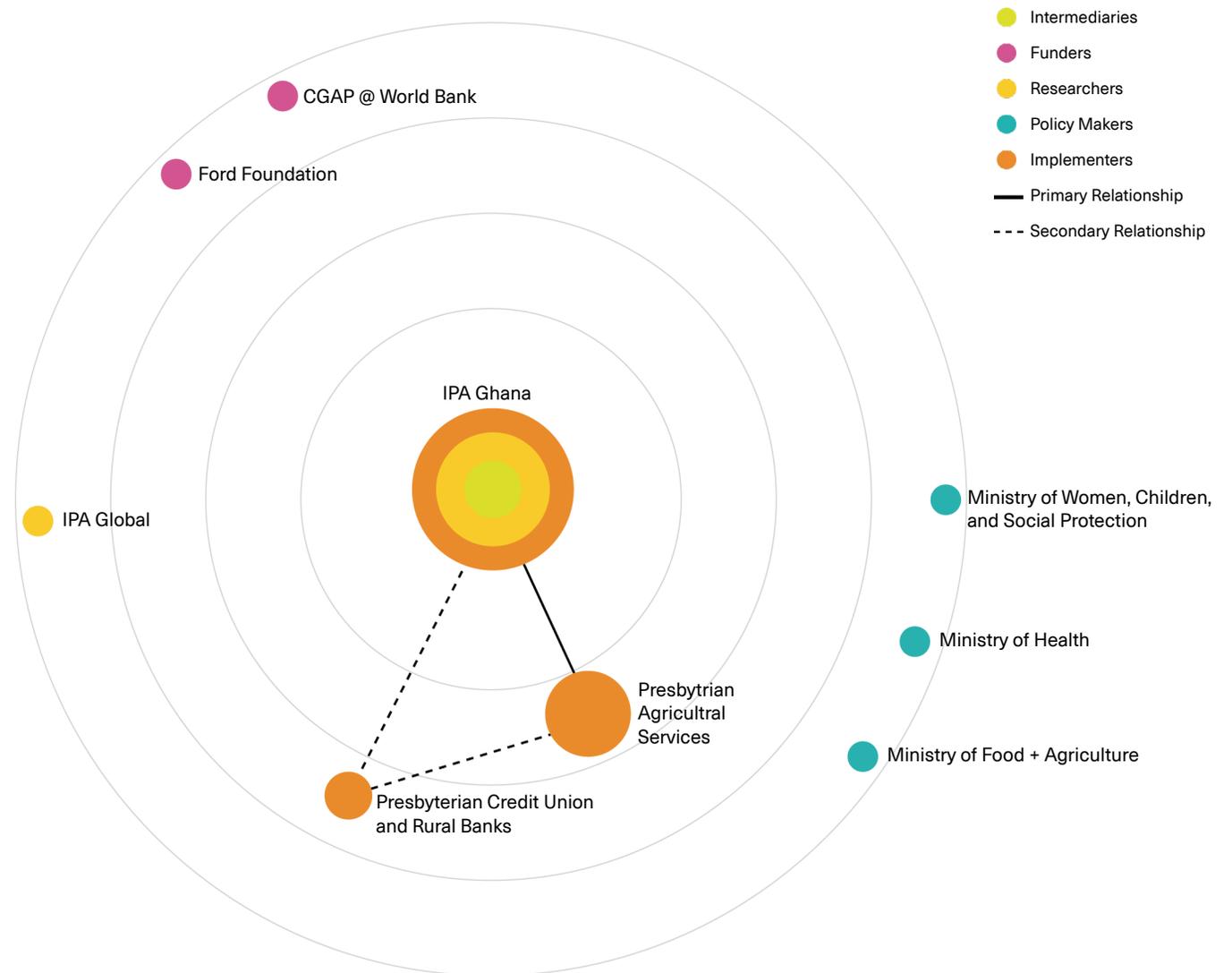


- Stakeholder Map
- Timeline
- Process Diagram
- Appendices
- Print

- View Alphabetically
- View Geographically
- View by Discipline

Stakeholder Map

This stakeholder map is a visual representation of the major stakeholders involved with this project. The importance of each of the actors is defined by their relative size, and their proximity to the center of the project. Their role is defined by the color; multiple colors indicate multiple roles. Primary relationships, denoted by solid lines, indicate the most directly significant relationships while secondary relationships, denoted by dashed lines, indicate indirect, but influential relationships. Actors not connected by lines are still involved with the project, but less directly.

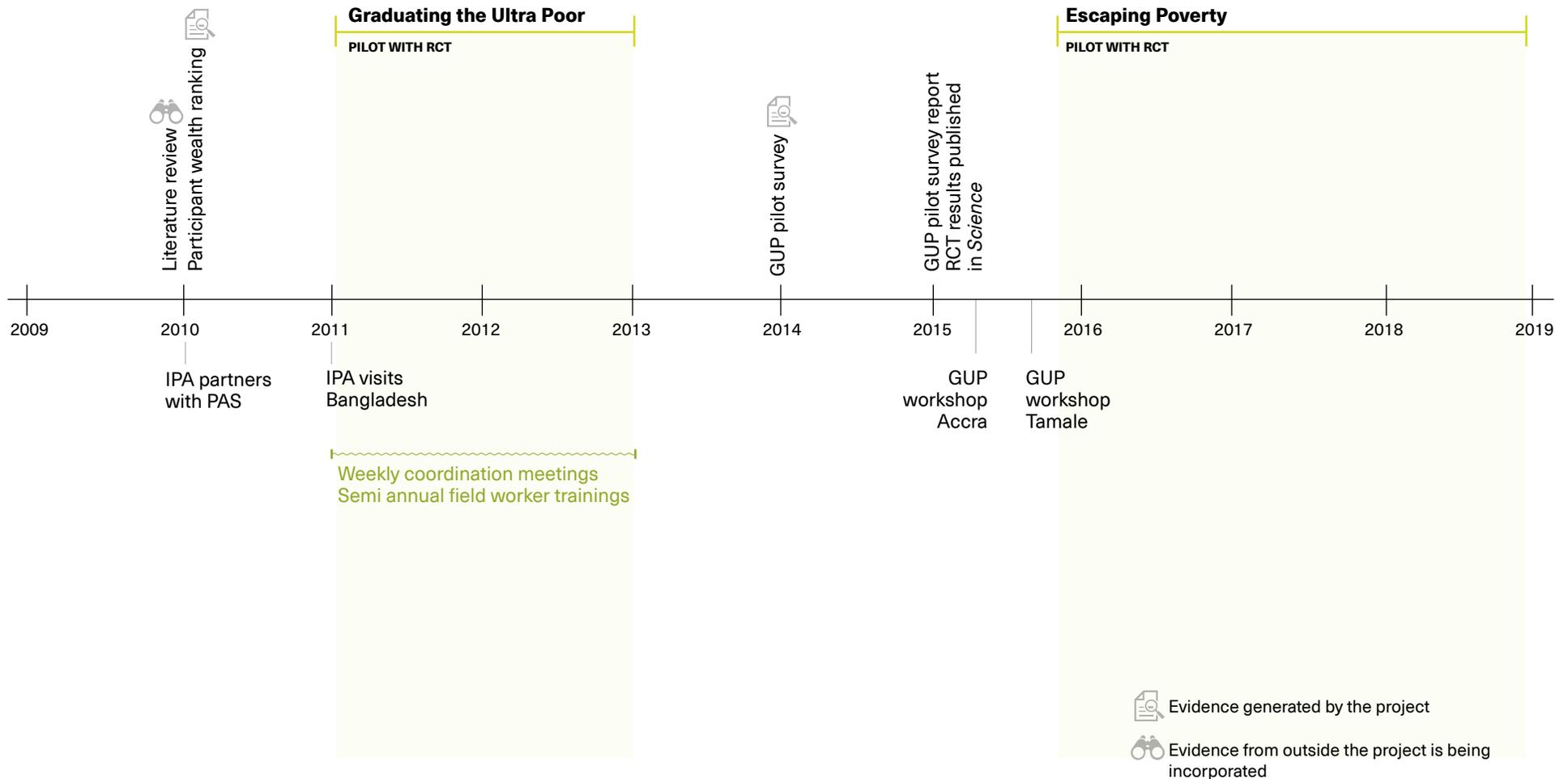




- Stakeholder Map
- Timeline
- Process Diagram
- Appendices
- Print

- View Alphabetically
- View Geographically
- View by Discipline

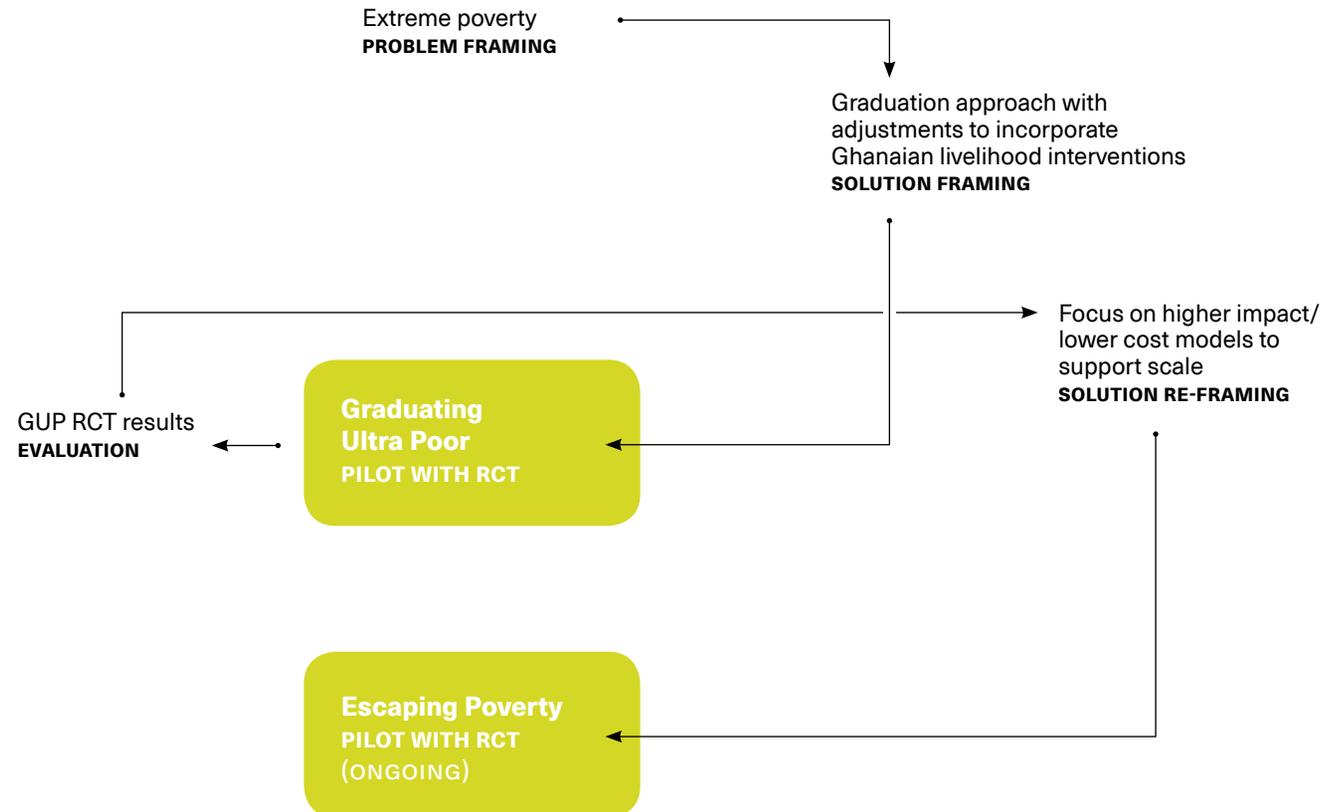
Timeline





Process Diagram

Ongoing implementation of Escaping Poverty seeks to enhance potential for in-country **scale-up** of the Graduation model, which has successfully **spread** internationally.





Stakeholder Map
Timeline
Process Diagram
Appendices
Print

View Alphabetically
View Geographically
View by Discipline

Appendix 1

The full GUP treatment arm was comprised of the following six components:³¹

1. Consumption support: During the lean season (14 out of 24 months), households received weekly cash transfers.
2. Health: all GUP clients and three dependents each were registered on the National Health Insurance Scheme (NHIS) for the first two years of the program
3. Productive asset transfer: One-time transfer of a productive asset, such as goats or guinea fowl.
4. Technical skills training: Training on running a business and managing their chosen livelihood. For example, households who selected livestock were taught how to rear the livestock, including vaccinations, feed and treatment of diseases.
5. Savings account: All treatment households were encouraged to save. GUP helped all households open savings accounts and field agents collected the deposits from these households on a weekly basis.
6. Household visits: Weekly visits by PAS staff to provide monitoring, coaching, and encouragement.

The SOUP arm only received the savings portion of the GUP treatment; a part of this group also received a 50% match of their savings to test whether this would further incentivize saving. The AO arm only received the productive asset transfer portion of the program.

³¹ <http://www.poverty-action.org/study/graduating-ultra-poor-ghana>



Contents

Stakeholder Map
Timeline
Process Diagram
Appendices
Print

**Full Report
Summary + Findings**

Case Studies

View Alphabetically
View Geographically
View by Discipline

Appendix: Resources

Selected, Published Background Research

Banerjee, Abhijit, et al. "A multifaceted program causes lasting progress for the very poor: Evidence from six countries." *Science* 348.6236 (2015): 1260799.

"Can the Graduation Approach Help to End Extreme Poverty?" *Yale Insights*. February 7, 2017. <http://insights.som.yale.edu/insights/can-the-graduation-approach-help-to-end-extreme-poverty>

Hashemi, Syed, and Aude De Montesquiou. "Reaching the poorest: Lessons from the graduation model." *focus note* 69 (2011): 1-15.

"Graduation from ultra poverty in Ghana: Final grantee report." *International Initiative for Impact Evaluation*. August 2017. http://www.3ieimpact.org/media/filer_public/2017/08/22/gfr-ow2206-ghana-ultra-poor.pdf



Appendix: Data + Methods

The research design for the Evidence in Practice project consisted of three broad components. First, we conducted expert interviews (31) with individuals who had spent a significant portion of their professional lives attempting, researching, or promoting the integration of evidence into development practice.^{A1} This included academics, government officials, foundation program officers, NGO practitioners, and think-tank directors. To identify these experts, we first contacted individuals who had either published extensively and prominently on the topic or who had actively funded research or programs with the explicit goal of integrating evidence into practice. From this first set of experts we conducted snowball sampling until we reached a saturation point.^{A2} This initial set of interviews informed and directed the next two components, as they resulted in an initial map of the relevant stakeholders in the “evidence-to-practice ecosystem” and the hypothesized and actual paths that seemed to link them together.

Second, we conducted a matched comparison of eight cases of development programs or interventions where rigorous evidence was integrated with varying degrees of effectiveness. These cases were matched on structural, geographic, and programmatic characteristics—as well as on the extent to which evidence had informed practices—to better identify the critical factors that allowed actors in certain cases, and not others, to integrate rigorous evidence into practice.^{A3} This matching process led us to identify pairs of cases across four different countries, leveraging temporal and cross-sectional variation between them as seen in table A2.

A1 By development practice, we mean the work of government actors, NGOs, and others who are responsible for designing and executing development projects and programs.

A2 Data saturation is difficult to define and is dependent on the field of study. In this case, we defined saturation as the moment when, in a sequence of several expert interviews, no interviewee gave us information that we had not encountered before.

A3 George, A. L., & Bennett, A. (2005). Case studies and theory development in the social sciences. MIT Press, Cambridge, MA. Chapter 5.



- Stakeholder Map
- Timeline
- Process Diagram
- Appendices
- Print

- View Alphabetically
- View Geographically
- View by Discipline

Table A1. Expert Interviews

	Researchers	Funders	Intermediaries	Policymakers	Implementers	Total
Phase 1: February 2015 – May 2016	10	8	6	2	5	31
Phase 2: September 2016 – June 2017	10	7	6	3	8	34

For each case, we first identified, through existing literature and interviews with subject experts, a series of key informants who had detailed knowledge of the case’s history and protagonists. These initial interviews with case experts led to the creation of a detailed actor/stakeholder map for each case, where we identified the key stakeholder groups that either participated in or were affected by the program, as well as the specific individuals who played an active role in the program’s evolution.^{A4} These stakeholder maps were validated with several informants for each of the cases. We then conducted interviews with each of the key individuals across stakeholder groups. Interviewees were asked to relate chronologies of objective events, behaviors, choices at critical junctures, and facts of the processes described.^{A5} In every instance, the goal was to identify the individuals responsible for the particular evolution of a case, as well as the specific

tactics they employed throughout the process, to better understand the rationale behind their decisions as well as the factors that led them to succeed or fail. In total, we conducted 161 interviews across the eight cases. Interviews were complemented with a wealth of archival information including media articles, private documents (donor reports, internal presentations and communications, etc.), and public documents (announcements, academic articles, editorial pieces). These data were used to trace the chronological list of events for the overall development of each case. Each storyline was developed in an extensive document that established the causal links described by the subjects and ensuring a balanced consideration of different stakeholders.^{A6}

The third component, conducted in parallel to the eight case studies, consisted of interviews with prototypical representatives of each of

the stakeholder groups, or individuals who would clearly describe the typical experience of enacting a particular stakeholder role. Using the stakeholder map and initial hypotheses as starting points, this stage focused on the dynamics that shape the interactions between stakeholder categories. The work consisted of 34 in-depth interviews with representative actors from each stakeholder group. The interviews focused on each individual’s needs, assumptions, operational constraints, main concerns, professional and ideological backgrounds, timelines, and aspirations—especially concerning the development, dissemination, and use of novel evidence in development practice. This in-depth analysis resulted in a more nuanced and detailed stakeholder and system map that more clearly identified both breakdown points and paths of connection that hinder and facilitate the exchange of knowledge and information across stakeholder groups, as well as a refined

^{A4} See Canales, R. (2016). *From ideals to institutions: Institutional entrepreneurship and the growth of Mexican small business finance*. *Organization Science*, 27(6), 1548-1573.

^{A5} Davis, J. P., & Eisenhardt, K. M. (2011). *Rotating leadership and collaborative innovation: Recombination processes in symbiotic relationships*. *Administrative Science Quarterly*, 56(2), 159-201.

^{A6} *Ibid.*



Table A2: Case Studies

Country/Program	Description	Dates of Intervention	Number of Interviews	Primary Stakeholders
South Africa				
Collaborative Analysis of Labor Intervention Effectiveness	Employment program introducing new elements to vocational training	2011 – 2016	42	Government, Researchers
FUEL: Feed, Uplift, Educate, Love	School nutrition program	2007 – present		NGO
Ghana				
Teacher Community Assistant Initiative	Remedial education program for primary school children in reading and math through teaching assistants from local communities	2010 – 2013	30	Researchers, Government
Graduating the Ultra Poor	Poverty alleviation program integrating elements of social protection, livelihoods development, and financial services	2010 – 2013		Researchers, NGO
India				
Teaching at the Right Level	Remedial education program for primary school children in reading and math	2001 – present	51	NGO, Researchers, Government
AQUA+	Water purification drops for retail sale	2010 – present		NGO
Mexico				
Progresa Oportunidades	Poverty alleviation program using conditional cash transfers	1997 – present	38	Government
Programa Primer Empleo	Employment program using government incentives for the private sector	2007 – 2012		Government



set of hypotheses about the breakdown of communication and about possible interventions to solve it.

Across the three components, we conducted a total of 226 interviews. All interviews were in-depth and semi-structured, with an average length of around 90 minutes (minimum of 60, maximum of over 120). Around two-thirds of them were done in person and the rest were conducted remotely. All interviews were recorded and transcribed verbatim.

Data analysis was conducted in several stages. Each of the 226 interview transcripts was coded extensively to identify first-order concepts related to the integration of evidence into development practice. First-order concepts include “concerns about reputation” or “short-term decision-making”. This required multiple readings of interview transcripts, field notes, and archival data to associate nearly every passage of text with one or more codes. These codes were then grouped into second-order themes,^{A7} always contrasting them with current research on the integration of evidence into practice. Second order themes included “incentive structures” or “timing misalignments”, each of which was developed extensively in a memo that explored the characteristics, tensions, and contradictions of each theme. In stage

three, we mapped the codes to each of our case narratives to detect patterns of activities, constraints, and decisions that defined the evolution of each case at critical junctures. This allowed us to identify similarities and discrepancies across cases, as well as to create comparable counterfactuals that could account for differing outcomes.^{A8}

In stage four, we created process maps, concept maps, data tables, and detailed case synopses that linked key challenges, events, and decisions to the specific alternative tactics employed by actors and then to their subsequent consequences for the development program or intervention in question. This final set of analyses revealed a somewhat consistent set of factors faced at comparable stages by actors across our different settings. Throughout our analysis, we iterated between emerging insights, existing theory, and matched comparisons across cases to identify the mechanisms that operated at critical junctures.

It is worth mentioning that, at two moments of the project (the first after our first set of expert interviews was over and the second after the completion of our initial case narratives) we hosted a workshop with two different groups of highly experienced representatives from each of the stakeholder groups. During these

workshops, we discussed our emerging findings and we gathered additional, essential insights from participants. The workshops served to validate and deepen our understanding of emerging insights.

A7 Glaser BG, Strauss AL (1980) The Discovery of Grounded Theory: Strategies for Qualitative Research (Aldine Publishers, Hawthorne, NY).

A8 We ensured consistency in coding across the different cases and authors through several mechanisms, including: a) a selection of interviews was coded by two or more coders, after which they reviewed discrepancies and agreed on their resolution, b) a common project book where all the codes were collectively kept, aggregated, and analyzed, c) a weekly meeting to review coding process and to develop a joint coding standard, d) memos were developed jointly, with contribution from and verification by the different team members, among others. [Access here.](#)