



**Evidence in
Practice**

Full Report

Print

Case Studies

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Evidence In Practice: Toward an Integrated Approach

By Rodrigo Canales, Tony Sheldon, Tory Grieves, and Brendan Lehan





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Table of Contents

Introduction	03
Existing Paradigm	10
Clusters of Encouraging Practices	20
Developing a New Paradigm by Leveraging Encouraging Practices	28
Appendix: Data + Methods	37
Case Study: Teacher Community Assistant Initiative	42
Case Study: Graduating the Ultra Poor	71
Case Study: Teaching at the Right Level	97
Case Study: Aqua+	121
Case Study: Progresa Oportunidades	143
Case Study: Programa Primer Empleo	165
Case Study: Collaborative Analysis of Labor Intervention Effectiveness	189
Case Study: FUEL: Feed, Uplift, Educate, Love	209

Image Credits

Page 9 + 60: Yooku- Ata Bedu, USAID/Ghana
Page 13 + 160: María del Mar Gutiérrez
Page 15 + 79: Kasia McCormick/USAID
Page 16 + 171: [Geraint Rowland](#)
Page 19 + 166: [Bud Ellison](#)
Page 21 + 197: [Hans-B. Sickler](#)
Page 22 + 100: Ryan Lobo/Pratham Books
Page 26 + 76: Kasia McCormick/USAID
Page 33 + 186: [Jan Truter](#)
Page 34 + 104: Ryan Lobo/Pratham Books
Page 36 + 97: Ryan Lobo/Pratham Books
Page 36 + 141: [Enrique Saldivar](#)
Page 42: Innovations for Poverty Action
Page 44: Yooku Ata-Bedu, USAID/Ghana
Page 46: Innovations for Poverty Action
Page 50: Stella Benezra/Innovations for Poverty Action
Page 52: Innovations for Poverty Action
Page 54: Innovations for Poverty Action
Page 63: USAID/A. Kauffeld
Page 70: Susan Quinn/USAID
Page 80: Yooku Ata-Bedu, USAID/Ghana
Page 82: Susan Quinn/USAID

Page 83: USAID/Ghana
Page 87: USAID/A. Kauffeld
Page 89: USAID/A. Kauffeld
Page 95: Ryan Lobo/Pratham Books
Page 101: Pratham Books
Page 102: Ryan Lobo/Pratham Books
Page 108: Ryan Lobo/Pratham Books
Page 109: Ryan Lobo/Pratham Books
Page 111: Pratham Books
Page 118: [Daniel Bachhuber](#)
Page 121: [Prato9x](#)
Page 123: [sdfgsdfgasdr](#)
Page 125: [Blink O'Fanaye](#)
Page 127: UN Women/Ashutosh Negi
Page 129: [Nithi Anand](#)
Page 130: Oxfam International

Page 139: [Bud Ellison](#)
Page 143: [Eli Duke](#)
Page 145: [Son of Groucho](#)
Page 147: [Darij + Ana](#)
Page 148: [Doris Hausen](#)
Page 154: [Kashfi Halford](#)
Page 175: [Bud Ellison](#)
Page 183: [David McAughtry](#)
Page 188: Kate Holt / AusAID
Page 193: [Jan Truter](#)
Page 202: [Megan Trace](#)
Page 204: Trevor Samson / World Bank
Page 208: Trevor Samson / World Bank
Page 215: [Steve Evans](#)
Page 215: [Randy OHC](#)
Page 215: [Steve Evans](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Introduction

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 research 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. 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.¹

International development work seeks to substantially improve living conditions for low-income households, using scarce resources judiciously and efficiently. Achieving this requires that organizations and governments base their work on the best available knowledge and evidence—about both the science that may support policies and programs as well as the organizational, political, and social mechanisms that can ensure effective design and delivery. Too often, however, practice in the field of international development, encompassing both government and NGO policies and programs, lag behind state-of-the-art research and evidence. In some cases, research and evidence fall outside of

what is realistic or actionable because of a gap between academic research and its practical implications. In others, relevant and actionable research fails to reach practitioners because of breakdowns or misalignments in the communication, incentives, language, timing, and relationships across actors. To ensure that policies and programs are designed using the best available evidence, it is first critical to understand why key actors, many of whom are committed in principle to integrating evidence into practice, struggle to systematically do so. This requires an exploration of how structural constraints and incentives—which we conceptualize as the assumptions, operational constraints, field-realities, career and organizational pursuits, and conceptual models—are shaping the current “evidence-to-practice ecosystem.” Our research seeks to better understand where incentives between actors are in alignment, where they are misaligned, and the potential strategies and structures that various stakeholders can take to leverage the former to mitigate the latter.

¹ See the detailed description on [Data + Methods here](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Our work takes a practice-centered approach² to analyze the predominant paradigm for the integration of evidence into practice, identify common barriers, and look for leverage points. Our analysis illustrates key practices that actors can engage in to transform how problems are defined and solved, as well as how and when different stakeholders collaborate to integrate knowledge and evidence into practice and thus achieve broader impact. Taken together, the proposed changes suggest a new paradigm that shifts its emphasis from the *translation* of evidence *into* practice towards the *integration* of evidence *in* practice.³

² See, for example, Bechky, B. A. (2011). *Making organizational theory work: Institutions, occupations, and negotiated orders*. *Organization Science*, 22(5), 1157-1167.

³ This is consistent with Carlile's (2004) discussion of the effective integration of knowledge across organizational and disciplinary boundaries. See Carlile, P. R. (2004). *Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries*. *Organization Science*, 15(5), 555-568.

⁴ 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.

⁵ Further information on design thinking in development can be found in the following articles: M. Mintrom and J. Luetjens (2016), "Design Thinking in Policymaking Processes: Opportunities and Challenges" *Australian Journal of Public Administration*, 75, no. 3: 391-402; Mark Evans and Nina Terrey (2016), "Co-design with citizens and stakeholders," in *Evidence-Based Policy Making in the Social Sciences: Methods that Matter* (Bristol: Policy Press), 243-262; Emma Blomkamp (2017), "Co-Design for Government: Magic Bullet or Magical Thinking?," [Access here](#); M. Mintrom and J. Luetjens (2017), "Creating Public Value: Tightening Connections Between Policy Design and Public Management," *Policy Studies Journal* 45: 170-190; Andrew Gunn & Michael Mintrom (2017), "Evaluating the non-academic impact of academic research: design considerations," *Journal of Higher Education and Policy Management* 37: 20-30.

Evidence-based Practice in Context

The field of international development has become increasingly sensitive to a basic question: How do we know if what we are doing is truly helping? This has reoriented the field around the specific needs of and outcomes for intended beneficiaries,⁴ which in turn has opened two broad and complementary lines of inquiry: **How do we know if we are asking the right questions?** And, **how do we know if we are investing in the right solutions?** The first question is influenced by the expansion of design thinking principles and methodologies into different fields, including economic development. This user-centered approach to development involves a set of rigorous and systematic tools to deeply consider the intended beneficiaries' perspectives in the identification of problems and potential solutions.⁵ The second question is influenced by the development of increasingly rigorous analytic tools in the social sciences. This evidence-based approach to development focuses on the rigorous evaluation of the extent to which a proposed solution actually delivers its desired outcomes.⁶ Our work centers on the latter question and seeks to identify current roadblocks that inhibit the integration of the best-available evidence into the design and delivery of development solutions.

⁶ Our research focus on evidence-based practice is an extension of recent work to incorporate findings from experimental economics, behavioral science, cognitive psychology, and behavioral economics into the various stages of public policy analysis and design. See especially: Abhijit V. Banerjee, Esther Duflo (2011), *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty* (New York: Public Affairs); M. Gopalan and M.A. Pirog (2017), "Applying Behavioral Insights in Policy Analysis: Recent Trends in the United States," *Policy Studies Journal* 45: S82-S114.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

A variety of actors, including academics, philanthropists, impact investors, and development practitioners,⁷ have grappled with the question of how to identify, support, and implement the “right solutions.” Their efforts to understand what constitutes evidence, how to apply it to a particular context, and the implications for each party’s respective roles have led to vigorous and often contentious debate on the meaning and implications of a growing “evidence imperative” within development. These debates have evolved alongside broader changes in the priorities of public and private funders and an increased separation between implementing organizations and those that provide funding and oversight.

Our research starts from the current state of the relationships among development actors and its implications for the integration of evidence into development practice. It is important to acknowledge that what is considered relevant, timely, and useful evidence is itself contested. Similarly, there is disagreement on what it means to actually integrate evidence into practice. Thus a broad understanding of the various types of knowledge that qualify as ‘evidence’ and the different approaches that qualify as ‘integration’ is an essential starting point to the study of how to better integrate evidence into development practice.⁸

Based on the identification of the structural barriers encountered in the integration of evidence into practice, we articulate several encouraging practices which have been adopted by a wide array of actors to mitigate or overcome these barriers, summarized below. These observations illuminate a set of structural adjustments that could be made to the existing system to shift it in the direction of a new paradigm centered on the ethos⁹ of evidence-based practice. We summarize the five clusters of encouraging practices here and then explore them in depth later in this report.

From Incentive Misalignment to Value Alignment

Encouraging practices mitigate misaligned incentives by: 1) convening various organizations to develop a shared understanding of a problem before launching program design, and 2) negotiating a project structure that allows actors from diverse organizations to provide their unique contributions to the integration of evidence into a project while also complying with their own, disparate incentives. A negotiated collaboration among stakeholders can make explicit the contrasting (and converging) incentives and constraints across the various actors, as well as compromises and adaptations that may be required to ensure integration of evidence into practice.

A new paradigm of “evidence in practice” would thus start from the identification of a compelling, shared problem as a convening place where the common values that are shared by all and the unique needs of each actor are specified upfront.

7 Throughout the text we will use the term “practitioners” to refer to the government actors, NGOs and implementation partners, and others who are responsible for designing and executing development projects and programs.

8 Alternative definitions of evidence are addressed in Julius Court, and John Young (2003), “Bridging research and policy: Insights from 50 case studies,”; G Reid et al. (2017), “Minding the gap: the barriers and facilitators of getting evidence into policy when using a knowledge-brokering approach” Evidence & Policy: A Journal Of Research, Debate & Practice 13 no. 1: 29-38; Craig Mitton et al. (2007), “Knowledge transfer and exchange: review and synthesis of the literature,” Milbank Quarterly 85, no. 4: 729-768.

9 By ethos we mean the interdependent set of values and practices that determine how a community of practice defines its work. See Fayard, A. L., Stigliani, I., & Bechky, B. A. (2017). How nascent occupations construct a mandate: The case of service designers’ ethos. Administrative Science Quarterly, 62(2), 270-303.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

From Disparate Definitions of Evidence to Bodies of Evidence

For stakeholders to develop an understanding of the critical principles and features of a particular intervention, the evidence they use must have both a certain level of robustness and contextual nuance.¹⁰ In our research, this happened most effectively when actors across stakeholder categories focused on cultivating *bodies of evidence* for a *particular problem*, comprised of complementary types of evidence, including RCTs, qualitative studies, and practitioner experiences.

A new paradigm would explicitly create the mandate and the supporting structure to incorporate multiple sources and types of evidence. This would start from the collective understanding of a shared problem, would emphasize *situated* (vs. abstract) knowledge, and would identify which types of evidence can be most relevant and useful to different types of stakeholders and at different junctures in the process.

From Timing Misalignments and Lack of Trust to Long-term Collaborative Relationships

Asynchronous time cycles across stakeholder groups get in the way of starting or maintaining trusting relationships. They can be mitigated through the deliberate cultivation

Taken together, the proposed changes suggest a new paradigm that shifts its emphasis from the *translation* of evidence into practice towards the *integration* of evidence in practice.

of long-term relationships that purposefully span the cycles of various stakeholder groups. Strong yet flexible relationships are key to developing bodies of evidence and integrating them into practice. Building relationships that span multiple instances of evidence generation and integration can create a “scaffolding” that supports the pursuit of broad-based, lasting change.

In a new paradigm, formal structures would naturally convene (and demand the participation of) the necessary diversity of stakeholders around a compelling and shared problem. Such structures would provide funding, project management, and governance mechanisms to explicitly establish long-term, collaborative, transparent, and fluid relationships.

From Fear of Failure to Learning in Process

The most successful examples of evidence integration lessen the distinction between evidence generation and incorporation. Instead, they design iterative approaches that simultaneously generate (different types of) rigorous *situated* evidence and integrate it into practice. With an emphasis on *learning*, these projects transform the need to negotiate evidence generation and integration into an asset rather than a roadblock. This

¹⁰ This questions the frequent treatment of ‘knowledge’ as an abstract, codified, and transferable asset and emphasizes the nature of situated knowledge, or the fact that knowledge is always enacted in practice. See, for example, Bechky, B. A. (2003). *Sharing meaning across occupational communities: The transformation of understanding on a production floor*. *Organization Science*, 14(3), 312-330.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

To ensure that policies and programs are designed using the best available evidence, it is critical to understand why key actors, many of whom are committed in principle to integrating evidence into practice, struggle to systematically do so.

explicit commitment to learning opens the door for different types of information flows across stakeholders to share experiences, perspectives, and insights.

In a new paradigm, collective learning to better solve a shared problem would be the central mandate of a collaboration structure, paired with dedicated funding, resources, and mechanisms to promote and disseminate learning.

From 'Champions' to Structured Collaboration

In our research, successful integration of evidence into practice was possible because of the (often extraordinary) actions of individuals at both the informal and the institutional levels. These champions' passionate commitment to exploring a problem allowed them to focus more on integrating evidence to enhance impact than on the role, job, or career imperatives imposed by their stakeholder affiliations.

In a new paradigm, the collaboration structure would formalize many of the tasks that champions must currently perform, which they often discover only through trial and error and which often entails substantial personal risks. The new paradigm would still benefit from such champions, of course, but it would provide a structure to quickly identify them and formalize their work so it can be appropriately valued.

We explore these themes in more depth below, after our exploration of the existing paradigm and the barriers within it.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The Pathways of Integration

At the most basic level, to integrate evidence into practice means to identify what works, to do more of it, and to do it in more places. In turn, it also means that actors will become better at the process of identification, replication, and transfer of evidence. Thus, a useful framework for identifying whether evidence has been integrated into practice is to consider three pathways that are mutually reinforcing but have differing scopes and timelines. We will refer to them broadly as *scale*, *spread*, and *structure*:

Scale

The most basic goal of evidence-based practice is to identify 'what works' and to do more of it. *Scale* therefore describes the extent to which specific programs are improved and expanded in their original settings. The accumulation of evidence on the effectiveness of a program justifies increasing the resources invested in it; thus if the evidence generated shows positive outcomes, there will be a higher probability of scaling.

Spread

A second goal of evidence-based practice is to promote learning across settings and organizations. *Spread* refers to the adaptation and application of key concepts to new organizations and contexts. The spread of evidence-based practice depends on the ability to evaluate whether evidence for a certain intervention is timely, relevant, and useful for adaptation to different contexts. This mode of integration is by definition more diffuse, harder to measure, and longer term, but is often the way that integration achieves a broader impact. Spread often occurs through social and organizational networks or across levels of an organization. Spread can be supported by deliberate efforts to move beyond simple broadcasting of (non-situated) information towards the creation of communities of practice that build connections between relevant actors, organizations, and disciplines; help identify and situate relevant evidence and useful practices; and jointly create situated evidence to improve collective practices.

Structure

A third goal of evidence-based practice is for participating actors to increase their organizational capacity for integrating evidence over time. *Structure* thus describes enhanced organizational capacity to generate, analyze, and integrate evidence into the design and implementation of interventions, as well as into overall performance evaluation.

Addressing the Three Pathways

The case study on Pratham offers an insight into all three pathways discussed above: it has achieved major scale in its own operations in India; its work has spread to adaptations of its pedagogical methodology in other countries (including Ghana); and its ongoing partnership with researchers to generate and analyze ongoing evidence is reflected in an organizational structure based on evidence-informed practice.

At the most basic level, to integrate evidence in practice means to identify “what works,” to do more of it, and to do it in more places.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The Pathways in Action

The Indian NGO Pratham, and the scale and spread of its signature remedial tutoring program, “Teaching at the Right Level” (TaRL), provide a helpful illustration of the three pathways in action. In 2001, researchers from the Massachusetts Institute of Technology approached Pratham because it was considered unusually data-driven and able to scale interventions quickly and inexpensively. The researchers conducted evaluations which demonstrated that Pratham’s TaRL program was cost-effective and led to significant learning gains for children. The evaluations also confirmed Pratham’s ability to administer programs at scale. Over the ensuing 16 years, the researchers and Pratham have continued their collaboration, expanding Pratham’s signature program to 29 states in India as of 2017. This growth represents significant **scale**, informed by (and enriching) the evidence of TaRL’s effectiveness.

As the evidence accumulated it created increased opportunities for the **spread** of evidence-based approaches to remedial education. One example is Ghana’s Teacher Community Assistant Initiative (TCAL), an education reform project coupled with a randomized control trial that was conducted between 2010 and 2013. Pratham’s “Teaching at the Right Level” tutoring program provided the evidence and key design aspects for TCAL. This project was a direct result of the professional and personal connections between the researchers working in India and those wanting to launch a similar program in Ghana. The impetus to adapt the program to Ghana was both to advance the compelling research agenda and to work with a national government partner to explore implementation at a national scale.

In addition to reaching an impressive scale in its educational programs, Pratham is considered unique among education NGOs in India for its investment in building internal capacity for monitoring and evaluation,



its willingness to expose itself to evaluation by external researchers, and its commitment to sharing evaluation results publicly. Pratham’s commitment to learning is summarized well by CEO Rukmini Banerji and co-founder Madhav Chavan:

“The major lesson for Pratham was that the journey of transformation of communities and of systems is a long one, with continuous learnings at every step. Internally as an organization and externally as a major player on the Indian scene, Pratham learned that it is important to be flexible and nimble, to seize and to create opportunities, and to continue to push the learning agenda on every available forum.”

This capacity for continuous learning exemplifies the **structures** that have resulted from Pratham’s long-term investment in evidence-based practice. Pratham’s approach has also transformed how the government evaluates education, as Pratham helped design the national survey that is now at the heart of policy design and evaluation.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Existing Paradigm

Better integration of evidence into practice in international development requires an assessment of the ways in which evidence is (and is not) currently incorporated into policy and practice. A conceptual model of this process is useful for identifying the structural reasons behind the failures, as well as the useful practices behind the successes that can become leverage points for broader change.

Our research involved, among other things, exploring how actors in international development view themselves and others within the field. Their perspectives broadly and consistently described the integration of evidence into practice as residing within an ecosystem where a set of archetypical stakeholder groups interact. This is not a perfect or comprehensive description (e.g., some organizations and especially individuals fall within more than one stakeholder category), but it helps identify—as *defined by actors themselves*—the most frequent and critical roles, incentives, and relationships that define the complex dynamics between “evidence” and “practice.” These stakeholder categories, and examples of organizational actors within them, are:

Researchers: research institutions, universities, research think-tanks, expert consultants, national statistical data aggregators, some international agencies.

Financiers: foundations, multilateral agencies, private sector social investors, development banks, high net worth individuals, bilateral government funders.

Intermediaries: knowledge translation platforms, some think-tanks, self-appointed “translators” in other stakeholder groups.

Policy Makers: national, regional, and local governments.

Implementers: practitioner NGOs, government implementing agencies.

Beneficiaries: households, recipient organizations, communities involved in projects.

Influencers: media, lobbyists, influential individuals, public intellectuals.

Each stakeholder group is described as mostly immutable, constrained by the formal organizations (and professions) that people belong to. These organizational structures and incentive systems, as described below, not only define each stakeholder group, but also create a number of structural barriers in the relationships across groups. In consequence, there are enormous perceived constraints to the integration of evidence into practice, which by definition occurs only through cross-stakeholder relationships.

Interactions between actors, who mostly self-identify as belonging to a given stakeholder category, are currently conceptualized as a linear model of *translating* evidence to practice. In this simplified model, actors described evidence as a form of abstract knowledge that is passed along from researchers to policymakers and implementers, through a series of distinct steps that translate it until it is amenable to practice. This flow, however, is impeded by a number of structural barriers. Each step in the sequential process is perceived as the purview of a particular stakeholder type, defined and regulated by that group’s norms and incentive systems. The sharp distinctions between stakeholders reinforce the perceived



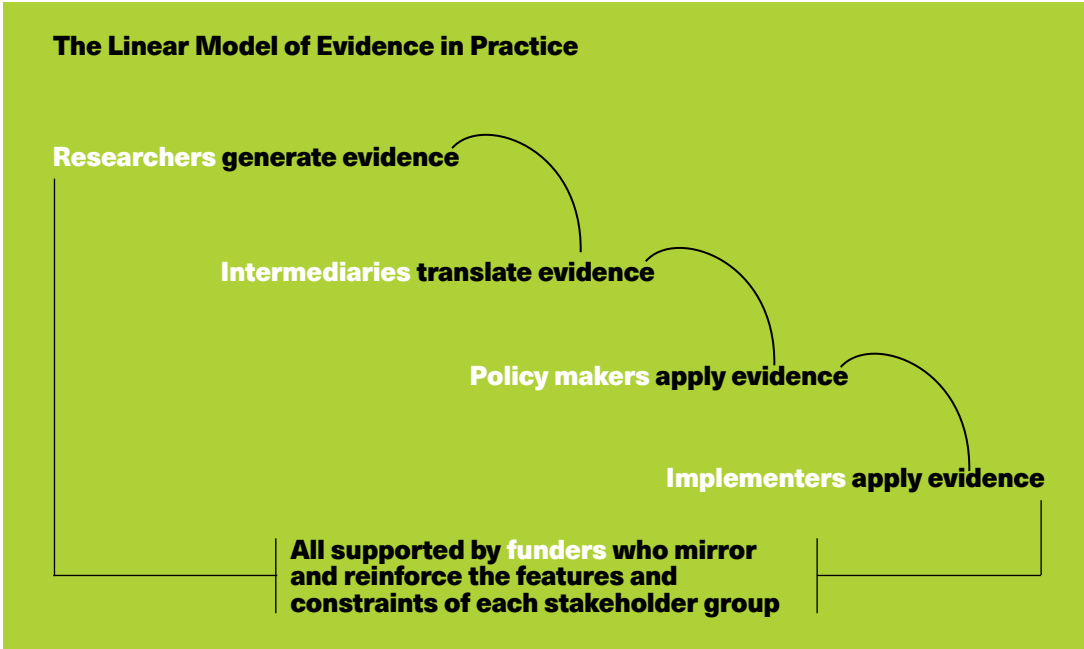
Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

need for a sequential approach to ‘translation.’ Academic evidence, for example, is often generated and published only for academic audiences, so it is abstracted from operational realities and unavailable or inaccessible for actors in other stakeholder groups. Formal translation is then required for policymakers and implementers to consider such evidence in the design of new interventions. This translation can only happen through organizations or individuals who act as

intermediaries between evidence producers and practitioners, and so on. Individual actors can (and do) work outside of the established paradigm, of course, but given formal and informal norms and incentive structures, such deviations entail significant personal risks and opportunity costs.

Transaction costs refer to the time, energy, and social/political capital required to establish, cultivate, and maintain relationships with other stakeholders in pursuit of certain goals. Generally speaking, actors are less willing to develop relationships outside of their established expectations, norms, and incentives because of the costs involved. An implementer, for example, generally would need to invest more time to develop trust and mutual understanding with a researcher than with a fellow implementer.

Opportunity costs describe the personally valuable activities that an actor must sacrifice in order to pursue a goal outside of the established paradigm. In the case of researchers, this might include spending time to package research for use by policymakers, for which researchers are not explicitly rewarded. In this case, the opportunity cost is time that could have been spent writing and publishing research in academic journals, which is explicitly rewarded by the tenure system. As in this example, opportunity costs are frequently quite high.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Structural Barriers to Evidence Translation

Barriers within the existing, linear model of evidence translation can be grouped into two categories based on their effect on the behavior of stakeholder groups: those that make relationships *across* stakeholder groups more challenging and thus isolate each category, and those that regulate an actor's behavior *within* its stakeholder group. In addition, many of these barriers enhance a fear of failure and a lack of trust that permeate the entire system.¹¹

While these barriers heavily constrain the integration of evidence into practice, they do so unintentionally. That is, the formal structures that underpin these barriers were not built with the explicit goal of impeding the integration of evidence. Rather, the barriers are the unintended consequence of a set of norms, organizational processes, incentive structures, and patterns of behavior that evolved in the field of international development to solve specific (and important) problems at a time when the integration of evidence into practice was not a central concern. Consider, for example, grant reporting cycles and strict monitoring practices that donors have established with their grantees. There is an excellent reason why grantees create detailed budgets with proposed, programmatic uses of a donor's funds. As a secondary consequence, however, these reporting principles today get in the way of experimentation, flexible

partnerships between grantees and other stakeholders and, ultimately, evidence integration.

Within each stakeholder group, behavior is largely dictated by a set of distinct incentives and norms that have the unintended consequence of rendering each group more isolated rather than more collaborative—despite the broad mission to improve the lives of beneficiaries, which is shared across stakeholder groups.

We first turn to the barriers that isolate stakeholders from one another.

¹¹ Many of these barriers (alternative definitions of evidence, complexities of collaboration, variant decision timelines, different core audiences, and negative past experiences) are discussed in the literature as leading to a lack of trust: Julius Court,

Within each stakeholder group, behavior is largely dictated by a set of distinct incentives and norms that have the unintended consequence of rendering each group more isolated rather than more collaborative.

and John Young (2003), "Bridging research and policy: Insights from 50 case studies," [Article access](#); G Reid et al (2017), "Minding the gap: the barriers and facilitators of getting evidence into policy when using a knowledge-brokering approach" *Evidence & Policy: A Journal Of Research, Debate & Practice* 13 no. 1: 29-38; Craig Mitton et al. (2007), "Knowledge transfer and exchange: review and synthesis of the literature," *Milbank Quarterly* 85, no. 4: 729-768.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Barriers that Isolate Stakeholders From One Another

Incentive Misalignments

Formal and informal incentive structures are frequently not conducive to—and are often in contradiction with—the integration of evidence into practice.¹² 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). Organizational incentives rarely encourage an explicit search for external evidence, much less the generation of internal evidence that could lead to continuous adaptation of programs and policies as new learning emerges. For example, politicians must win elections. Overall, even when individuals may have the motivation to integrate evidence into practice, structurally misaligned incentives may render such work highly costly, if not impossible.

¹² Incentive misalignments are also discussed in recent academic literature: Brian W. Head (2010), “Reconsidering evidence-based policy: Key issues and challenges,” *Policy and Society* 29 no. 2: 77-94. [Article access](#); Ashley Thomas Lenihan (2015), “Institutionalising evidence-based policy: international insights into knowledge brokerage,” *Contemporary Social Science* 10, no. 2:114-125. [Article access](#)

“Politicians say researchers are too picky or too slow or too complicated and researchers say politicians are... not really worried about doing the right things. So it’s very tricky. Incentives are very different in each of the sides as are the needs of what type of evidenced-based research and evidenced-based recommendations you are really trying to promote.”

POLICYMAKER



Programa Primer Empleo | Mexico

Mexico's Programa Primer Empleo (“First Job Program”), launched in 2007 during the first months of a new presidential administration, was intended to be a major national intervention to create new permanent jobs in the private sector. However, the new administration's incentive to get a signature program implemented as quickly, and with as much public fanfare, as possible led them to design the program with virtually no consultation of the business sector that the government sought to engage. The incentives offered by the government were seen by the business community as poorly framed and full of underlying risks, so the program failed to meet its goals.



[Teacher Community Assistant Initiative](#)
[Graduating the Ultra Poor](#)
[Teaching at the Right Level](#)
[Aqua+](#)
[Progresa | Oportunidades](#)
[Programa Primer Empleo](#)
[CALIE](#)
[FUEL: Feed, Uplift, Educate, Love](#)

“Practitioners have very little time and bandwidth to pick up new information. It’s a tremendous constraint. They have so many things to attend to. The rules and procedures grow ever more complicated because they are responsible to do all kinds of things to mitigate fiduciary risk, environmental and social risk, and to meet all the processing guidelines. It takes up their bandwidth for learning.”

FUNDER



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Alternative Definitions of Evidence

Definitions of what constitutes “evidence” vary by stakeholder group: Academic researchers generally refer to findings derived from rigorously designed studies, ranging from randomized controlled trials to in-depth qualitative studies. In contrast, implementers often think that academic studies oversimplify or underestimate contextual factors and complexities; to them, evidence is what is learned on-the-ground, from experiences where positive impacts were observed.

Moreover, researchers tend to view evidence as a hierarchy, where some forms of evidence (e.g. an RCT) are categorically more legitimate than others, “a ladder of the quality of evidence as to how much we can believe in the results.” In contrast, implementers tend to view types of evidence as different but not hierarchical, existing along a spectrum or ‘menu’ of options. This directly follows researchers’ tendency to think of evidence as abstract, ‘universal’ knowledge, while implementers have learned that knowledge is always and necessarily *enacted and situated in practice*, where few universal principles seem to hold across multiple complex contexts. In consequence, even when actors agree that integrating evidence is important, they differ—sometimes in irreconcilable terms—on what actually constitutes valid evidence and one actor’s “evidence” may be misunderstood or discounted by another group.



Graduating the Ultra-Poor | Ghana

The Graduating the Ultra Poor (GUP) project was a randomized controlled trial and pilot program conducted in northern Ghana between 2010 and 2013. The GUP program, in conjunction with nine other programs around the world, sought to test the effectiveness of the “Graduation” approach originally developed by BRAC in Bangladesh. In Ghana, the program was designed by the research firm, Innovations for Poverty Action, which co-implemented the program with Presbyterian Agricultural Services, a local NGO.

GUP was based on, and contributed to, rigorous evidence about a novel economic development intervention. The program was immensely successful in terms of learning from and contributing to a global body of evidence on poverty alleviation among the ultra-poor, which was its intended purpose. But Ghanaian participants involved with the project noted that GUP left a vacuum in terms of developing capacity for scaling of the approach on the ground in Ghana. The focus on generating academically rigorous evidence (the stated goal of the program) did not accord comparable importance to other kinds of “evidence” (such as feedback from field operations on the organizational implications of the program) which could have nurtured local capacity building and eventual scaling of the program.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“[The Researcher] had to have a publication that would pass muster with peer review in the behavioral economics world, which is an entirely different concern, of course, than what the practitioners have. [Implementers] want to have a program design that is going to actually benefit their clients and because they’re trying to run a financially self-sustaining organization, actually retain their clients and attract new clients, and give them a competitive advantage over competitor organizations.”

IMPLEMENTER

Timing Misalignments

Researchers, policymakers, and implementers operate within discordant timeframes, hindering efforts to coordinate, let alone collaborate, on evidence-informed approaches. Electoral cycles and political windows differ from NGO funding and academic publishing cycles. Actors may recognize that timelines are misaligned, but are still constrained by the timeframes of their formal stakeholder groups.¹³

¹³ The effects of differing timelines are discussed in the literature in the following articles: Ashley Thomas Lenihan (2015), “Institutionalising evidence-based policy: international insights into knowledge brokerage,” *Contemporary Social Science* 10, no. 2: 114-125, [Article access](#); Craig Mitton et al. (2007), “Knowledge transfer and exchange: review and synthesis of the literature,” *Milbank Quarterly* 85, no. 4: 729-768. [Article access](#).



Programa Primer Empleo | Mexico

As noted above, Mexico's Programa Primer Empleo (“First Job Program”) was launched in the first months of a new presidential administration. Out of a desire to deliver quickly on campaign promises to generate new employment opportunities at a national scale, the program design and implementation were rushed, allowing for little consultation with the business community that was supposed to generate the new jobs, which came to view the program with suspicion and never participated at the anticipated levels. Once launched, however, the program became bound by government budgetary cycles and rules, which limited its flexibility to adjust to emerging evidence of its shortcomings.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Barriers that Regulate an Actor's Behavior Within Its Stakeholder Group

No Practice of Devoting Time and Resources to Integrating Evidence

Few organizations carve out explicit time for managers to explore emerging evidence in their field.¹⁴ Even fewer assign staff to find relevant evidence and translate it into accessible formats for the organization. As a result, the role of preparing and sharing evidence that is timely, useful, and relevant for practitioners is sometimes explicitly played by formal intermediaries (e.g., certain think-tanks). More frequently, an actor who holds a formal role within another stakeholder group informally takes on the (additional) responsibility of trying to integrate evidence, leaving no actor formally responsible for the process and creating no trace of institutional memory or learning.

Not Operationalizing Evidence

Even organizations with strong monitoring and evaluation departments often do not transform operational data into knowledge that can be widely used by the organization—or other stakeholders—to learn from past or existing programs. Data is thus used to evaluate retrospective operations, but not to improve the design of new initiatives.¹⁵ For example, monitoring and evaluation in government agencies tend to focus on the judicious use of resources, but rarely codify core programmatic lessons for the improvement of future policies. Failure to

¹⁴ The issue of organizations not providing space for managers to evaluate new evidence is highlighted in the literature as well: Jermey M. Grimshaw et al. (2012), "Knowledge translation of research findings," *Implementation Science* 7, no. 1: 50

¹⁵ This issue has been highlighted in recent academic literature: M S. Reed et al. (2014), "Five principles for the practice of knowledge exchange in environmental management," *Journal of Environmental Management* 146: 337-345.

"The entire field of evidence from whatever source is very poor on its interpretation and use. There's a lot of evidence generated that is never ever used and that's partly because I think we under-invest in but also don't understand properly the best processes...that information can be absorbed and lead to change."

IMPLEMENTER



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

operationalize evidence inhibits the transfer of experiential learning, which may be rigorous and convincing, to new contexts and prevents evidence from reaching key stakeholders after it is produced, as it remains linked internally only to a given initiative. “Accountability” is often limited to performing within agreed upon budgets and activity flows, and not extended to achievement of actual outcomes.

As a system, these barriers contribute to the propagation of two additional systemic inhibitors to the flow of evidence into policy and practice: an **intolerance of failure** and a **generalized lack of trust**.¹⁶ Incentive structures across stakeholder groups tend to heavily punish failure and reward conservative approaches. This dissuades experimentation with novel, evidence-based approaches that could yield invaluable learning—especially when such approaches originate from a different stakeholder group. Such risk aversion can further hinder the integration of novel evidence into practice, even when stakeholders recognize its value and applicability.

¹⁶ *Intolerance of failure is described in the literature: Michael Howlett (2009), “Policy analytical capacity and evidence-based policy-making: Lessons from Canada,” Canadian public administration 52, no. 2: 153-175.*

“...I’ve been unbelievably frustrated with the number of academics that have absolutely no conception whatsoever of the practical implications of the research they do, of the fact that they never look to see what their colleagues are doing at all when they do their own research, and how much of this is wasted.”

INTERMEDIARY

Aqua Plus | India

Aqua Plus is a water purification product developed by TARA (Technology and Action for Rural Advancement). While rigorous evaluation of the underlying technology was done by TARA’s sister organization, Development Alternatives, TARA was charged with developing the final product and marketing it to Base of the Pyramid customers. Due to several factors, including limited funding and capacity constraints, TARA did not devote a comparable level of attention to assessing market development as Development Alternatives had to assessing the underlying technology. Moreover, while TARA was proactive in collecting internal evidence in the form of reviews and feedback from the communities where they sold Aqua Plus, TARA lacked a rigorous, internal monitoring and evaluation process to analyze these data and operationalize them into new practices on the ground. This limited investment in understanding the behavioral aspects of potential customers’ motivations contributed significantly to Aqua Plus not achieving its projected level of scale.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Misaligned incentives, alternative definitions of evidence, complexities of collaboration, differing decision timelines, incompatible core constituencies/audiences, and often negative past experiences, make it difficult and costly to initiate and sustain cross-stakeholder relationships. Without a relational history, actors have a low baseline of trust to build on, further complicating potential collaborations.



Programa Primer Empleo | Mexico

Mexico's Programa Primer Empleo ("First Job Program") was significantly inhibited by mutual distrust between the government and the private sector. Policymakers, wary of employers' potential to bend the rules and distort benefits of the government employment program, intentionally made the program's rules complicated and difficult to circumvent. This only fostered the companies' mistrust of the government's actions; employers were suspicious of the rigid rules imposed on them and resented that they had not been included in the program's design process.

"I think between the government and the funders, again, in an ideal world it should be very collaborative, but we have seen instances where there is animosity because in some ways the government feels that the funders are 'trying to do what I failed at' and 'they want to showcase that they've done a better job than me.'"

FUNDER



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Clusters of Encouraging Practices

Real and constraining as the barriers we described above may be, in our research, particularly the eight case studies, we identified an array of actors who experienced, understood, and sought to overcome these barriers in an attempt to integrate evidence into practice. These efforts by unusually committed individuals revealed a set of encouraging practices that can mitigate or overcome existing barriers. These observations illuminate a set of structural adjustments that could be made to the existing system to shift it in the direction of a new paradigm that more productively embraces an ethos of evidence-based practice.

We identified five clusters of encouraging practices, which address the current paradigm's key barriers, which are discussed below.

From Incentive Misalignment to Value Alignment

The incentive structures that currently confine actors to their stakeholder groups are not only deeply ingrained, but also exist for valid reasons, making them difficult to change or circumvent. We observed a cluster of encouraging practices that seek to mitigate misaligned incentives by: 1) convening various organizations to develop a shared understanding of a problem before launching program design, and 2) negotiating for a project structure that allows actors

from diverse organizations to provide their unique contributions to the integration of evidence into a project or program while also complying with their own, disparate incentives.

A common starting point for effective negotiations is a shared focus on solving a particular problem for which an intervention (program, policy, etc.) is appropriate. This is not because other entry points are not possible, but because, according to our observations, it is where actors across

A common starting point for effective negotiations is a shared focus on solving a particular problem for which an intervention (program, policy, etc.) is appropriate.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

stakeholder groups are most likely to have a shared interest and language: at the core of a problem or issue that affects a population of interest. The overarching, joint question then becomes: what factors are important in the design and development of an intervention?

Identifying a set of shared values can create the trust and relational space for more complex negotiations on the process to design and develop an intervention, as well as where existing evidence can be incorporated and new evidence needs to be generated over the course of a project. Shared values also allow the various stakeholders to determine, from the beginning, what each party needs and hopes to get out of a particular project. This allows for a project structure that can meet the otherwise conflicting criteria of its key stakeholders.

An open negotiation among stakeholders can make explicit the contrasting (and converging) incentives and requirements across the various actors, as well as compromises and adaptations that may be required to ensure integration of evidence into practice. Such a *negotiated collaboration* can generate the relational, political, and social capital needed to sustain a project through its various and unavoidable ups and downs, with a foundation of understanding, trust, and



CALIE | South Africa

As explored in the case study on South Africa's "Collaborative Analysis of Labor Intervention Effectiveness" (CALIE), researchers and policymakers came together in an upfront process to design an intervention to grapple with the country's high unemployment rate. The research questions were explicitly linked to policymakers' concerns, so that piloting an evidence-informed intervention (the addition of a reference letter to an ongoing career counseling program) allowed the academics to engage in rigorous research while tackling a high-profile issue for the government. Both sides paid careful attention to aligning incentives at multiple levels. For example, the project would require no additional expenditures by the Department of Labour, and evaluation metrics aligned with field staff's existing goals as well as responding to their expressed desire for better measures of the outcomes of their work.

reciprocity among the parties. It can also cultivate an appreciation not only for the unique contributions that other actors can make, but also for the structural constraints they face, which helps uncover "currencies" to offset such constraints (e.g., policy makers have access to unique and valuable data, a valuable currency for academics; academics, for their part, can provide a seal of objectivity or legitimacy that can become a political asset for policymakers). The willingness of each group to compromise, where possible, on the project's design, governance, and assignment of "credit" for success, are important ways to increase the likelihood that relevant, timely, and useful evidence is generated and integrated in a program.

A new paradigm of "evidence in practice" would thus start from the identification of a compelling, shared problem as a convening place where common values and unique needs are specified upfront. It would create a collaboration structure where the participation of the necessary diversity of stakeholders is mandated, where their specific contributions and needs are recognized ex ante, providing clear incentives for participation, and where funding and project management mechanisms allow for transparent and efficient coordination without the need for heroic champions.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love



Pratham | India

Pratham's approach to implementation of remedial education programs is founded on a three-way collaborative process between itself, the Jameel Poverty Action Lab (J-PAL) as researcher, and the government of each state in which it seeks to operate. Pratham's commitment to scale requires a model that is cost-effective and simple enough to implement in extremely diverse contexts. As a result, even though RCTs have shown significant gains in learning outcomes, Pratham has insisted on adjusting and testing the model in order to identify the most impactful, cost-effective components. Understanding these essential features enabled Pratham to be flexible when adapting the program to a new context in partnership with government stakeholders. Pratham launched each new adaptation or expansion by convening the relevant stakeholders to jointly explore how best to improve learning outcomes for students. The partners overcame common design and implementation challenges due to an integrated, collaborative approach that leveraged stakeholders' complementary resources, networks, and insights to achieve their deeply shared goal of improving remedial education.

From Disparate Definitions of Evidence to Bodies of Evidence

To inform practice effectively, evidence for a particular concept requires both a certain level of robustness and contextual nuance. In our cases, this was most effective when actors across stakeholder categories focused on cultivating *bodies of evidence for a particular problem*, comprised of complementary types of evidence, including RCTs, qualitative studies, and practitioner experiences.¹⁷ Such bodies of evidence, which develop over time and encompass diverse settings, seek to identify, for a given intervention (a) the critical principles that must be present regardless of context, (b) the specific useful practices that can best support these principles across a variety of contexts, (c) the features that *may* be adapted to better fit different contexts, and (d) the features that *must* be adapted to respond to local contexts.

There is a second, distinct aspect to reconcile disparate definitions of evidence. As discussed above, a collaborative approach to program design allows for a process where not only different types of evidence are

¹⁷ The value of developing a rich body of evidence is discussed in: Lois Orton et al. (2016), "Putting context centre stage: evidence from a systems evaluation of an area based empowerment initiative in England, Critical public health 27, no. 4: 477-489.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

brought to the table, but also where actors develop an appreciation for and agree to integrate types of evidence that matter most to each group. In our case studies, this translated into a deliberate effort to codify diverse types of evidence that would support key decisions throughout project implementation.

At the culmination of projects (or periodically during programs) actors in our most effective cases identified, collected, and shared lessons from across levels of decision making. This consolidated evidence and associated lessons contributed back to the bodies of evidence for particular concepts and clarified the challenges of integration with implementation. In the current paradigm, it is typical to find that only a certain type of evidence is captured and shared (e.g. the analysis of an RCT). But a more representative reporting of the various types of evidence generated during a project is also important for engaging stakeholders of various types, enriching the bodies of evidence pertaining to a particular problem.

This approach to integrate complementary types of evidence across levels and organizations within a project had the secondary, unintended consequence of building fluency in different types of evidence among participants, which persisted beyond the focal project and set the stage for more effective integration of evidence into practice in the future.

A new paradigm would explicitly create the mandate and the necessary structure for the incorporation of multiple sources and types of evidence as relevant and useful to multiple stakeholders in this collective understanding of a shared problem.

Progresa (Oportunidades) | Mexico

The Progresa case shows the roles that evidence played in a major national social protection policy, from analyzing available evidence in the design of the program to committing to generate, evaluate, and use evidence in the course of implementation. The decision to launch the program came before a randomized controlled trial had demonstrated the effectiveness of the model. But evidence clearly showed that generalized and in-kind subsidies—which had been the cornerstone of social policy in Mexico until then—were ineffective. The Progresa team thus made it a priority not only to research the best available evidence, but also to generate new data and evidence in order to confirm (or contradict) their underlying hypotheses and to inform the final design of the program.

Policymakers chose to pilot and evaluate initial results before launching a full-scale program in order to understand the impacts of the program and the components that needed to be improved. Also, precisely because of the rigor and transparency with which the evidence was generated, it provided critical political strength to carry on with the rollout and scale-up of the program against enormous political resistance.

The use of evidence did not stop at the design phase, but generation and use of evidence was built into the very structure of the ongoing program, so that quality and progress could be assessed and adjusted on an ongoing basis according to emerging evidence.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

From Timing Misalignments and Lack of Trust to Long-term Collaborative Relationships

Each stakeholder category is bound by a different time cycle. For example, policymakers are profoundly affected by electoral cycles. An annual funding cycle dramatically shapes the actions of funders. These relatively fixed constraints create asynchronies across stakeholder groups that can get in the way of starting or maintaining trusting relationships. They can also be mitigated through the deliberate cultivation of long-term relationships that purposefully span the cycles of (and job tenures of actors within) various stakeholder groups.

Strong personal and institutional relationships are key to developing bodies of evidence and constructing effective evidence-informed projects and programs. Building relationships that last over multiple instances of evidence generation and integration can create a “scaffolding” that supports continued efforts to bring about broad-based, lasting change. Regular communication between actors outside the context of a particular program or project to develop a shared understanding of problems and questions that need to be answered was consistently highlighted as key to ensuring a successful relationship. The trust that emerges from these partnerships

Building relationships that span multiple instances of evidence generation and integration can be an important catalyst for lasting change.

allows multiple stakeholders to develop a deeper understanding of particular bodies of evidence and, more broadly, of effective processes to integrate evidence and practice.

We observed several mechanisms that facilitated longer-term commitments. One was framing partnerships, collaborations, or initiatives around a broader problem, rather than a specific program. This signaled a commitment to multiple iterations of projects involving the same stakeholders and provided incentives for actors to develop “relationship capital” over time. It helped when actors were able to access sources of funding

Pratham | India + TCAI | Ghana

Since its first collaboration in 2001, Pratham has built a strong partnership with the Jameel Poverty Action Lab (J-PAL), based on long-standing professional relationships between the principals of the two organizations. As one Pratham staff member described the partnership:

“I think we’ve been very lucky because nobody forced any of these [evaluations] on us, so it was a [voluntary] coming together of two sides ... We were very lucky to have them as partners because we have done some work with other people as well, but we see that they [J-PAL] really treat you as a partner and as an equal.”

Over the course of than 15 years, J-PAL and Pratham have collaborated on the design, implementation and evaluation of dozens of education programs in India. And the connection between J-PAL and Pratham led directly to the pilot and evaluation of the Teacher Community Assistant Initiative (TCAI) in Ghana.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

that explicitly spanned multiple iterations of specific projects. Such cycle-spanning structures reduced the disruption of some timing misalignments and allowed actors to align on shared values rather than to compete on accessing resources.

In a new paradigm, formal structures would naturally convene (and demand the presence of) the necessary diversity of stakeholders around a compelling and shared problem. Such structures would allow for funding, project management, and governance mechanisms that explicitly establish long-term, collaborative, transparent, and fluid relationships matched to the time cycles and needs of the collective understanding of and progress in addressing the problem at hand.

From Fear of Failure to Learning in Process

As noted above, the most successful examples of evidence integration lessen the distinction between evidence generation and application, and focus on designing approaches that simultaneously generate (different types of) rigorous evidence and develop an iterative process for integrating evidence into practice. These projects turn the need to negotiate evidence generation and integration into an asset rather than a roadblock.

An explicit emphasis on learning shifts the conversation away from “success vs. failure” and focuses on the broader set of goals that each project aspires to accomplish, what questions it hopes to answer, and how those two aspects can enhance long-term results.

In that sense, the best examples of evidence integration resulted from programs with robust, explicit learning and evidence sharing agendas. This commitment to learning opens the door for different types of linkages and information flows across stakeholders to share experiences, perspectives, and insights with the explicit (and non-threatening) goal of learning. As a result, the most successful cases frame stakeholder collaboration around a specific program as a launch pad for both program-specific and broader learning.

This reframing around learning can have the double effect of lowering the perceived risk of failure (as it generates valuable learning) while

at the same time increasing the perceived value of the collaborative relationship. An explicit emphasis on learning shifts the conversation away from “success vs. failure” and focuses on the broader set of goals that each project aspires to accomplish, what questions it hopes to answer, and how those two aspects can enhance long-term results.

In a new paradigm, collective learning to better solve a shared problem would be the central mandate of a collaboration structure, paired with dedicated funding and resources to promote learning and the formal mechanisms to disseminate it.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love



Graduating the Ultra-Poor | Ghana

The Graduating the Ultra-Poor (GUP) project 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 underlying “Graduation approach” in about 100 countries worldwide. From the global Graduation pilot program’s inception, CGAP and the Ford Foundation defined knowledge sharing and outreach as key tenets. In addition to quantitative and qualitative research, CGAP and the Ford Foundation created a community of practice to bring together implementers and researchers across ten pilot sites to share experiences and lessons learned. The community of practice also offered exchange visits, ongoing technical assistance, and an annual convening of implementing organizations with international NGOs, funders, policymakers, and academic partners interested in the Graduation approach.

From ‘Champions’ to Structured Collaboration

In our case studies, successful integration of evidence into practice was possible because of the (often extraordinary) actions of individuals at both the informal and the institutional levels. At the informal level, successful integration of evidence was driven by self-motivated individuals who, acting alone or in collaboration with others, persisted in their efforts to identify, understand, and overcome the structural barriers that stood in the way of evidence integration. This usually started from a deep commitment to a problem or cause that sustained their drive throughout the complexity and struggles of the integration process. Their passionate focus on a problem allowed these champions to focus more on bringing evidence to bear on achieving impact than on the role, job, or career imperatives imposed by their stakeholder affiliations. It also led them—through prior experience or empirical discovery—to identify and cultivate relationships with critical stakeholder groups, and individuals within them, that would need to participate to achieve impact. These champions sought to understand the priorities and constraints of diverse stakeholders, which in turn led to a holistic view of the system they were working in. As a result, they were able to

better navigate the system, build a common ground across stakeholder groups, and make true collaboration possible. It helped that, often, champions developed this unusual identity through professional experiences that straddled several of the critical stakeholder roles (e.g., a policymaker who was trained and worked as an academic researcher, or an NGO manager who had worked in government) which gave them the ability to understand and translate between different “languages.”

At the institutional level, we often found more formal intermediaries, or organizations with the stated mandate to translate evidence for practice. These actors worked to enhance the flow of information between different stakeholder groups, particularly between researchers and practitioners, by positioning themselves as an “honest broker” who explicitly gathers, packages, and transfers evidence. While the presence of such organizations was often helpful, because of the structural barriers discussed previously, the presence of such intermediaries was not enough to make integration possible. Rather, intermediaries, when present, either were an effective platform for or an invaluable ally to champions who, as described above, worked beyond formal constraints on behalf of the integration of evidence to practice.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Regardless of their origin and institutional affiliation, we found a remarkably consistent set of competencies and behaviors that champions used to perform their work of evidence integration:

- **Be problem-oriented.** Champions were adept at using the problem or question at hand as a north star around which to facilitate relationships and interactions across stakeholders.
- **Identify and balance interests of stakeholders.** With an understanding of stakeholders' diverse incentives, champions could facilitate a collaboration in which each stakeholder could meet the non-negotiable demands of its core constituency.
- **Identify opportunities to use evidence.** With their broader framing of the problem, champions were able to identify critical areas of collaboration, in which an exchange of stakeholders' "currencies" would be most effective (e.g. academics providing legitimacy for a government initiative in exchange for policymakers providing support for the program), making the mutual value explicit.

In the current paradigm, champions' behavior—the particular maneuvers they employ, the effort they exert—is rarely acknowledged and even more rarely rewarded.¹⁸ In a new paradigm, the collaboration structure would formalize many of the tasks that champions must currently perform, which they often discover only through trial and error and taking substantial personal risks. The new paradigm would still benefit from such champions, of course, but it would provide both a structure to quickly identify them and formalize their work so it can be appropriately valued—a scaffolding to support their evidence integration efforts.

¹⁸ *The role of intermediaries who effectively translate and communicate evidence is discussed in: Ashley Thomas Lenihan (2015), "Institutionalising evidence-based policy: international insights into knowledge brokerage," Contemporary Social Science 10, no. 2:114-125; Jermey M. Grimshaw et al. (2012), "Knowledge translation of research findings," Implementation Science 7, no. 1: 50; Fadi El-Jardali et al. (2014), "Capturing lessons learned from evidence-to-policy initiatives through structured reflection," Health Research Policy and Systems 12. [Access here.](#)*



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Developing a New Paradigm by Leveraging Encouraging Practices

Based on extensive conversations with diverse stakeholders, we have made an effort to extrapolate lessons from the encouraging practices we've observed to imagine what a holistic and systematic approach could look like. We believe that the path forward entails a combination of changes in both mindsets and structures for how international development evidence is generated, funded, and integrated. Our hope is that the new paradigm we lay out can provide a common language and a useful framework for organizations and individuals that are interested in increasing the integration of evidence into practice in international development.

Taken as a whole, the encouraging practices help develop a vision for a more productive ecosystem which integrates evidence effectively at a systemic level. Several themes can be drawn from practice that help elucidate a vision for the future: strong collaboration around a shared problem across all types of stakeholders, a broader understanding of the definition of evidence and stakeholder roles in its generation and use, and built-in systems of learning that are rooted in trust.¹⁹

We have endeavored to understand where the characteristics of the current ecosystem are most fixed and immutable, and where more mutable and potentially advantageous

opportunities exist for change. The new paradigm for integrating evidence and practice implied by the encouraging practices seeks to acknowledge the fundamental constraints and strengths of the various stakeholders, and we have highlighted lessons below that we hope can serve as guideposts for moving toward a new paradigm.

¹⁹ *The importance of stakeholder collaboration is discussed in: Mark Evans and Nina Terrey (2016), "Co-design with citizens and stakeholders," in Evidence-Based Policy Making in the Social Sciences: Methods that Matter (Bristol: Policy Press), 243-262; Emma Blomkamp (2017), "Co-Design for Government: Magic Bullet or Magical Thinking?". [Access here.](#)*

**Taken as a whole,
the encouraging
practices help
develop a vision for
a more productive
ecosystem which
integrates evidence
effectively at a
systemic level.**

Integrating Evidence and Practice

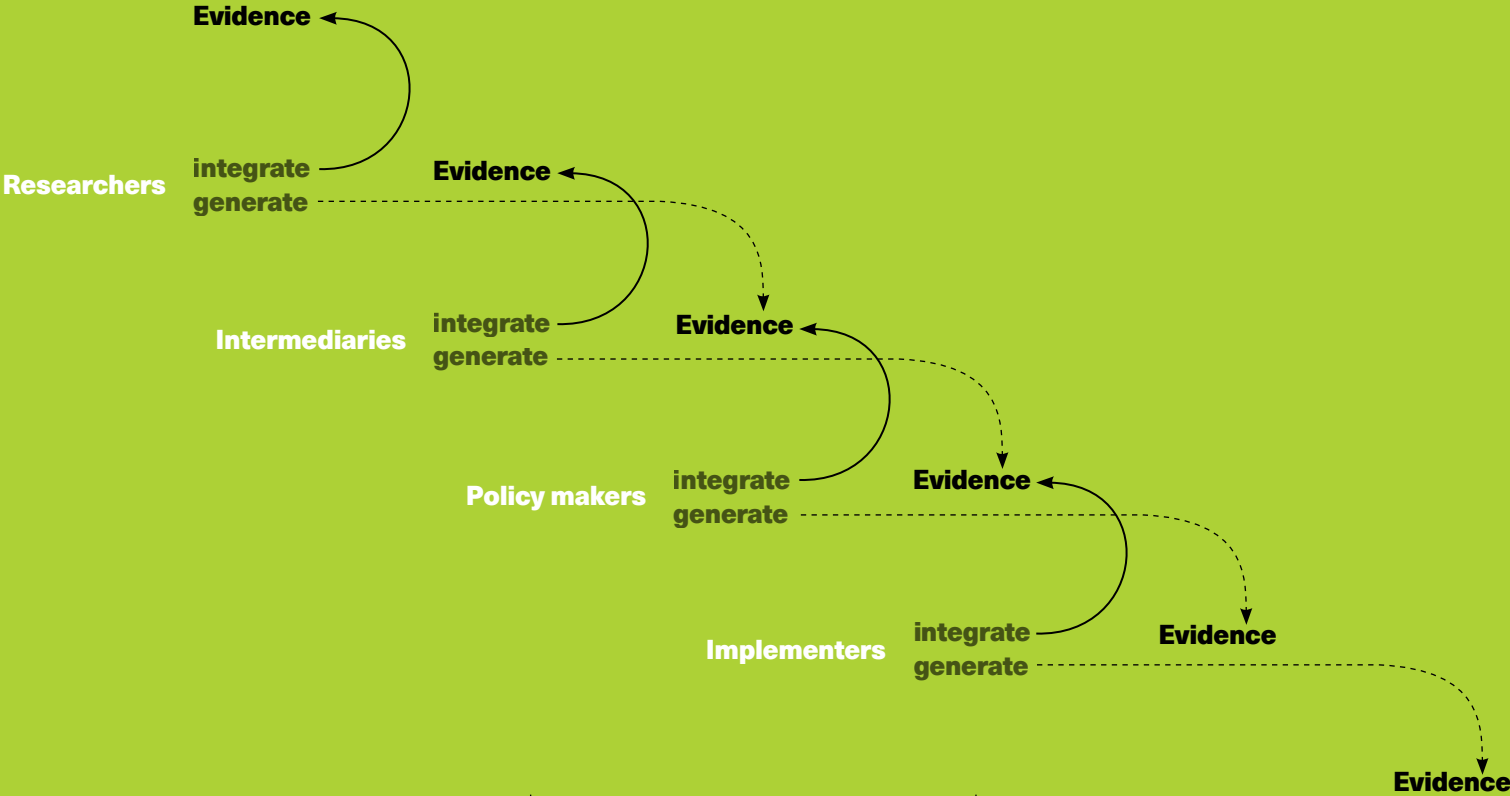
Better integration of evidence into practice requires an interconnected set of changes in the mindsets, relationships, processes, and roles that connect evidence and practice. At the highest level, a new paradigm moves away from the predominant conceptualization of the flow of evidence to practice as a sequential process of generation, translation, and application, and towards an integrated model for development of evidence and practice. This shift will be driven by a mindset



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

An Integrated Model of Evidence in Practice

A model in which evidence is **generated** and **integrated** by all stakeholders, based on mutual trust and a shared purpose centered on addressing key problems.



A model in which **funders** support cultivating and sharing bodies of evidence amongst all stakeholders.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

that focuses on compressing the gap between evidence generation and evidence-based practice, and requires a shift from a hierarchy of the types of evidence towards an approach that seeks to match the various types of evidence to the types of decisions where they are most pertinent, generating bodies of evidence for specific social challenges that multiple projects may be addressing.

Each stakeholder group will continue to have a unique role and contribution, but that contribution is made collaboratively, not sequentially or in isolation. An ex-ante, explicit recognition of each stakeholder's contributions and constraints could lower transaction costs, drive a shared understanding of the problem, and prioritize learning over time.

Collaboration as the Norm Rather Than the Exception

Our research suggests that the systematic generation and use of evidence grows out of an ongoing process of collaboration and negotiation among stakeholders. The mutual recognition of a shared problem plays a critical role as the 'glue' holding together the collaboration, providing critical motivation, common ground, a shared commitment to

Explicit acknowledgement and negotiation of potential common ground can help to establish and maintain relationships of mutual value.

using outcomes as the metrics for progress, and naturally emphasizing the need for collective learning.

The expectation of continued collaboration around shared problems helps spread the transaction costs (which are currently borne upfront by one or two groups) across the system as a whole as a long-term investment. This effectively creates economies of scale for evidence integration. In an improved system, stakeholders actively look for ways to support the integration of evidence into practice, viewing this as part of their mandate.

Identify the Currencies of Exchange

Because stakeholders possess different kinds of 'currencies,' an exchange of value can take place between stakeholder groups, providing the basis for meaningful negotiation. For example, researchers can offer the legitimacy of their work in exchange for access to government data that will enable their research; policymakers are willing to share their data with researchers in exchange for the legitimacy that comes with rigorously produced evidence. While in the current paradigm, such an exchange is desired but rarely openly discussed, our research suggests that explicit acknowledgement and negotiation of potential common ground can help to establish and maintain relationships of mutual value, which are essential for the integration of evidence to occur. This process also increases trust among stakeholder groups, as it requires open dialogue and deeper understanding of exactly what each group can offer the others.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Currencies of Exchange

Recognizing specific areas of compatibility is key to the identification of “currencies” that can be exchanged between stakeholders as a basis for negotiated collaboration. Often such currencies are readily available to one stakeholder (such as a rigorous evaluation conducted by researchers from a prestigious university) and can prove of great value to another stakeholder (such as granting external legitimacy to a policymaker’s intervention). Our research pointed to several key areas that underpin currencies of exchange, such as unique resources, needs, and priorities. This table, a distillation of our findings, is not meant to be exhaustive, but rather an attempt to provide examples of entry points for a negotiated collaboration.

STAKEHOLDER	TIME CYCLES	KEY AUDIENCES	TYPICAL INCENTIVES	RESOURCES
Funders	Typically annual funding cycles.	Board of directors, government regulators, the public (e.g. taxpayers, media, implementing partners, beneficiaries).	Ability to deploy money effectively. Alignment with organization’s long-term strategy. Their own issue area expertise/credibility.	Access to funds. Access to diverse sources of situated knowledge (through their grantees).
Researchers	Peer-reviewed article publishing timelines. Tenure timelines. Grant timelines.	Senior researchers making decisions on journal publications and academic tenure.	Publishing in peer reviewed journals. Tenure. Awards. Fundraising.	Recognized fluency/mastery in separating complex problems into their constituent questions, framing hypotheses, identifying causality, and assessing other kinds of evidence.
Intermediaries	Windows of opportunity for evidence translation, in which the timelines of other stakeholders align.	Policy makers. Implementers. Researchers.	Recognition and funding for being evidence-based.	Speaking multiple “languages”, ability to translate ideas across stakeholder categories, convening power, serving as bridges across timelines.
Policymakers	Typically annual budget cycles. Election cycles.	Elected officials (for bureaucrats). Political party, media, public opinion (for elected officials).	Promotion and avoiding mistakes (for bureaucrats). Positive press and public opinion (for elected officials).	Ability to scale programs. Access to data. Convening power.
Implementers	Grant timelines. Budget cycles. Needs of beneficiaries/clients..	Funders. Beneficiaries/clients.	On time and on budget delivery of agreed upon projects. Low overhead costs.	Access to beneficiaries/clients. Contextual sensitivity and intuition. Resourcefulness, resilience, agile problem solving. Lessons and data from programs.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

STAKEHOLDER	NON-NEGOTIABLE NEEDS	EVIDENCE PRIORITIES	CURRENCY OF EXCHANGE
Funders	Compelling justification for use of funds. Avoidance of fraud or misuse of funds.	<ol style="list-style-type: none">1. Their own experience2. Other funders (via word of mouth and observing capital allocation trends)3. Evidence generated by their programing4. Popular discussion (e.g. in media, public dialogue)5. Academic research papers	Willing to provide capital in exchange for results/recognition. Convening power. Support of cross-organizational exchanges.
Researchers	Need to be able to publish and otherwise advance their careers in academia.	<ol style="list-style-type: none">1. Their own evidence or other academic evidence in their field published in peer-reviewed publications.2. Popular research in their field (e.g. in media or elsewhere).	Ability to confer legitimacy through affirmation of being “evidence-based” in exchange expanded opportunities for ongoing access to data/information that furthers their research.
Intermediaries	Willingness of involved parties to engage in some compromise and make time for negotiation.	<ol style="list-style-type: none">1. Evidence and themes popular among policymakers and funders2. Evidence with widespread implementation and replication potential3. Published academic research	Willing to tap into networks, make markets for ideas, and translate between different stakeholder languages in exchange for access to other stakeholders and influence in decision-making.
Policymakers	No negative optics/scandal. Abiding by legal mandate of institutions.	<ol style="list-style-type: none">1. Their own experience2. What other policy makers tell them3. Evidence from own country4. Evidence from other countries5. Peer-reviewed journal articles	Willing to support a program in exchange for affirmation of their policies. Convening and mediation power/ability.
Implementers	Enough resources to implement effectively.	<ol style="list-style-type: none">1. Their own experience.2. Pragmatic evidence (e.g., best practices on how to integrate/implement an idea).3. Useful for fundraising (e.g., how can we design a project that is likely to be funded?).	Willing to provide access to beneficiaries/ clients and related data in exchange for funding and autonomy. Situated understanding of how to apply knowledge. Agile adaptation to emerging challenges of implementation.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love



CALIE | South Africa

The researchers involved in South Africa's "Collaborative Analysis of Labor Intervention Effectiveness" (CALIE) set out to build long-term relationships with policymakers at the Department of Labour (DoL). In order to do so, the "currencies" that each had to offer were explicitly identified and factored into the collaboration.

The researchers wanted to establish an ongoing relationship because they saw great value both in having access to the DoL's extensive database on labor statistics—a rich resource for ongoing research—and in the opportunity to generate new data in partnership with the DoL.

The DoL was interested in the collaboration in order to bring better evidence to bear on grappling with the issue of unemployment, which it was mandated to address, and to be seen by citizens as proactive in addressing a pressing social issue. In coming together around the shared problem, the DoL opened its data to the researchers, and the researchers always described CALIE as a DoL initiative.

Field staff were open to the collaboration because results of the research were shared with them in periodic workshops, giving them timely and useful information data on the impact and quality (rather than only the quantity) of their services, which they would not otherwise be able to collect. The project design also aligned with field staff's existing goals and evaluation metrics, bringing a sense of greater transparency to their ongoing operations.

Ongoing Collaborations Increase Synergy and Understanding

One of the clearest lessons from our research was the value of repeated opportunities for collaboration and engagement. Long-term relationships with repeated interactions create incentives for all parties to invest in relationship building and to be willing to make compromises when developing and managing projects or programs.²⁰ This process also transforms the identities of the stakeholders involved in the interactions, making participants into evidence integrators themselves as time goes on.

Flexibility in Design and Implementation

Developing this type of evidence integration process requires flexibility and foresight in both design and implementation of a program. Evidence, as illustrated above, is most useful when it provides the parameters for adaptation, rather than prescribing a single path forward. Flexibility during a project's design can help identify and accommodate the needs of various stakeholders, which encourages authentic integration into specific contexts. Projects that are also flexible in successive iterations of

²⁰ The value of long-term collaboration and trustbuilding is discussed in: M. S. Reed et al. (2014), "Five principles for the practice of knowledge exchange in environmental management," *Journal of Environmental Management* 146: 337-345.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

implementation create space for the real-time, nimble maneuvering required to effectively integrate evidence into practice on the ground.²¹ A new paradigm would be structured to create this level of flexibility through collaboration- and learning-oriented funding and governance, founded on an explicit acknowledgment of the contributions and needs of each stakeholder group.

Negotiated collaboration, rooted in mutual understanding of the needs and drivers of each stakeholder, is pivotal both at the outset and throughout this iterative process. Learning from various, evidence-based practices feeds directly back into subsequent problem identification.

Integrating evidence into practice is a complex process, but holds the promise of policies and programs in international development becoming both more effective and more efficient. Our research has shown that there is much we can do to make the process of evidence integration more systematic and productive. These changes will not happen easily or quickly. But with concerted effort from all stakeholders, we see a future where the new paradigm can increasingly become the norm.

Deliberately involving individuals from various contexts (within and across organizations and geographies) increases the probability that evidence-informed approaches can be successfully adapted to new contexts.



²¹ The importance of designing projects that can be flexible over time is discussed in: "Lois Orton et al. (2016), "Putting context centre stage: evidence from a systems evaluation of an area based empowerment initiative in England," Critical Public Health 27, no. 4: 477-489; M. S. Reed et al. (2014), "Five principles for the practice of knowledge exchange in environmental management," Journal of Environmental Management 146: 337-345.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The Pathways of Integration—Revisited

As discussed earlier, the incorporation of evidence happens mainly through three pathways (scale, spread, and structure), which are mutually reinforcing but have varying timelines and scopes. When programs explicitly consider the three paths of integration, they can create “economies of scope” that increase the likelihood of success. Evidence can be woven into practice in both direct and indirect ways, considering both the needs of multiple stakeholder groups and the stage in a given problem’s life cycle. That is, *scale*, *spread*, and *structure* happen on different timescales and are most critical at different times in the learning process. As we have learned, full integration can sometimes take a decade or more after the initial generation of rigorous evidence supporting the approach to the problem.

Scale

While there are numerous factors that contribute to scale, our research has highlighted two aspects of particular importance: the ability (1) to identify where you are in the life cycle of a problem, and then (2) to identify the most relevant and useful scale to seek at that stage. For an untested intervention, a modest pilot with limited reach and a rigorous, multi-faceted evaluation is likely the most appropriate scale. While plans for scaling will be dependent on whether evidence indicates the project is achieving its intended outcomes, it is nonetheless important to consider what kind of partner would be desirable should results prove to be positive, and engage them early on with an eye toward potential future scaling. Building relationships with these partners early in the process can create genuine ownership and management capacity among the potential partners, so that the evidence generated will have a higher probability of actually scaling.

Spread

The spread of evidence-based practice depends on the ability to evaluate whether evidence for a certain intervention is timely, relevant, and useful for adaptation to different contexts. If we start from a shared problem, we can then also identify, globally, where we stand on the collective understanding of that problem. Spread also depends on an understanding of different types of evidence and their value. Deliberately involving individuals from various contexts (within and across organizations and geographies) increases the probability that evidence-informed approaches can be successfully adapted to new contexts. Spread can be supported by deliberate efforts to disseminate information on new evidence and evidence-informed practices.

Structure

Structure describes the enhanced understanding and capacity of implementing parties to integrate evidence generation and learning into the design of their interventions, practices, and overall performance evaluation. This type of integration includes both evidence-based interventions (the ability of an organization to implement programs based on evidence) and evidence-oriented practices (using evidence as part of the ongoing culture, routine and structure of an organization).



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love



Pratham | India + TCAI | Ghana

As noted earlier, Ghana's Teacher Community Assistant Initiative (TCAI) grew out of the professional and personal connections between the J-PAL researchers—working on Pratham's Teaching at the Right Level (TaRL) in India—and the global leadership of Innovations for Poverty Action, a sister research organization to J-PAL and the initiator of TCAI. Research in India (and subsequently in Kenya) had demonstrated that the basic TaRL methodology delivered reliable and substantial benefits. At the time, Pratham, despite being one of the largest education NGOs in India, could not operate at the scale of a national government, leaving unanswered the question of whether the TaRL methodology could generate substantive impact at a national scale if implemented in partnership with government. The impetus to take the TaRL methodology to Ghana was the opportunity to work with a partner in government who could further test the compelling TaRL research and ideally bring implementation to a national scale.



Progresa | Mexico

The new approach pioneered in Mexico by Progresa to integrating evidence into policymaking was so influential that it not only informed the evolution of the program itself but also led to the creation of the National Council of Evaluation (CONEVAL), which now brings rigorous methodologies to the evaluation of all social programs in Mexico. This fundamental change in how the federal government evaluates its portfolio of social programs illustrates how the effective integration of evidence into practice can translate to new structures for routinely integrating evidence into practice in other areas.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

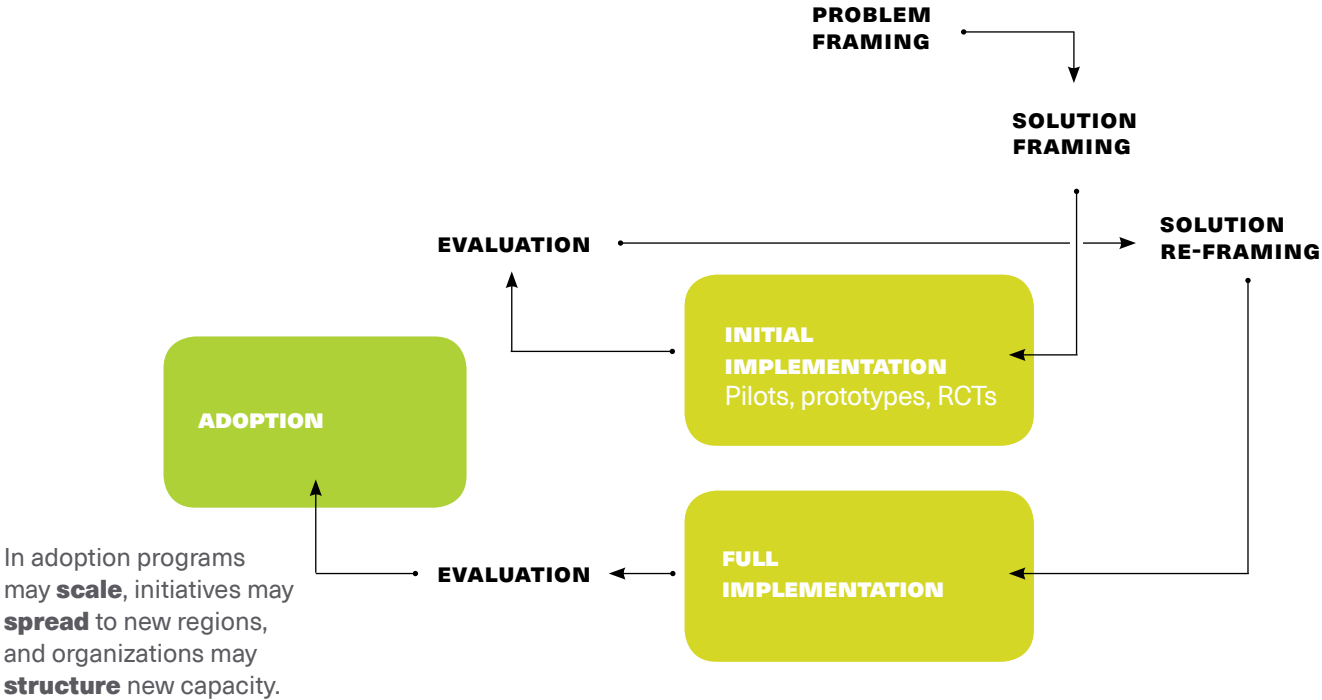
Principles Into Practice: Focus on Process

Our research, rooted in conversations with stakeholders around the world—in the U.S., Mexico, Ghana, South Africa, India, and elsewhere—has helped to clarify a path forward from the current paradigm to a new one. Steps toward this new paradigm will be based on: 1) a commitment by all stakeholders to frame their work together as a negotiated collaboration centered on addressing a shared problem; 2) an appreciation by all stakeholders of each stakeholder’s incentives, constraints, and “currencies” of exchange; and 3) an understanding of how to strategically leverage shared values to form long-term relationships.

Another critical element of an “evidence in practice” approach involves leveraging those relationships and exchanges to build a robust and rigorous process through which evidence is generated and integrated into practice.

In the new paradigm, the iterative process of integrating evidence into practice leading to broader Adoption is comprised of five key elements: Problem Framing, Solution Framing, Initial Implementation, Evaluation, and Full Implementation.

Rather than being prescriptive, this diagram is intended to be emblematic of the improved process of evidence integration that our research suggests. As such, we invite its adaption for real-world use as well as further questioning.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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.²² 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.²³ 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.²⁴ 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.

²² By development practice, we mean the work of government actors, NGOs, and others who are responsible for designing and executing development projects and programs.

²³ 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.

²⁴ George, A. L., & Bennett, A. (2005). Case studies and theory development in the social sciences. MIT Press, Cambridge, MA. Chapter 5.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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.²⁵ 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.²⁶ 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.²⁷

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

²⁵ See Canales, R. (2016). *From ideals to institutions: Institutional entrepreneurship and the growth of Mexican small business finance*. *Organization Science*, 27(6), 1548-1573.

²⁶ Davis, J. P., & Eisenhardt, K. M. (2011). *Rotating leadership and collaborative innovation: Recombination processes in symbiotic relationships*. *Administrative Science Quarterly*, 56(2), 159-201.

²⁷ *Ibid.*



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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,²⁸ 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.²⁹

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.

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

29 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.](#)



**Evidence in
Practice**

Full Report

Print

Case Studies

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Page **42** of **220**

GHANA | EDUCATION

Teacher Community Assistant Initiative

A national pilot for remedial education shows the value of proactive negotiation of evidence-based projects.

By Brendan Lehan





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part I: **The TCAI Story**

The Teacher Community Assistant Initiative (TCAI) was an education reform project coupled with a randomized control trial conducted in Ghana between 2010 and 2013. The project was designed by Innovations for Poverty Action (IPA), funded by Children's Investment Fund Foundation (CIFF), and implemented throughout the country by the Ghanaian government through Ghana Education Services (GES) and the National Youth Employment Program (NYEP). TCAI adapted and further tested evidence from India and Kenya showing that using teaching assistants from local communities to teach remedial classes can raise learning levels for primary school students who are below grade-level.

Reflecting on the impact of TCAI, it appears that the project increased the Ghanaian government's awareness of the importance of remediation as part of its education approach, and influenced important individuals in the Ministry of Education regarding evidence-based practice in remedial education. This influence is despite the fact that an agreement to scale up the program following the end-evaluation was ultimately abandoned. Project implementation involved challenging choices regarding how to adapt the evidence from India and Kenya to the Ghanaian context and deliberately prepare the project for scale via government adoption. The decision

to test remedial education at scale with a government partner required IPA to balance the tension between conducting rigorous research and creating programs that are viable at scale. It also required balancing the tension an intermediary can face between creating rigorous evidence and striving for success for its participants. Designing the TCAI implementation and evaluation showed the importance of collaboration across various implementing stakeholders from early on to test treatments that would be sustainable at scale, despite the risk that they might not show as strong an effect as smaller, more intensive and controlled pilot programs.

In addition, the management of the implementation and evaluation in partnership with GES required the researchers involved to make compromises to the research design in order to make implementation more practical and demonstrate their commitment to the overall success of TCAI. The decisions made by IPA and the government throughout the project, and the legacy of TCAI in Ghana and at IPA, reflect the complex ways in which large-scale and hands-on evidence creation can have both direct and indirect effects, such as more evidence-oriented policy and practice by government, strengthened relationships between researchers and government, and increased awareness of the importance of remedial education throughout the education system.

The story of TCAI also illuminates the role of champions inside and outside government in generating evidence, translating evidence, and preparing for scale based on evidence. The design, development, and management of TCAI demonstrates the importance of mechanisms to establish and maintain trust among stakeholders and the ways in which early involvement and subsequent ownership by policy makers and



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

government implementers is an important precursor to long-term adoption and effective implementation at scale. Four years after the project's conclusion it is clear that translation of evidence can be non-linear: while the TCAI program in Ghana hasn't been adopted at scale, there continues to be interest from neighboring countries and new development partners to incorporate the concepts of TCAI into new initiatives. The aftermath of TCAI suggests that not scaling a project immediately after evidence is generated does not imply that evidence isn't credible, useful, or in demand, but rather reveals the challenges and long timelines required to find funding and implementation partners.

Teaching at the Right Level

During the early 2000s researchers from the Abdul Latif Jameel Poverty Action Lab (J-PAL) initiated randomized controlled trials in India to test how lessons tailored to students' learning levels and delivered by teaching assistants could be used as a cost-effective approach to improve education quality and student learning, a concept they later termed Teaching at the Right Level (TaRL). Over the course of approximately a decade, research on this topic in India (partnering with the NGO Pratham), and later in Kenya, generated strong evidence from evaluations that programs to provide remedial education to students based



on their learning level have significant effect on learning outcomes.

The basic concept underpinning TaRL (and the eventual treatment arms used in TCAI) is that a typical classroom of students is heterogeneous in terms of their literacy and numeracy levels, and that accurately identifying the learning levels of each student, and providing targeted instruction to that level, can lead to a significant improvement in education outcomes. Early TaRL research investigated the use of teaching assistants to either: (i) pull students out of the classroom or keep them after school for more individualized



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“TCAI is an exception in which we had a nationally-run representative example implemented by the prime consumer of the evidence.”

RESEARCHER

tutoring, or (ii) break students into groups within the classroom for more individualized instruction by group.

The TCAI project in Ghana was a direct result of the professional and personal connections between the J-PAL researchers working in India and Kenya, and IPA's global leadership, as well as IPA's established presence in Ghana. Research in India and Kenya had demonstrated that the basic TaRL methodology, executed in randomized controlled trials, delivered reliable and substantial benefits. What was less clear was whether and how the methodology could be adapted and scaled to deliver the system-level benefits it was designed to pursue. At the time, Pratham, despite being one of the largest NGOs in India, could not operate at the scale of a national government, leaving unanswered the question of whether the TaRL methodology could generate substantive impact at a

national scale if implemented in partnership with government. The impetus to take TaRL to Ghana was the opportunity to work with a partner in government who could further test the compelling TaRL research and bring implementation to just such a national scale.

Implementing at Scale in Ghana

The decision to further test the TaRL concept in Ghana was driven by three key factors: (i) the existence of an IPA office in Accra, and the organization's knowledge and experience of working in the country, (ii) IPA's existing relationships with the Ministry of Education, which put IPA in a position to propose collaboration with a government partner, and provided a higher likelihood of ongoing collaboration, and (iii) the awareness among IPA and the government that there was sufficient need for remedial education among Ghana's school children. Encouraged by these factors, the initial IPA discussions

in Ghana regarding TaRL were launched by Annie Duflo, then IPA's Vice President and Research Director, and Wendy Abt, a J-PAL consultant and advisor. They were supported by Kelly Bidwell, the IPA Country Director at the time.

Testing TaRL at scale would require the engagement of Ghana Educational Services, which was responsible for administering Ghana's centralized primary and secondary education systems, as well as of a funder willing to support a randomized evaluation focused on large-scale implementation. In 2009, armed with the evidence from J-PAL's research in India and Kenya, Annie Duflo and Wendy Abt approached the Ghanaian Ministry of Education to initiate the project, and at the same time sought a supportive international donor to fund evidence generation at scale.

In the case of TCAI, the Ghanaian government, the international donor community, and IPA all played important roles in launching an ambitious project to further develop the global evidence base for TaRL and set the stage for the translation of TaRL into an education system at scale. While each party played a critical role in spreading TaRL evidence, none of the stakeholders could have accomplished TCAI on its own.



IPA: Intermediary, Evidence Producer, and Implementation Partner

Two key aspects of IPA leadership's approach (and their significant commitment to the project) were instrumental in testing the concept of TaRL at scale and its effect in Ghana. First, the previous involvement of IPA's leadership in J-PAL's India research made them knowledgeable about the evidence, and credible and motivated champions for the potential of TaRL. Second, testing TaRL concepts at a national scale was a deliberate decision to advance the evidence base for TaRL and set the stage for scaled adoption in Ghana. Together, these two aspects illustrate the way that researchers can become vectors for the spread and evolution of evidence for a particular development concept.

In this capacity as a vector for evidence, IPA combined the roles of intermediary and evidence producer to support the translation of evidence into policy and practice. As intermediary, IPA focused on how evidence could be used to achieve a significant change in how education was delivered in Ghana, namely through the adoption of effective interventions for remedial education. As an evidence producer, IPA sought an opportunity to evaluate whether the TaRL lessons from India and Kenya were relevant to national, government-led implementation. In this way,



IPA played a significant role in determining both the direction in which evidence on TaRL would evolve and the amount of time and resources that the concept received.

From IPA's perspective, Ghana was a good fit for this type of work because of its presence in the country and its solid relationships with the Ministry of Education and Ghana Education Services. The personal

relationships that IPA had already established in Ghana, together with the credibility of the evidence that had been produced in India and Kenya, were critical for convincing the government and other actors (e.g. donors and development partners) that TCAI was worth the government's time and attention, and was not going to undermine other education initiatives, projects, or priorities.

Also important were visits to Accra by Pratham, the NGO that pioneered TaRL in India, to help IPA convince the Ghanaian government of the value of both TaRL and randomized evaluations. Pratham visited several times while IPA was working to build support within the Ghanaian government and find a project funder, and later IPA organized trips to India for GES staff to observe Pratham's implementation first-hand. The involvement of Pratham played a critical role in establishing the credibility of the project, but at the same time it emphasized that the project was coming from outside the country, an impression that IPA and others had to work to mitigate to get Ghana Education Services fully onboard.

At the time, IPA was a relatively young research organization, whose mission was to develop rigorous evidence on international development programs, primarily through randomized controlled trials. The choice to

NGO: a non-governmental organization is any non-profit, voluntary citizens' group which is organized on a local, national or international level.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

partner with a government implementing agency presented both risk and opportunity for the organization. On the one hand, government implementation is known to be fraught with management and logistical challenges that can undermine program viability³⁰, and therefore the broader validity of experimental evaluation. At the same time, implementation at scale is dependent on interventions that are often only achievable through government adoption on a national level.

To help mitigate the risks, IPA added a third dimension to its role: implementing partner. As described below, the Ghanaian government implemented TCAI, with significant control of, and direct involvement in, the program's delivery. However, IPA hired its own implementation team in Ghana to work alongside the government to enable IPA to manage and troubleshoot implementation in a hands-on way. Making this additional investment in implementation support was part of an effort by IPA to ensure that an ambitious, large-scale project and evaluation had the best possible chance of accurately testing the TaRL concept. Through the creation of the TCAI Steering Committee, IPA and the government worked together to address both the risks and opportunities as they launched the project and managed it over time.

“You are engaging with government and they have to believe that you will not let them down. You’re not leading them down a path that they’re going to fail.”

IMPLEMENTER

Government as Implementing Partner

Another critical factor to the spread of TaRL was the Ghanaian government's receptiveness to partnering with IPA on a national project and evaluation. At the time that Abt and Duflo were looking to spread TaRL, Dr. Stephen Adu was the Director of Basic Education at Ghana Education Service. TCAI would not be the first time that GES had worked with IPA, and Adu was familiar with IPA from earlier education work in Ghana. Building on their existing relationship, IPA approached Adu and others at the Ministry of Education to test the feasibility of the project. Adu was responsive to IPA's desire for a potential project for several reasons, including:

- an existing relationship with IPA and a belief that it was a credible research organization that would be committed for the long term
- awareness of the extent to which Ghanaian school children were below grade level, and a desire to bring as many resources as possible to bear on the challenge. For example, a 2006 assessment supported by USAID had shown that less than 2% of students at the Primary 2 level were reading at the expected level
- deep personal experience performing education research and working at GES, particularly on teacher training

Stephen Adu's PhD training in education and 35-year career at GES, with a focus on teacher training and basic education, gave him a combined understanding of the pedagogical issues at stake with TaRL and the institutional dynamics of GES and the broader Ghanaian

³⁰ See for example: Tessa Bold, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a and Justin Sandefur, "Scaling-up What Works: Experimental Evidence on External Validity in Kenyan Education (March 27, 2013)," Center for Global Development Working Paper No. 321. [Available here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

education sector. His support of the IPA proposal went beyond mere receptiveness: he became a critical champion of the project during initiation and implementation, and many sources cited his support as a critical factor in getting it off the ground and through implementation.

Government support for the project did not come without serious concerns, including the financial viability of the project, the value of spending scarce time and human resources on a randomized controlled trial, the potential that the project might undermine professional teachers, and a wariness that the project would recreate unsuccessful initiatives that had already been tried. As discussed in the following sections, the credibility of IPA, the support of a committed funding partner, and accommodations in the design of the TCAI project were key to addressing these concerns.

Financing TCAI

The availability of a funding partner to support the project's implementation and evaluation was a critical factor. As IPA's leadership was building support within the Ghanaian Ministry of Education, they were also searching for a funder to support the project.

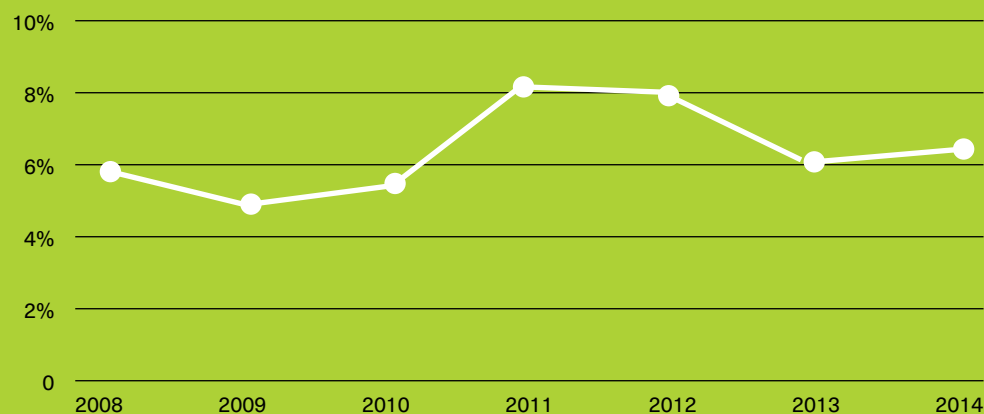
IPA secured initial, small-scale funding to pilot potential treatment arms for a TaRL project in Ghana and lay the groundwork for recruiting the government as the primary implementing partner. While this initial funding was critical for convening a modest piloting of potential interventions, a much more substantial investment was required to achieve a full-scale national project and the related, costly evaluation.

The vast majority of Ghana's education budget went towards salaries, and there were limited discretionary funds to support projects like TCAI. Ghana was one of the first development partners for major agencies like USAID, and there has been a long, stable history of

development partnership in the country. Since the early days of these partnerships, education has been one of the priority issue areas for most major development partners. Because of this, Ghana has been a hub for international development actors interested in education, and development organizations vie for the attention and endorsement of the Ministry of Education so that they can implement their projects. Given the jockeying for government attention and its fiscal constraints, external funding was a necessary precondition for government support and thus for the viability of any TaRL project.

While the initial pilots were underway, the IPA team searched for potential funders of the full-scale implementation and evaluation. In early 2010, The Children's Investment Fund Foundation (CIFF) began to seriously consider support and made several trips to conduct due

Ghana: Public Expenditure on Education as Share of GDP



Source: Global Partnership for Education



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“The people who are in charge must understand, must want it, and must be able to defend it with the ordinary people and say, ‘This is really good. We are doing this particular one just for this year. We will learn from it, and this will take us places.’”

IMPLEMENTER

diligence on a potential project. The effort of IPA and the Ghanaian stakeholders to convince CIFF that TaRL interventions would have an effect, and that the TCAI project could lead to permanent adoption of TaRL concepts in Ghana, eventually paid off. CIFF’s investment committee approved the funding for TCAI’s implementation and evaluation in November 2010.

Designing TCAI

The process of planning TCAI, and several associated decisions, exemplify how both Ghana Education Service (GES) and IPA

worked together to design a program that would provide credible evidence about, and set the groundwork for, TaRL implementation at scale in Ghana.

A first step in this process was the joint decision of IPA, GES and the project funder, CIFF, to create a steering committee. The committee was designed to bring together, from the early stages of project design, the key stakeholders in the project (especially from the Ministry of Education, including GES, and the National Youth Employment Program) to ensure that the project would have the

support necessary to be successful, and the high-level involvement required for eventual scaling.

This steering committee was designed to explicitly include skeptics of the project as well as supporters: adversaries and champions. This approach allowed the project’s designers to gather input from a diverse range of potential implementation stakeholders during the planning phase, and to generate a broader sense of ownership over the project, particularly among members of the Ghanaian government.

Participants described how the process of crafting the project proposal was designed to stage a negotiation to ensure that the resulting program was interesting to researchers, viable for implementation in the context of the Ghanaian education system, and aligned to funder priorities. The skeptics among the government were most concerned about the viability of the project to scale and succeed in the Ghanaian context, rather than outright opposed to the idea or the supporting evidence. This hesitation was largely alleviated by the fact that outside funding for the project meant that no resources—other than the opportunity cost of labor—were being diverted away from those stakeholders. The design of the project included a number



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

of choices intended to balance the logistical complexity of the program with the rigor of the evaluation, and adapt TaRL to the Ghanaian context. The three major areas of focus were (i) considering the long-term financial sustainability of the proposed treatments, (ii) resolving pedagogical concerns and differences, and (iii) prioritizing Ghanaian ownership of the project.

While TCAI was being designed, the belief among the steering committee was that the National Youth Employment Program (NYEP) was a stable government agency, and that integrating this existing, funded government apparatus into the design of TCAI was a sound tactic to prepare for the full-scale rollout of the lessons from the project. GES and the teachers' union were both interested in testing treatments delivered by classroom teachers, which would not require additional human resources, so they pushed the IPA team to test treatments that did not rely on outside volunteers. As a result, the project used the NYEP, which hired recent secondary school graduates, to provide the community assistants for TCAI, and also added a teacher-led treatment arm that would test the impact of a teacher-only version of TaRL. These choices could help test whether the program could be successful using existing Ghanaian resources. This partnership between GES



and NYEP, which was intended to increase the financial feasibility of the program, was unprecedented and required significant and complex negotiation to put into place.

Establishing Ghanaian Ownership

From the start, TCAI's emphasis was on the Ghanaian government's "ownership" of the project. The structure of the steering committee and decisions regarding program design were a concerted effort to put implementation in the hands of the

government and align the project to the needs of its education system. A defining feature of TCAI was the explicit creation of an implementation partnership between the government and a research organization, and the risk and opportunity this implied for program and evaluation viability, as well as for long-term scalability. One of the key approaches to mitigate the risks was an intentional emphasis on adapting the project to Ghana, and the willingness to balance the academic incentives that often animate research and evaluation in order to support



that adaptation. Annie Duflo and Jessica Kiessel, the two principal investigators for the project, participated in TCAI in their roles as IPA staff and in addition to their other full-time responsibilities, which helped insulate the team from the pressures that academics often face to prioritize the most novel research topics and seek rigor of evidence to the exclusion of other factors. It also enabled IPA to play the role of both researcher and intermediary, focused on empowering the Ghanaian implementing partners and setting the stage for eventual scale-up. Because of this, the IPA team could focus its energy on the pragmatic, if difficult, question of how to design and implement for the Ghanaian context, while still applying and testing the body of TaRL evidence in a scientifically rigorous way.

IPA's deep support for and instigation of many of the design choices for TCAI required compromise from a researcher's perspective. The decision to add teacher treatment interventions and to transfer so much implementation responsibility to GES was a departure from the methodology developed in India, and put the quality of implementation at risk. The results of the India research had convinced the IPA team that volunteer-led interventions were the most effective, raising the question of how the project could best incorporate classroom teachers in the

The IPA team could focus its energy on the pragmatic, if difficult, question of how to design and implement for the Ghanaian context, while still applying and testing the body of TaRL evidence.

delivery of the intervention. In addition, using GES as the project implementer put the IPA implementation team in a position of advising and supporting the government as they implemented, rather than a more traditional research-only role.

To manage the inherent tensions in the approach, IPA decided to play a hands-on role in implementation. While the program was executed by GES, using GES staff and volunteers from NYEP, IPA stepped into the day-to-day details of implementation to a

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

significant extent, and worked side-by-side with GES across many levels to support the Ghanaian government's implementation. In addition, as noted above, IPA invested in efforts like learning trips for GES leadership to observe Pratham's implementation of TaRL in India. These visits were frequently cited by GES staff as a critical moment for understanding the potential of the TaRL model and its potential impact in Ghana.

The IPA team was organized into two separate groups, which focused on the implementation and evaluation of TCAI, respectively. IPA's implementation team focused on working with GES and a local NGO (School of Life) to develop teacher training materials and train master trainers (primarily retired or senior GES teachers) who then trained volunteers and teachers. The IPA evaluation team was responsible for working with GES to develop assessments and overseeing a team of 260 surveyors that performed the midline and endline surveying. It is important to note that Maame Nketsiah, TCAI's Implementation and Policy Director, reported to the Steering Committee and GES rather than the IPA Country Director. These types of staffing and management decisions were important to all parties for emphasizing the Ghanaian ownership over TCAI's implementation.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Launching TCAI

Between November 2010 and March 2011, the project conducted a baseline survey as part of the randomized controlled trial. The baseline results showed very low learning levels across Ghana. While the GES members of the steering committee were not surprised by these results, given their already deep concern about student achievement, they were anxious about publishing the baseline results without being able to also show that they had found a solution. After much discussion and disagreement, the steering committee decided to hold off on sharing the baseline results publicly until there were initial evaluation results that could show whether TCAI was providing a solution. This compromise is another example of the good faith efforts the IPA research team made to orient its efforts towards the government's success and to build long-term goodwill and support for the program.

As a result of the initial pilots and the deliberations of the steering committee, the team proposed a final project design for TCAI that tested four different interventions, reaching 25,000 pupils. Five hundred schools across 42 districts, chosen to be nationally representative, were randomly assigned to one of five groups, for students in first through third grades:

- **In-school remedial TCAs.** Teacher community assistants (TCAs) taught remedial classes **during school hours**, with a focus on basic literacy and numeracy skills. The remedial sessions were targeted **to the weakest pupils**.
- **After-school remedial TCAs.** Teacher community assistants taught remedial classes **after school hours** with a focus on basic literacy and numeracy skills. The remedial sessions were also targeted **to the weakest pupils**.
- **Normal curriculum TCAs.** Teacher community assistants **pulled out students randomly** (i.e. *not* remedial students only) to review the teacher's lessons on literacy and numeracy for a few hours a day during class time. The assistant alternated which students were pulled out of class.
- **Targeted lessons training for teachers.** Public school teachers were **trained in how to provide small-group instruction** targeted at pupils' actual learning levels. Starting in the second year of implementation, these teachers split their students by ability levels, rather than grades, for one hour daily. This design decision is discussed in more detail in later sections.
- **A control group** that received **no TCAI intervention**.

Consistent with the overall, shared objective to use this project as the foundation for a national scale-up of the program, the funding included a commitment that the TCAI model would be scaled nationally if the evidence demonstrated positive results. As we will see, the scale up of TCAI never happened, despite evidence of positive effects from several of the interventions.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Making it Work

The launch and ongoing implementation of TCAI was characterized by efforts by GES and IPA to: (i) cultivate trust among the participating parties, (ii) increase local involvement, and (iii) remain adaptable to try to ensure the project's success. As TCAI kicked off in 2011, GES led work to recruit and train a cohort of community assistants from across the country to deliver the program, which involved working with local communities to recruit recent secondary school graduates to work in their local primary schools. In each region, IPA randomly selected schools to participate and worked with GES to provide the learning materials and training to both community assistants and classroom teachers.

As implementation proceeded, IPA adjusted its engagement to build trust with GES and increase the government's implementation capacity. At the start, IPA made a point of hiring local Ghanaian staff to manage the project from both the evaluation and implementation sides and to be on the front lines working with the GES staff implementing the program. The foreign IPA leadership deliberately stayed in the background to emphasize Ghanaian leadership of the project. IPA staff at all levels oriented their approach to working with GES to focus on

how they could first and foremost help GES to successfully implement the project, such as working to find local education NGOs that could help develop teacher training materials and troubleshooting data analysis questions. The effect was a deepening trust between the IPA team and GES as they saw that IPA was genuinely committed to GES' success and ownership over the project. As the project progressed, IPA continued to make decisions, such as inviting GES leadership to present on TCAI's progress at conferences, that put the Ghanaians front and center. The trust that developed between IPA and GES was the most cited reason for the success and impact of the TCAI project.

Another important, ongoing aspect of the project was the extent to which GES staff, both across the country and at all levels of the agency, were involved in TCAI. TCAI was designed so that it was primarily implemented by GES' existing staff, and this structure, combined with the national scope of the project, resulted in a significant number of GES staff being personally involved in the work of TCAI. On top of this, there was a deliberate effort by Stephen Adu and his team to engage GES senior leadership in TCAI. For example, we heard from senior GES staff about the strong impression made

by their monitoring visits to TCAI schools, and how those visits were key to their own belief in the potential of the program. These types of involvement are significant because they helped to disseminate awareness and experience with TaRL across the Ghanaian education system, and provided opportunity for firsthand experience of the impact of the program. As a result, TCAI had an elevated level of support among GES' senior leadership throughout implementation, and organizational knowledge about the concept of TaRL was built across the country and throughout the organization.

One important aspect of engaging GES staff throughout the organization was the project's use of midline data that was collected between May 2011 and April 2013. Sharing this data in an ongoing way was critical to building trust with GES staff throughout the project. GES employees became more invested in the project data once they realized that it revealed new insights. For example, many seasoned education professionals were surprised that the after-school program was having positive results—or even worked at all—and that it worked better than the in-school program. Initially, many educators had theorized that students would not stay after class or would be too hungry or tired



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“I think TCAI is special, in a way. They are not like a university or a research institute that only needs you for data collection and then goes off for two years and comes back. They are always working.”

IMPLEMENTER

to learn after school. The sharing of interim results was one significant way for the project to maintain and increase buy-in across the various levels of GES that were involved.

As the project's first endline evaluation concluded, the project was faced with a significant decision: how to respond to feedback from teachers implementing the program about the difficulty of a key aspect of the design. The steering committee's response demonstrates the way in which the project focused on adaptability, and was able to make trade-offs in favor of the programs results, rather than favoring the “purity” of the evaluation design.

As mentioned, the teacher-led treatment condition was a high priority for GES, which wanted to understand how its existing staff could implement TaRL concepts. When the project started, the teacher treatment required that teachers divide their students by learning level and design lesson plan variations that were aligned to the needs of each group in their class. This required significant effort from teachers to design multiple lesson plans and to manage their classroom in small groups. The effort was causing complaints from teachers, and GES wanted a more practical version of the targeted lessons for the teacher-led intervention. Despite the potential that

changing the treatment in midstream could lessen the rigor of the evaluation, IPA agreed to the change, so that teachers could now divide their students by learning level and then send them to different classrooms, requiring each teacher to now work with only one level.

In this case, the decision to change a treatment condition supported the project's objective of having interventions that were practical for Ghanaian teachers, while at the same time complicating evaluation of the effect of the teacher-led treatment by changing it mid-course. These decisions further deepened trust between IPA and GES, by prioritizing the impact that TCAI would have on participating teachers and students, and reassured the government that the TCAI program was feasible within their existing system.

The End of TCAI

Unfortunately, one of the decisions designed to bolster the sustainability of TCAI ended up ultimately undermining its success and viability for long-term scaling up. The use of the National Youth Employment Program (NYEP) was designed to leverage an existing government program to reduce the cost of TCAI's community assistants and set the stage for scale if the program proved successful.



However, it also meant that important administrative controls, like payroll for community assistants, were the responsibility of NYEP and not directly under the control of GES or IPA. This caused serious challenges early in the project when paychecks were delayed, affecting community assistant retention and morale. Because of their reliance on a separate government agency, the TCAI team had to try their best to cajole NYEP to resolve the payroll issues.

In the second half of TCAI, a second, graver issue arose with the NYEP, which undermined the impact and scalability of TCAI. During the second year of the project a scandal at the NYEP caused the minister in charge to resign and resulted in a significant shake-up in the structure and objectives of the NYEP program. This shake-up again disrupted payments for many community assistants and derailed plans to use NYEP as a pipeline for full-time certified teachers and as a vehicle for full scale implementation of TCAI. While the full effect of NYEP's scandals cannot be known, there were reports of community assistants abandoning their posts as the financial support from NYEP was disrupted, and uncertainty about long-term employment opportunities increased. The shadow of the NYEP scandal clouded the final months of TCAI implementation and likely had an effect on the results of IPA's evaluation.

In the summer of 2013 IPA administered the second endline survey for TCAI, and its analysis showed that the program significantly improved children's basic skills in numeracy and literacy. Of the four treatments that were tested, the in-school and after-school remedial teacher community assistants (TCAs) had the largest impact on average achievement. IPA's analysis found that the effect of treatment persisted for students after one year, but found that there was significant variation in impact across regions, which appeared to be correlated with variations in implementation quality. Specifically, the monitoring of the project showed that in some regions the program was being implemented with remarkably low fidelity to the design. These gauges of implementation quality were significant cause for concern that the government would be unable to effectively scale up the interventions tested in TCAI, despite their demonstrated positive results.

While the evidence from the study was positive, it was still disappointing from the perspective of the Ghanaian government. IPA found that although it was slightly cheaper to provide either small group instruction through normal curriculum TCAs or targeted lessons through classroom teachers, these interventions did not improve test scores as much as the combination of teaching at a child's level and providing community

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love



assistants. The government's preferred, teacher-led intervention had demonstrated the lowest impact, although researchers attributed the relative low effect to teachers' low compliance with the program.

In the year after completion of TCAI's implementation, Maame Nketsiah, TCAI's Implementation and Policy Director, led an effort to travel to the participating districts and share the results of the evaluation.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

There was significant interest among local GES leaders in the results of the study since it provided previously unavailable data on students' learning levels, as well as important administrative data such as teacher attendance and other measures of teacher performance. According to participants, local leaders were quite engaged in these conversations and concerned about the low learning levels of their students, but their primary concern was how the findings from TCAI could be operationalized locally. While Ghana had made efforts to decentralize the school system, there was very little actual discretion for decision-making by local schools to implement changes based on the lessons from TCAI. Without support from the Ministry of Education, it was hard for local schools to implement changes that would require new incentives to compel teachers to take on additional responsibility, or coordinate local community assistant programs based on the TCAI model.

At the national level, the decisions on how to proceed after TCAI concluded were constrained on several fronts by limited resources. First, the shake-up at NYEP meant that a major co-implementer of TCAI had been reconfigured and the associated resource commitment withdrawn. At the same time, concerns about the varied implementation

“The whole nuance about the ways that we have advanced the agenda of evidence-based policymaking through TCAI, is to me quite strong.”

RESEARCHER

quality raised doubts that scale could be achieved in an effective and cost-efficient manner. Also, the potential for new USAID funding around education diverted the Ghanaian government's attention from TCAI. The confluence of these factors meant that government resources were not available to scale up to create a national program, or to help local communities that might want to adopt elements of the TCAI model. IPA and GES worked to share the findings and

cultivate support from the international donor community. These efforts to engage the donor community continue to this day, but haven't yet resulted in funding commitments.

Four years after the final endline study, many of the senior GES staff involved with TCAI remember the program as a success both in terms of the effect it had on Ghanaian students, and on the way in which the government and researchers were able to collaborate to test an impactful program at a significant scale. Despite the problems that led to the end of the program, there is anecdotal evidence that several schools continued to implement some aspects of the TCAI program on their own, without resources from GES or elsewhere. For example, one district reported that teachers from a treatment school had instituted their own “Learning Hour” to provide remedial tutoring for an hour a day.

While no formal measure was taken of how many schools continued to use TCAI materials or methods after the end of the program, we heard from several stakeholders that the program had an effect on students that participated and also on the way that teachers and administrators thought about the value of remedial education.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“Success,” “Failure,” or ?

In considering the ways in which TCAI was a “success” and/or a “failure” in translating evidence to policy and practice, it is important to consider several angles and ask: what was the purpose of TCAI? At the highest level, the project had two ambitions: to rigorously test the concept of Teaching at the Right Level with a national government and to lay the groundwork for long-term strengthening of the Ghanaian education system. The significant effect demonstrated by the TCAI RCT achieved the project’s first ambition, but at first blush the second ambition has not as clearly been met. Despite TCAI’s significant effect, the original commitments to scale the program nationally have fallen through as the collapse of the NYEP and unavailability of funding for scaling have thus far prevented TCAI’s from being adopted nationally. Despite this apparent failure to directly scale TCAI or create large-scale change within GES, there is nonetheless reason to believe the project had significant, if indirect, impact within Ghana and in IPA’s ongoing work, and that it contributed to a larger national, and even global, conversation on the importance of differentiated learning.

Traces of TCAI are still evident in Ghana, and the IPA country team is actively working on new opportunities to scale the evidence for TaRL in Ghana and other locations, including Zambia and Peru. In addition to the spread of the TaRL concept, TCAI has influenced IPA and its relationship with the Ghanaian government, and set the stage for new practices and attitudes in Ghana regarding evidence-based practice and remedial education. IPA has sustained its institutional knowledge of TCAI, but the institutional knowledge of TCAI is starting to fade at GES as senior leadership retires and other initiatives take on prominence in the country. As IPA and others consider the future of TaRL in the country, it is hard to know whether the potency of the indirect effects of TCAI is waxing or waning.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part II: **Key Themes + Insights**

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

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.

Randomized controlled trials were the standard of evidence for enabling and validating TCAI. This “gold standard” of evidence was the basis by which TaRL was evaluated in India and Kenya, and the method chosen by the IPA team and Steering Committee to determine whether the TCAI project was effective. It is also important to note the observation shared by some participants that education is unique compared to agriculture or public health, in

the sense that there are fewer accepted standards of evidence, and there are stronger ideological perspectives among donors and governments on what priority should be place regarding different educational goals. This ambiguity can create significant “noise” regarding what constitutes robust evidence and how the government and others should determine priorities.

Despite TCAI's emphasis on rigorous statistical significance and a robust counter-factual for evaluation, it is important to step back and consider the project in the context of TaRL as a concept that has been developing over time. TaRL as an approach is broader than any individual example of “evidence” of successful implementation, and the cumulative evidence from several projects serves less to isolate a single key intervention, but rather to broaden and deepen the field's understanding of how and why targeted instruction based on learning level can improve educational outcomes. We heard from several interviewees that

TCAI was instrumental in reinforcing the concept of remedial education within GES, an effect which we observed in talking to several different participants in the project. In addition, the project was also part of a broader global conversation on the importance of differentiated learning and remedial education.

Despite the official emphasis on RCTs as evidence, anecdotal evidence, local knowledge, and direct personal experience played significant roles in both launching the project and the longer-term effects of TCAI in Ghana. For example, visiting Pratham's program in India was instrumental in convincing GES' senior leaders and others that the program could be effective in Ghana. Also, a deliberate effort over the course of the project to have GES' senior leadership observe implementation first-hand was critical to their assessment of the project's success and continued support. Last, the anecdotal evidence that a number of schools continued to use TCAI interventions after the program ended suggests that, even without the results of the RCT, local school staff were confident that the benefits of the program were worth their time and effort. These direct and anecdotal forms of evidence have had a much more lasting and potent effect on those

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.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

involved in the project than the results of only the RCTs, and may be the most significant evidence generated in terms of the long-term influence of TCAI.

Trust-building

The cross-stakeholder collaborations required for evidence-informed policies and practices are often difficult to initiate, develop, and sustain. Particularly when institutional incentives are lacking, personal trust, respect, and buy-in between individuals across stakeholder groups become critical to fostering the effective flow of evidence into practice.

As described in the case study narrative, trust-building among the various project stakeholders was critical for getting TCAI launched and for cultivating Ghanaian ownership of the project over time. Intentional efforts by IPA and by Children's Investment Fund Foundation before, during and after the project to focus on relationship-building with the Ghanaian government left a lasting impression among GES staff that TCAI was an exceptional development project, characterized by genuine collaboration.

This approach to trust-building relied on a combination of both who was involved and how the project was designed and

The design, development, and management of TCAI demonstrates the importance of mechanisms to establish and maintain trust among stakeholders and the ways in which early involvement and subsequent ownership by policy makers and government implementers is an important precursor to long-term adoption and effective implementation at scale.

implemented. Who was involved had three important dimensions: design, governance, and staffing. On design, the negotiation of the final project design allowed for each party to have their priorities integrated, and developed a shared ownership over its design. On governance, the project sought to involve a wide variety of stakeholders

that were both champions of the project and skeptics, and developed a steering committee that included these stakeholders in decisions and created a forum for critical input. Practical considerations like giving GES direct control over the project's funds, having the Implementation Director report directly to the Steering Committee, and adjusting



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

the project's design over time to address the needs of the government, were concrete examples of flexibility and commitment to partnership. From the staffing perspective, IPA made a deliberate choice to hire mostly Ghanaian staff, and to partner with local Ghanaian organizations (e.g. NGOs, vendors, etc.) whenever possible. This effort to put Ghanaians at the heart of the project on both the government and research side was an important demonstration of commitment to local ownership and trust-building. In addition, the IPA and GES teams invested significant time in developing genuine personal relationships, co-locating their staff when possible, and going above and beyond to troubleshoot day-to-day challenges. This “all-in” approach to the project and willingness on IPA's part to let go of significant control in favor of influence was key to building trust on both sides.

The effect was significant in two ways. First, it resulted in greater commitment from both IPA and the government, which was critical for addressing minor and more significant challenges that came up over the course of implementation. Second, it provided the foundation for a lasting relationship between the Ghanaian government and IPA, which has led to open lines of communication and further collaborative projects around analysis

of government data. Trust-building was key to successfully pulling off a complex project, and doing so in a way that laid the foundation for further collaboration.

The shadow of the NYEP scandal clouded the final months of TCAI implementation but there were sound reasons why the project sought to partner with NYEP, and its demise likely could not have been predicted. The difficult decision to embed TCAI within other government initiatives bolstered the chances for long-term, sustainable implementation of TCAI, but it also exposed the project to risk outside of its direct control, and proved in the end to be an insurmountable barrier.

Time and Resources Required for Integrating Evidence

Few organizations provide incentives or carve out explicit time for managers to explore emerging evidence in their field. Even fewer assign staff to find relevant evidence and translate it into accessible formats for the organization. As a result, the role of preparing and sharing evidence that is timely, useful, and relevant for practitioners is sometimes explicitly played by formal intermediaries (e.g., certain think-tanks). More frequently, an actor who holds a formal role within another stakeholder group spontaneously takes on the (additional) responsibility for trying to





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

integrate evidence, with no actor formally responsible for the process. Discovering and integrating evidence requires time, energy and funding.

TCAI is an interesting example of a relatively small and new research organization able to leverage significant funding and develop relationships to influence government partners. IPA appears to have short-circuited the typical resource and relationship dynamics in international development by building personal relationships within GES, cultivating a significant international donor as a partner, demonstrating commitment to accommodating government priorities, and building government buy-in and ownership of the project. In addition, the credibility of J-PAL and IPA's relationship in India with Pratham and the strength of the existing evidence for TaRL increased its ability to bring other stakeholders on board. While raising outside resources to fund the project was essential, TCAI demonstrates how other factors, such as strong relationships and willingness to share control can be used to overcome typical resource dynamics. In this case, giving GES direct control over the project funds and putting them front and center on project implementation was important in getting the program launched.

“It’s not necessarily a translation from one physical context to another but from the context of one implementer to another. But the exact same questions arise.”

RESEARCHER

Another significant resource consideration is the investment that IPA made in taking on a supporting, advisory role in project implementation and in relationship building with GES staff. The collaborative approach that characterized TCAI involved much more time and energy on direct relationship building in order to successfully support implementation through influence, rather than direct control or a procurement-style relationship.

The last important consideration regarding time and resources is the role that resource availability played in the failure of the project to scale following evaluation. While a variety of factors, especially the reconfiguration of the NYEP, affected the opportunity to scale the evidence from TCAI, the lack of new, dedicated resources to support adoption was also a key barrier. This experience illustrates the significant challenge that exists in transitioning project-based evidence generation to become institutionalized in more sustainable and long-term modes of implementation.

Operationalizing Evidence

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.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“What we’re scaling up is not TCAI, actually. What we’re scaling up is targeted instruction, teaching at the right level.”

RESEARCHER

TCAI reflected a deliberate choice to evaluate the challenge of implementing TaRL concepts at scale, which required engaging with government as an implementing partner. This approach to both building the evidence base and operationalizing evidence presented potential trade-offs, due to the understanding among development organizations that government is often a less effective implementer than private NGOs. It is important to note that despite serious variation in the quality of implementation by the government, TCAI was still able to demonstrate a measurable effect. At the same time, there were also important effects on the participating organizations themselves.

TCAI depended on multiple independent government stakeholders, such as NYEP, with disproportionate levels of buy-in. This exposed TCAI to substantial risk from

management issues at NYEP, which had detrimental effects on TCAI, even though they did not prevent the community assistant treatment arms from having a measurable effect. Despite these effects, it is important to note that the government’s quality of implementation created serious doubt regarding its ability (with or without a stable NYEP) to effectively scale up implementation of TCAI. While the challenges associated with funding and administering the NYEP program reflect the types of ongoing constraints that a TCAI-like program would encounter at scale, they were not the only issues of management capacity facing the government.

For GES, implementing TCAI involved staff from across the country and various levels of the agency, from senior leadership to classroom teachers. This broad engagement led to a threefold impact:

- Wide exposure to new pedagogical theory, training, and teacher materials on remedial instruction
- Access to previously unavailable data on the operations of the education system (e.g. teacher attendance, school geolocation, etc.), and more regular data on student outcomes (e.g., the project’s midline survey results)
- Exposure to practical training on collecting and using data from schools

These structural after-effects of TCAI demonstrate the ways in which translation of evidence can create new structures and capabilities that persist even if the specific intervention doesn’t scale, and which potentially provide a foundation for other evidence-informed policies and practice in the future.

It is also important to note the way in which the project used the wealth of administrative implementation and evaluation data it generated. This use of evidence for performance management of the project illustrates the ways in which evidence can be deployed at more intermediary and local levels, rather than as just a global, retrospective measure of outcomes. It also suggests the importance of *who* information



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

from the project is shared with and *how* the information is presented (e.g., the relative and absolute performance of districts). These types of uses of project data and the new types of routines and decisions that they enable should be considered as another, indirect effect of using evidence-based practice.

At IPA, the TCAI experience was formative in shaping how the organization approaches partnering with policy makers and governments, and focuses more deliberately on the challenges of operationalizing evidence of effective practice. One of TCAI's consequences has been that IPA is investing more in its policy advocacy role, which allows it to better influence how governments match evidence and research with pre-existing priorities. While IPA has always focused on using RCTs to influence policy, the TCAI experience accelerated a focus also on developing a better understanding of the priorities of various public stakeholders and using those to inform and coordinate research priorities. Working on TCAI also helped IPA's predominately research-oriented staff to better understand the practical challenges of program implementation, a perspective that has been increasingly important as IPA operates as an intermediary for connecting rigorous research to practice.



Incentive 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.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

TCAI provides an interesting case for exploring how the incentives of policy makers and evidence producers can diverge, and some of the ways in which these incentives can be reconciled or aligned. Ultimately, the willingness of IPA and GES to undertake shared implementation, and the willingness of IPA to find a balance between its desire for a rigorous academic experiment and pragmatic program implementation were key to reconciling their incentives. The Ghanaian government was eager to partner with a project that aligned to its existing strategic priorities, while also attracting international philanthropic money to augment the effectiveness of GES and to direct resources to areas of clear student need. TCAI met these priorities by addressing a known gap in student educational attainment, and focusing on the role of government as an implementing partner. This alignment, as well as the personal support of champions with political capital within GES, gave TCAI an advantage in getting launched.

A key set of potentially conflicting incentives for IPA revolved around the importance of remaining credible and connected to the international academic research community. TCAI created a tension between wanting to make sure that the project achieved results and was aligned to government's needs,

and also being able to credibly say that it constituted a rigorous evaluation. Ultimately, IPA showed flexibility to the exigencies of pragmatic implementation, while ensuring that it maintained a rigorous evaluation. As discussed in the case narrative, the ability to tread this balance was facilitated by the fact that IPA and the principal investigators for the project were insulated from some of the traditional academic incentives (such as publishing and tenure considerations). The desire to advance the body of evidence around TaRL also created an additional incentive to work effectively with government to determine how the project could be most practically successful.

By placing the role of government at the center of its research question TCAI was able to partially reconcile the sometimes-conflicting incentives for policy makers, implementers and evidence producers, but not without significant effort and negotiation to arrive at a project design that worked for all parties, and sufficient trust between IPA and GES to navigate the ongoing implementation of the project and evaluation.

Responsibility for Evidence Translation

IPA was the driving force behind incorporating TaRL concepts into practice in Ghana, and IPA took on a significant role in project

implementation. Nonetheless, IPA views TCAI as an example of a project where it played a collaborative and influence-oriented role in implementation, rather than more direct control. Working to effectively transform TaRL into TCAI required IPA to give up some of its authority as a "pure" evidence producer, by inviting local stakeholders into genuine partnership and considering a give-and-take approach to design. This shift from pure evidence producer to combined producer and intermediary included three important aspects for IPA:

- Allowing local partners, such as GES, to make decisions about project design that met their needs, but were different than the researchers' priorities (e.g., including teacher-led treatment arms);
- Earning the trust of its Ghanaian government partners by giving them genuine ownership and leveraging local expertise (e.g., partnering with Ghanaian education experts and a local NGO to develop project materials);
- Choosing when to push back and spend the time and energy to convince its partners that certain aspects of the project were essential (e.g., pushing for more precise assessments to accurately determine student's learning levels).



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

TCAI illustrates the ways in which the translation of evidence into practice can be best thought of as a negotiation between the various incentives and priorities of a broad set of stakeholders. The more that stakeholders are proactively and explicitly engaged in this negotiation, the better the chance for successful translation and execution of evidence-based practice. To support this type of negotiation, however, an evidence-based concept needs to be both robust and malleable enough to allow for meaningful negotiation among stakeholders. TCAI shows how proactive negotiation of evidence-based projects can create important buy-in and ownership of projects, which is essential for resilient implementation at scale, and important for establishing the foundation for longer-term sustainability of a specific program and its broader, evidence-based concept.

One critical enabling aspect of this negotiation was IPA's decision to play the role of both researcher and intermediary, and focus on empowering the Ghanaian implementing partners. This unique combination of priorities made TCAI an opportunity to explore an intermediary role for IPA that focused on collaborative implementation with government, which it continues to play in more and more of its projects. For example, IPA increasingly has

**TCAI shows
how proactive
negotiation of
evidence-based
projects can create
important buy-in
and ownership of
projects, which
is essential
for resilient
implementation
at scale.**

policy-focused staff based in their local offices, who are focused on understanding local priorities and the opportunities to align evidence, as well as using the same perspective to set research priorities and shape the work of academics with whom they partner.

TCAI also illustrates the vital role of champions in translating evidence to practice. One

TCAI participant described a champion as someone who has the intelligence and resourcefulness to affect a project's overall outcomes and the willingness to work "for free" to make the project a success. In this sense TCAI had a number of champions in government and at IPA who were instrumental in getting the project launched and shepherding it through implementation. These catalysts for the project, despite the many complications and barriers, were critical for launching TCAI, bringing on government as an implementing partner, and helping to overcome roadblocks during implementation. TCAI illustrates the importance of having key participants who are knowledgeable about the intervention, personally committed to the project's goals, and possess the political capital required to bring other critical players to the table and influence how resources are deployed.

In addition to the early champions such as Stephen Adu, project managers at GES deliberately cultivated internal champions by engaging senior GES leadership during project implementation. These efforts, which included trips to observe TCAI's implementation in schools, were important for developing buy-in from across GES leadership, which in turn helped to reinforce that TCAI was an organizational priority and built awareness and support for TCAI.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

During implementation, another kind of champion was critical to pulling off TCAI, and doing so in a way that built trust among stakeholders. The IPA implementation project leadership adopted a “whatever it takes” approach that was critical to demonstrating to GES that IPA was committed to the success of the project and troubleshooting both the day-to-day and strategic challenges that arose over the course of the implementation.

Conclusion

As described in the case narrative, there are several ways to consider the question of whether TCAI was a success or a failure in integrating evidence into practice. On one hand, the results of the TCAI RCT were consistent with the effects observed in other countries. On the other, the original intention to dovetail TCAI with long-term and fundamental changes to the Ghana education system did not pan out. In general, the Ghanaians we spoke with described TCAI as a success in terms of the impact it had for students as well as the way that the project was executed, in spite of the fact that TCAI has yet to be scaled nationally in the country.

Perspectives among IPA staff on the project’s success varied considerably. These varying definitions of success, particularly within the same organization, highlight the

several different objectives of intermediary organizations. While TCAI failed to deliver the immediate changes to the Ghana education system that the project’s leaders had desired, it did have a number of indirect effects on GES and laid the groundwork for an ongoing relationship between IPA and the Ghanaian government, which is important for long-term influence. The question of whether this constitutes “failure” highlights the tension between creating direct impact through operationalizing evidence, and the more indirect, systematic change that intermediaries can also help bring about.

If we consider how TCAI helped to spread the concept of remedial education and enabled the use of the concept of learning levels to inform instruction, there is promising anecdotal evidence that the project has had impact. During our time in Ghana we heard from stakeholders at IPA and GES that the TCAI project exposed many educators across GES to new ways of thinking about remedial education and gave them tools to assess student need and adapt their teaching to the needs of different learning levels. Finally, there is still significant opportunity that TCAI will lead to further TaRL work in Ghana and in other countries dealing with issues of remedial education. As mentioned, IPA has since led efforts to apply TaRL concepts in Zambia and Peru.

In addition to considering the spread of the concept of TaRL, we heard of other, indirect effects of implementing TCAI. As mentioned, the use of administrative data and new processes was a valuable indirect effect of implementing TCAI. Another, more centralized effect has been a deepened relationship between IPA and GES, based on GES’ desire to increase its capacity for data analysis and using evidence in policy formulation. GES’s experience of collaborating with IPA has made it a trusted partner of the government, allowing for ongoing collaboration on how GES can better use data, and setting the stage for more evidence-informed collaboration in the future.

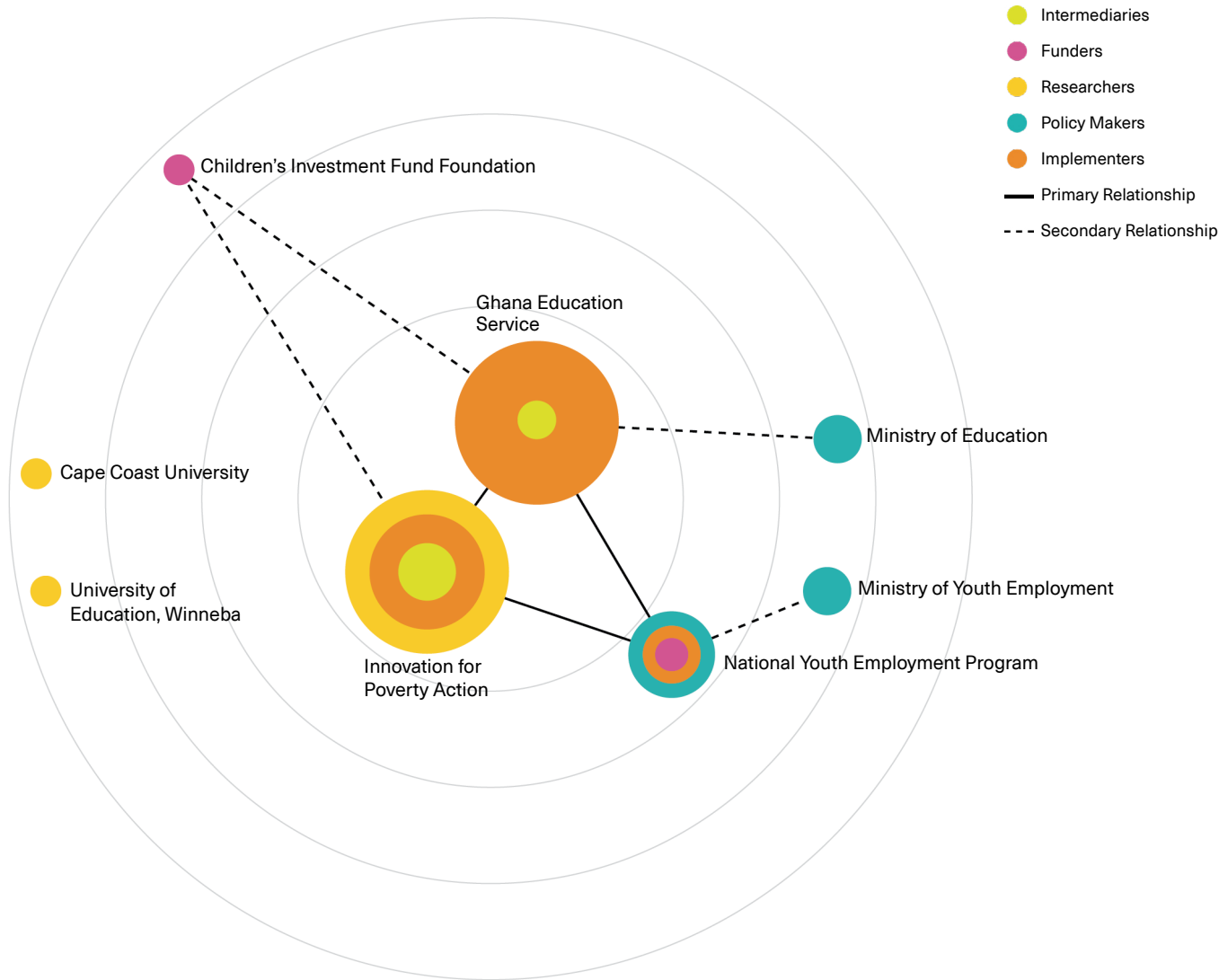


Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Stakeholder Map

The two primary stakeholders in the Teacher Community Assistant Initiative were the Ghana Education Services and Innovations for Poverty Action (IPA). Both of those organizations played multiple roles as implementers and intermediaries, with IPA acting also as the leading research stakeholder in the project.

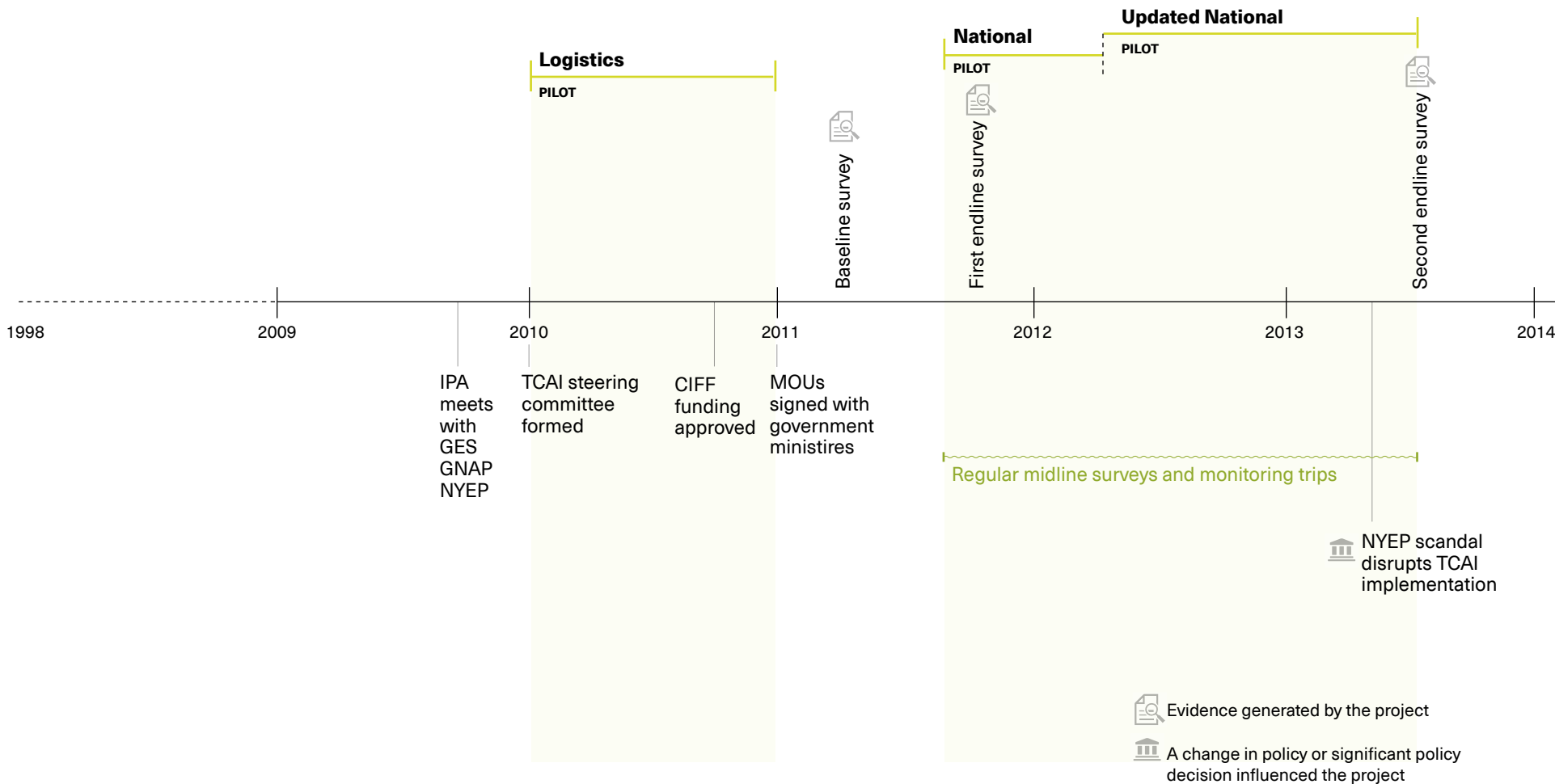
This stakeholder map is a visual representation of the major stakeholders involved with this project. The importance of each of the actors is defined by their relative size, and their proximity to the center of the project. Their role is defined by the color; multiple colors indicate multiple roles. 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.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Timeline

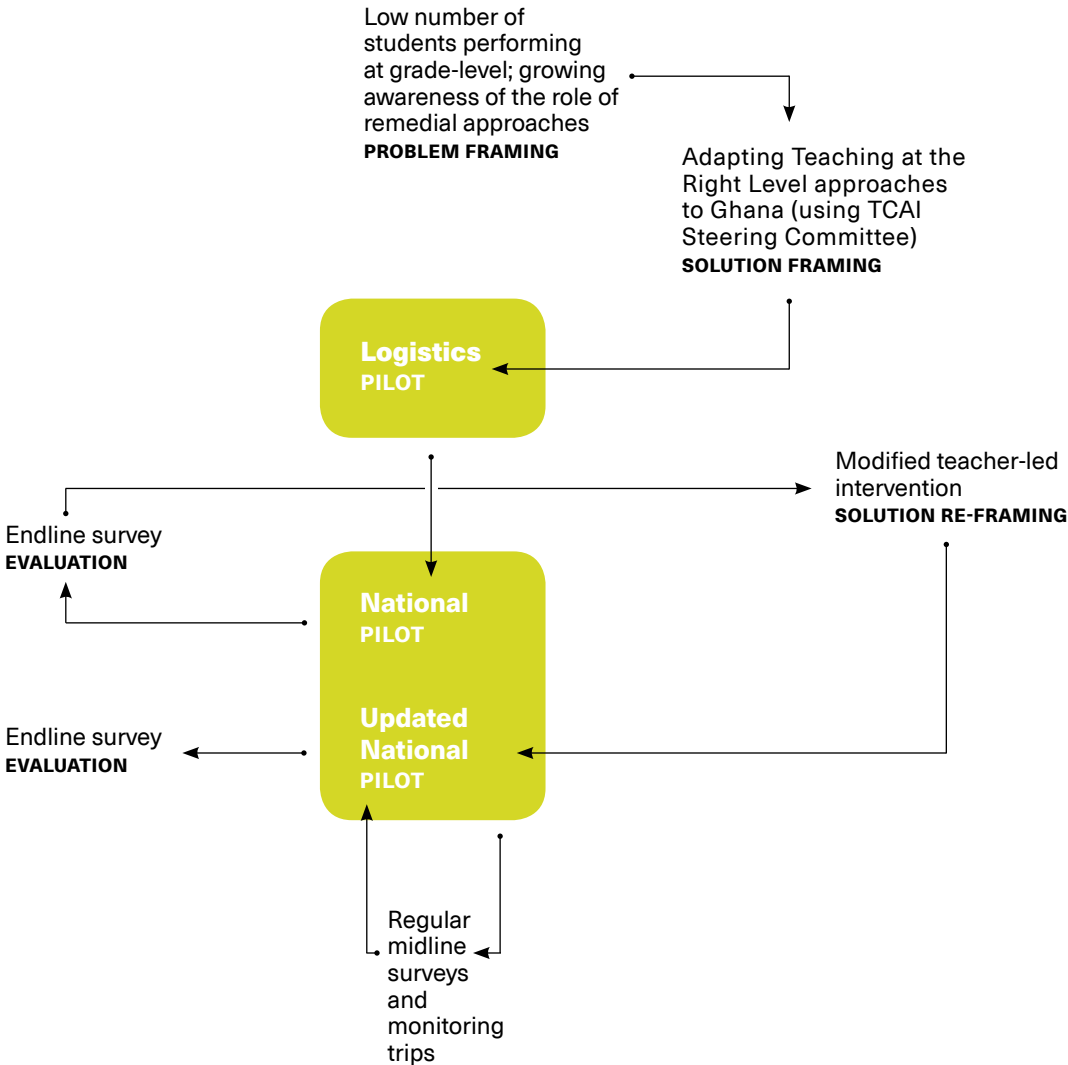




Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Process Diagram

TCAI never **scaled**, but its methods **spread** to other countries and **structured** new data capabilities at Ghana Education Service.





**Evidence in
Practice**

Full Report

Print

Case Studies

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Page **70** of **220**

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





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part I: **The Story of GUP**

The Graduating the Ultra Poor (GUP) project was a randomized controlled trial (RCT)³¹ 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*³² 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.”³³

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.

³¹ 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.](#)

³² 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.](#)

³³ *Ibid.*



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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:³⁴

- Identify the most vulnerable households within a community;
- Provide regular and time-bound consumption support to enable them to meet basic needs;
- Help families plan their livelihoods and then transfer appropriate productive assets to them;
- Develop their ability to save money as a tool to build resilience;
- Enhance their technical and entrepreneurial skills through livelihood training;
- 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 1.³⁵

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.³⁶

³⁴ “The Graduation Approach,” UNHCR. [Access here.](#)

³⁵ “Graduation Pilots Overview,” CGAP-Ford Foundation Program. [Access here.](#)

³⁶ “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



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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.³⁷ 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.³⁸ 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.³⁹

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.⁴⁰

The full GUP treatment arm was comprised of the following six components:⁴¹

- **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.
- **Health:** Households were enrolled in the National Health Insurance Scheme and received health and nutrition education.
- **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.
- **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.
- **Savings Account:** Households opened savings accounts and PAS staff collected deposits on a weekly basis.
- **Household Visits:** Through the program, PAS staff made weekly visits to each household to provide monitoring, coaching, and encouragement.

³⁷ 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).

³⁸ “Graduating the Ultra Poor in Ghana,” *Innovations for Poverty Action*. [Access here](#).

³⁹ 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](#).

⁴⁰ Further description of each treatment arm can be [found in Appendix I](#).

⁴¹ “Graduating the Ultra Poor in Ghana,” *Innovations for Poverty Action*. [Access here](#).



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“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.”

RESEARCH

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.⁴² 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.⁴³ (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.

⁴² *Ibid.*

⁴³ 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



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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)⁴⁴ and the Ghana Poverty Reduction Strategies of the early 2000s.^{45,46} 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).⁴⁷ 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.

⁴⁴ 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.

⁴⁵ 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.

⁴⁶ 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](#).

⁴⁷ *Ibid*.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

some of the Millennium Development Goals,⁴⁸ the three northern administrative regions still “harbor the poorest of the poor.”⁴⁹ 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.⁵⁰

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.

48 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.

49 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.](#)

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



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“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).

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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.”



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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.”⁵¹ In 2014, CGAP and Ford Foundation published a technical guide to implementing Graduation programs in various contexts, entitled “From Extreme Poverty to Sustainable Livelihoods,”⁵² 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.”⁵³

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.

⁵¹ S. Hashemi and A. de Montesquiou, “Reaching the Poorest: Lessons from the Graduation Model,” CGAP, March 2011. [Access here.](#)

⁵² A. de Montesquiou and T. Sheldon, “From Extreme Poverty to Sustainable Livelihoods: A Technical Guide to the Graduation Approach,” CGAP, September 2014. [Access here.](#)

⁵³ The graduation work has since been transferred from CGAP to the Partnership for Economic Inclusion (PEI) at the World Bank.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“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.⁵⁴ 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.⁵⁵

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,

⁵⁴ “Graduating the Ultra Poor in Ghana,” Innovations for Poverty Action. [Access here.](#)

⁵⁵ *Ibid.*



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

(and -198% in Honduras).⁵⁶ 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

⁵⁶ 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.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love



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



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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 PSNP's nearly 8 million beneficiaries have been included in activities modeled after the Graduation pilot.⁵⁷

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

⁵⁷ Matt Hobson, "Pathways to Graduation: A Work in Progress in Ethiopia," World Bank, December 2014. [Access here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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,



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Stakeholder Map

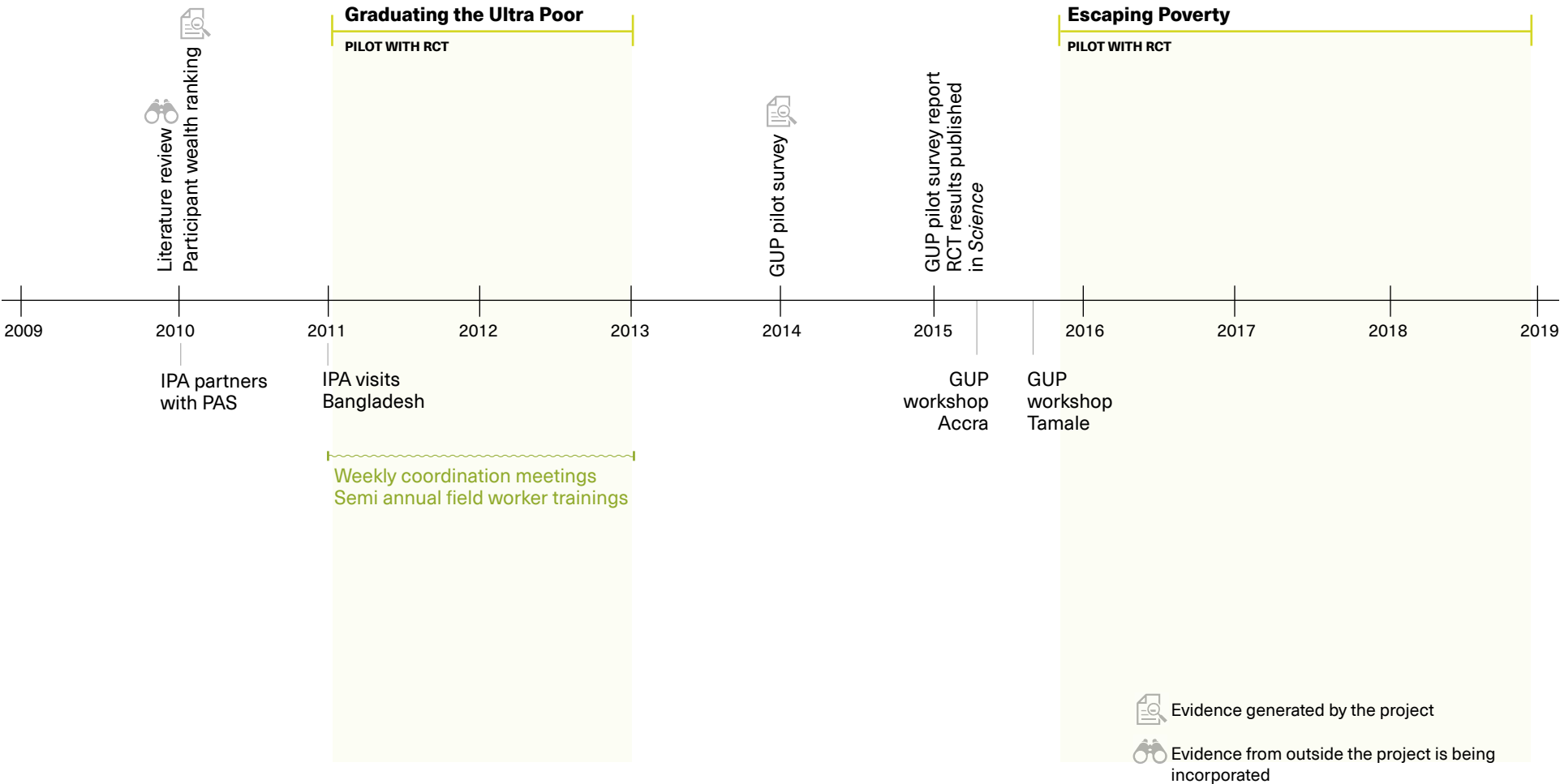
This stakeholder map is a visual representation of the major stakeholders involved with this project. The importance of each of the actors is defined by their relative size, and their proximity to the center of the project. Their role is defined by the color; multiple colors indicate multiple roles. 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.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Timeline

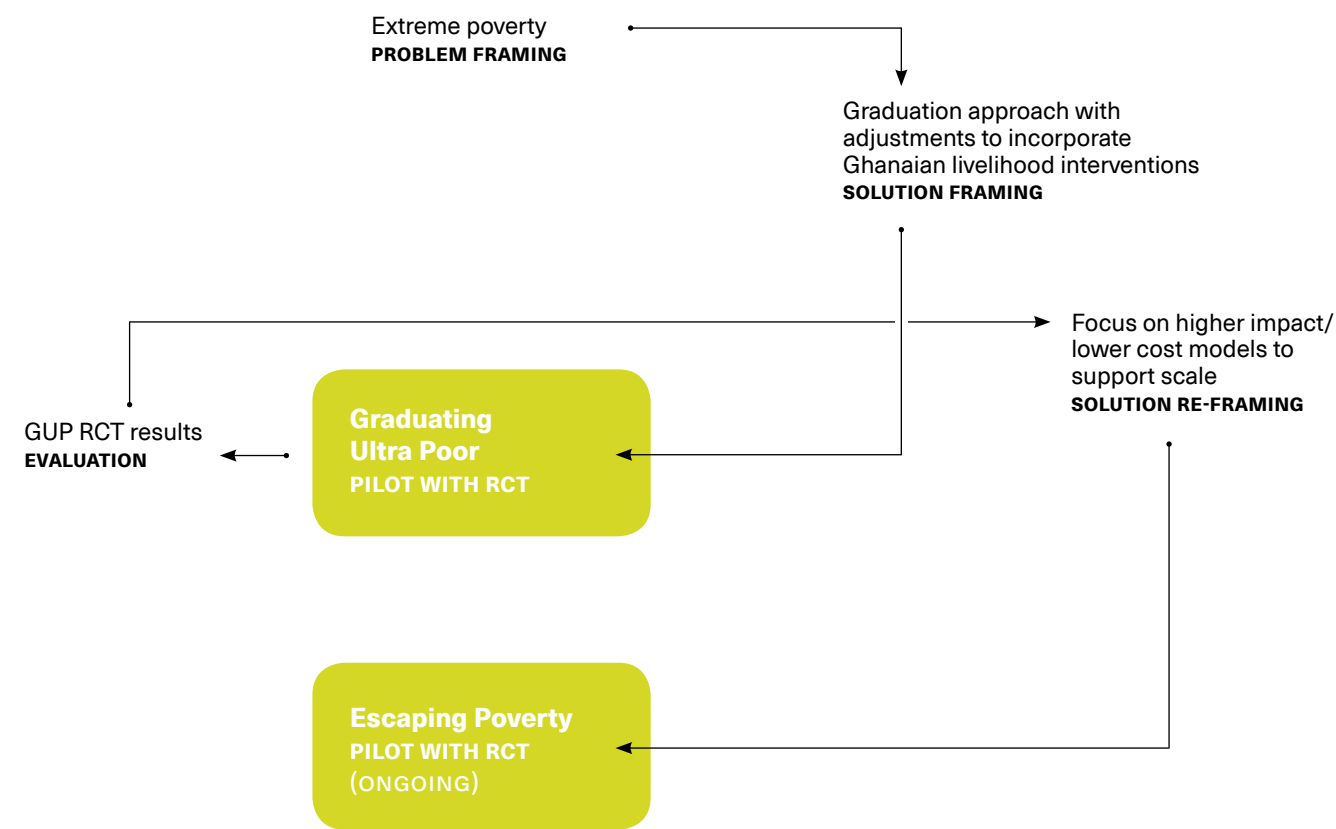




Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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.





Appendix 1

The full GUP treatment arm was comprised of the following six components:⁵⁸

- Consumption support: During the lean season (14 out of 24 months), households received weekly cash transfers.
- 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
- Productive asset transfer: One-time transfer of a productive asset, such as goats or guinea fowl.
- 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.
- 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.
- 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.

⁵⁸ <http://www.poverty-action.org/study/graduating-ultra-poor-ghana>



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



**Evidence in
Practice**

Full Report

Print

Case Studies

Page 95 of 220

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

INDIA | EDUCATION

Teaching at the Right Level

Pratham Education Foundation, one of India's largest educational non-governmental organizations, is recognized globally for its evidence-based pedagogical approach.

By Shira Beery





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part I: **The TaRL Story**

Pratham Education Foundation (Pratham), one of India's largest educational non-governmental organizations, is recognized globally for its evidence-based pedagogical approach, dubbed Teaching at the Right Level (TaRL) or Combined Activities for Maximized Learning (CAMaL, which means *amazing* in Hindi).⁵⁹ This case study explores how Pratham designed and scaled TaRL, from the perspective of an evidence-in-practice lens, and an analysis of the factors that encouraged different types of stakeholders to participate in the process.

Beyond TaRL, Pratham implements a diverse set of programs with funding from a range of individuals, corporations, foundations, and governments. Pratham is one of the largest NGOs in India with an annual budget of over US\$30 million.⁶⁰ Today, Pratham's programs serve children, young adults, and adult high school drop-outs across the country.

Background

Both academic and practical evidence have always shaped Pratham's approach. The 1990 UNESCO World Conference on Education for All in Jomtien, Thailand—emphasizing universal access to primary education and learning outcomes⁶¹—played an important role in Pratham's inception.⁶² In the years following, UNICEF and the Mumbai Municipal Corporation partnered to address education access in the city. The partners recruited

Dr. Madhav Chavan and Farida Lambay, who had research backgrounds and experience working in Mumbai slums, to conduct a secondary analysis using government data to estimate how many children were out of school in Mumbai.

The two researchers knew from experience that, in Mumbai slums, government data overestimated the number of unenrolled school-aged children. Through their own surveys and interviews with local teachers they learned that, in fact, children's lack of preparation for first grade was the more pressing issue. Rather than focus exclusively on enrollment, Chavan and Lambay concluded that pre-school programs—globally recognized for reducing the risk of dropping out of school—were needed to prepare children for success in school.⁶³

This was the genesis of Pratham's mission statement: "Every child in school and learning well."

With support from UNICEF and the Mumbai Municipal Corporation, Chavan and Lambay launched a network of low-cost, community-based pre-schools. In 1995, Pratham was incorporated as a charitable trust and by 1998 it ran 3,000 pre-schools.⁶⁴ During this period, Chavan was exposed to the concept of remedial education while visiting a school in Australia, where recent immigrants or children lagging behind received additional instruction

⁵⁹ In this case study, for simplicity's sake we will refer to the program as "TaRL."

⁶⁰ "Financials," Pratham USA, 2018. [Access here.](#)

⁶¹ "Education for All Movement," UNESCO, 2017. [Access here.](#)

⁶² Stakeholder Interviews conducted between December 2016 and March 2018

⁶³ "Pre-kindergarten," Center for Public Education. [Access here.](#)

⁶⁴ R. Banerji and M. Chavan, "Improving literacy and math instruction at scale in India's primary schools: The case of Pratham's Read India program," *Journal of Educational Change* 17, no. 4: 453-475. [Access here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

outside of regular class time. Chavan thought that this approach might work in India and, in 1998, Pratham began implementing a remedial learning program by pulling primary school students who were lagging behind out of class to work with a volunteer tutor (“balsakhi”) on basic numeracy and literacy skills. This was the first Pratham program to undergo rigorous evaluation and became the seed of the TaRL approach.

About TaRL

TaRL is designed to improve basic literacy and math skills for students in grades 3-5. The approach features: (i) grouping children by learning level rather than grade level (age), and (ii) teaching children at each learning level through engaging activities and tailored materials. TaRL evolved through an iterative design process from Pratham’s early remedial education work in the late 1990s. Researchers from MIT’s Abdul Latif Jameel Poverty Action Lab (J-PAL) worked with Pratham to test the approach in six randomized controlled trials (RCTs) conducted between 2001 and 2014 in six states across India. The results showed that TaRL led to learning gains when delivered directly by Pratham staff or volunteers, as well as indirectly through partnerships with government implementers. The methodology, or variations on it, has been implemented in 20 states and three Union Territories in India and in countries in Africa and South America.



As this case describes, three key factors have determined the evolution and scaling up of TaRL:

1. Pratham’s culture, which is driven by a commitment to learning and impact at a national level through deep, ongoing engagement with communities and with policymakers.
2. Long term partnerships with evidence producers (primarily J-PAL), to test program iterations, crystallize the core principles of TaRL, and identify effective

approaches to scaling, as well as with policymakers and funders.

3. The Annual Status of Education Report (ASER), a citizen-led survey that builds national awareness of low learning levels and drives demand from government for scalable solutions.

There are some critiques of the TaRL model: that it raises children to a common baseline rather than the grade-appropriate curriculum; and that remedial education addresses



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“I think we’ve been very lucky because nobody forced any of these [evaluations] on us, so it was a [voluntary] coming together of two sides...”

a symptom instead of the larger problem. However, TaRL is designed to address the critical challenge of low foundational literacy and numeracy skills in India, and evidence shows that it has successfully improved learning outcomes for millions of children.

Commitment to Ongoing Learning

TaRL is a pedagogical approach that is rooted in rigorous evaluation and has been implemented at a large scale. Pratham is unique among education NGOs in India for its investment in building organizational capacity for monitoring and evaluation, its willingness to expose itself to evaluation by external researchers, and its commitment to sharing evaluation results publicly. By contrast, one former policymaker describes the context in

India’s education sector:

“The tradition of evidence-based decision-making is very, very weak in our country. There is no premium or incentive on it. There is no centralized mechanism of validating that this program seems to have done well [...] compared with this one.”

Since 2001, Pratham has built a strong partnership with J-PAL, whose staff describe Pratham as “very, very results-oriented.” Pratham staff bring professional expertise in collecting and analyzing data and routinely attend trainings and workshops delivered by J-PAL and other evidence producers. One Pratham staff member describes the unique partnership:

“I think we’ve been very lucky because nobody forced any of these [evaluations] on us, so it was a [voluntary] coming together of two sides ... We were very lucky to have them as partners because we have done some work with other people as well, but we see that they [J-PAL] really treat you as a partner and as an equal.”

Institutionalizing Measurement, Monitoring + Evaluation

By 2013, Pratham was implementing diverse programs across India and collecting data to inform its work and report to funders. However, data collection and analysis was decentralized and varied by program and state. To improve efficiency, Pratham CEO Rukmini Banerji created a Measurement, Monitoring, and Evaluation (MME) unit to standardize and streamline data collection and analysis for all Pratham programs. The unit’s team, which by 2016 had grown to a staff of 80, aims to digitize data collection and present data in visually useful forms for various audiences, including field staff, volunteers, teachers, funders, and policymakers. It attracts top talent, despite a lower pay scale, with a focus on generalists who are passionate about using data for the public good.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The MME team believes that measurement is necessary for accountability. For example, to spur competition between two districts, the MME team may present policymakers with a rigorous report that compares learning levels in those Districts. As people become exposed to data, they often learn to value its usefulness. For example, Pratham experienced resistance when it created a tool to help teachers track individual children. After using the tool, teachers came to demand it because it helped them do their jobs better.

The MME team designs measurement tools in partnership with other Pratham program teams with complementary expertise, but it is designed to be independent. It shares an office and works closely with the ASER Centre, discussed below, which conducts national assessments of children's learning levels. As a result, and by design, information is gathered and shared collectively within the Pratham network:

“While it’s an internal unit it almost functions as an external unit because it’s not part of the program structure within Pratham. That has really helped as well. One of the main intentions was that we wanted the MME unit to be very evaluation, learning, research focused so the conscious decision was made to base it out of ASER Centre.”

“We wanted the MME unit to be very evaluation, learning, and research focused so the conscious decision was made to base it out of ASER Centre.”

In addition, Pratham stresses the importance of understanding the context of a community before conducting an assessment. Members of the teams involved in assessment are expected to travel to the field frequently, and any tools developed for a community are reviewed by local teams to ensure they are contextually appropriate. This learning collaboration between internal teams, and with communities, is illustrative of Pratham's culture and has played a significant role in TaRL's success.

Evaluating Pratham and Establishing a Proof of Concept

At critical junctures, J-PAL has evaluated Pratham's newest program adaptations to identify significant determinants of impact across time and context [*see Appendix 1 for a summary of the RCTs*]. The first RCTs were conducted between 2001 and 2004 by a team from the Massachusetts Institute of Technology, including Abhijit Banerjee and Esther Duflo, professors of Economics who co-founded J-PAL in 2003. The researchers approached Pratham because it was considered unusually data-driven and able to scale quickly and inexpensively.

These early RCTs determined that Pratham's remedial tutoring—the balsakhi—program was cost-effective and led to significant learning gains. They also demonstrated Pratham's ability to administer programs at scale. Thus, when Pratham wanted to adjust the model to implement the program through volunteers rather than paid tutors, J-PAL supported this experimentation.

Over the ensuing years, the partners systematically tested adjustments to improve impact and scalability, such as a 2005 evaluation for the World Bank, which showed that Pratham's volunteer-implemented *Learning to Read* model resulted in a 20% increase in participants' reading ability. This



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

meant “that the pedagogical idea behind the balsakhi program could survive the change in context and program design.”⁶⁵ However, the process evaluation also revealed flaws, including high volunteer attrition and low up-take by children who needed the program most.⁶⁶ Once again, Pratham redesigned its approach to address these challenges.

Experimenting with Pathways to Scale

In 2006, the Bill & Melinda Gates Foundation and the William & Flora Hewlett Foundation announced the Quality Learning Outcomes Initiative, a unique partnership established to improve learning outcomes, particularly in Sub-Saharan Africa and South Asia, where learning levels were lowest worldwide.⁶⁷ In 2007, after a due diligence process and field visits, Hewlett awarded a grant for the first iteration of Pratham’s new Read India program.

From 2007 to 2010, Read India was characterized by a focus on scale and experimentation, and was implemented directly by Pratham staff and volunteers, as well as with government partners. In the first two years of the program, Pratham scaled quickly, reaching 33 million children by 2008.

Between 2008 and 2010, Pratham and J-PAL evaluated two government partnerships in the Indian states of Bihar and Uttarakhand.⁶⁸ For



the first time, these RCTs tested the impact of Pratham’s approach when implemented by government teachers. In Bihar, the program evaluation also compared the results of teacher implementation to those of a volunteer-implemented “camp” outside of school hours. Surprisingly, the “introduction of Pratham’s methodology in schools during the school year, failed in both states.”⁶⁹ Qualitative data, including process monitoring and stakeholder interviews, identified that teachers’ low compliance with the model was responsible for the disappointing results. The results also indicated that teachers could successfully implement the program in the right context: when conducting summer camps outside of the school year, those same

teachers in Bihar did group children by level and students achieved learning gains on par with previous studies.

Based on interviews with teachers, Pratham and J-PAL determined that “teachers in both states seemed to believe the methods proposed by Pratham were effective and materials were interesting, but they did not think that adopting them was a part of their core responsibility.”⁷⁰ The partners concluded that there were two alternative paths to overcome low compliance during the school year: (1) work with school administrators and senior policymakers to ensure that teachers viewed TaRL as a core responsibility, or (2)

⁶⁵ A. Banerjee et al., “From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application,” *Journal of Economic Perspectives* 31 no. 4: 73-102. [Access here.](#)

⁶⁶ *Ibid.*

⁶⁷ “Gates and Hewlett Foundations Join to Improve the Quality of Education in Developing Nations,” Hewlett Foundation, December 2006. [Access here.](#)

⁶⁸ A. Banerjee et al., “From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application,” *Journal of Economic Perspectives* 31 no. 4: 73-102. [Access here.](#)

⁶⁹ *Ibid.*

⁷⁰ *Ibid.*



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

take teachers out of the equation and work through Pratham staff and volunteers, who consistently achieved positive results.

Although the RCTs on government partnerships in Bihar and Uttarakhand were disappointing, Pratham believed the most reliable path to scale would be through the government school system. Positive experiences working with a different state government, along with RCT data indicating that grouping by level led to learning gains, persuaded Pratham that government teachers could implement TaRL at scale under the right conditions.

Optimizing Pathways to Scale

In 2009 the Government of India passed the “Right to Education Act” (RTE), which replaced year-end examinations with “Continuous and Comprehensive Evaluations” and required that children move to the next grade level through grade nine, regardless of learning level. The new policy posed a challenge to Pratham’s learning level-based philosophy and generated significant debate. In this context, Pratham decided to scale down its efforts and reevaluate its strategy.⁷¹

In support of Pratham’s decision to take a step back, the Hewlett Foundation awarded Pratham a second grant from 2010 to 2013 for phase two of Read India. During this



phase, Pratham implemented its model in strategically located blocks throughout India “to illustrate to key players in education that rapid and cost-effective improvement in learning are possible using Pratham methodology and materials.”⁷² Pratham went from working in 300,000 villages in phase one to only 25,000 villages in phase two on the conviction that deeper relationships with local decision-makers and communities would more effectively encourage them to adopt Pratham’s approach. As a result, the underlying work would continue, irrespective of political cycles.

During this period, the state government of Haryana asked J-PAL to help build its internal capacity to conduct an RCT. J-PAL recruited Pratham as program partner. J-PAL was excited to evaluate TaRL under the new RTE regime; Haryana was eager to learn

⁷¹ R. Banerji and M. Chavan, “Improving literacy and math instruction at scale in India’s primary schools: The case of Pratham’s Read India program,” *Journal of Educational Change* 17, no. 4: 453-475. [Access here.](#)

⁷² “Raising the Bar on Hope: Pratham 2010 Annual Report,” Pratham, [Access here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

from a reputable research institution; and Pratham hoped to prove that TaRL could scale effectively in partnership with governments. The model in Haryana addressed the challenge of teachers' non-compliance observed in Bihar and Uttarakhand by ensuring that teachers viewed TaRL as a core responsibility. This was achieved by

- emphasizing government ownership of the program
- training cluster resource coordinators who monitored and mentored teachers
- adding a government-mandated hour of TaRL instruction to the school day
- requiring teachers to physically re-group children by learning level.

The 2012-2013 RCT revealed learning gains that exceeded expectations. Process monitoring data demonstrated much-improved compliance, with 90% of schools being grouped by learning level during the TaRL hour. Interviews suggested that cluster resource coordinators played a critical role in these outcomes.

In 2012, Pratham received a three-year grant from USAID to conduct an evaluation of an alternative TaRL implementation model that did not rely on government implementation. Instead, TaRL was delivered during a one-hour



session during school hours by Pratham staff and volunteers in short bursts of 10 to 20 days for a total of 40 days a year. From 2013 to 2014, J-PAL conducted a RCT of what became known as the “learning camp” model, which demonstrated “the relative ease with which apparently daunting learning gaps can be closed.”⁷³

Based on the successful results shown by these RCTs, Pratham was confident about the essential components of TaRL and the two different approaches to scale (directly through Pratham staff and volunteers and indirectly through government partnerships).

It launched the third phase of Read India (2013-2016) with support from Hewlett focusing on scaling TaRL both directly and through government implementers. As of 2017, the teacher-led model has been implemented in over 100,000 schools across 14 states, reaching almost 5 million children, while the in-school, volunteer-led model has been implemented in over 4,000 schools across India, reaching over 200,000 children.⁷⁴ Pratham has also leveraged the results of these RCTs to secure hundreds of district-level government partnerships.⁷⁵

The Annual Status of Education Report

Alongside its efforts to test and identify essential components of TaRL, Pratham worked to shape the public debate around education in India. The Annual Status of Education Report (ASER) continues to be one of the organization's most significant contributions in this respect. ASER is now a highly anticipated annual report on

⁷³ A. Banerjee et al., “From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application,” *Journal of Economic Perspectives* 31 no. 4: 73-102. [Access here.](#)

⁷⁴ *Ibid.*

⁷⁵ S. Chatterji Dutt, C. Kwauk, J. Perlman Robinson, “Pratham's Read India Program: Taking Small Steps Toward Learning at Scale,” *Brookings*, 2016. [Access here.](#)



learning levels across India and has spurred policymakers to adopt programs that will improve their state- or district-level results.

In 2004, the newly elected Government of India formed the United Progressive Alliance (UPA). The UPA government was committed to education reform, launching a 2% tax that would fund initiatives to improve access to and quality of education. Senior policy advisors familiar with Pratham's work invited Madhav Chavan to join the National Advisory Council advising the federal government on various education issues, making one of the co-founders of Pratham a senior policy advisor in the new government.

Within the government, there was debate about the best approach to improve education outcomes. The Ministry of Education had always collected an enormous amount of data, but it focused only on aggregate outcomes instead of individual learning levels. Given new policies aimed at improving learning outcomes, Pratham gave a presentation recommending a new assessment that could be quickly implemented and would measure learning levels of individual students across India. The Planning Commission encouraged Pratham to run a pilot and report back.

Pratham's standardized assessment tool for literacy and numeracy consisted of a single

“The reason we’re focusing on learning in India is because ASER made the problem in learning visible. Everyone knew people weren’t learning, but they put a vocabulary around it. A vocabulary everyone understood.”

sheet of paper with four levels of text—letters, common words, sentences in a paragraph, and stories. A similar tool was created for numeracy. The content was adapted to the local language and state-specific literacy and numeracy expectations for grade two.

Through partner organizations in rural districts across India, Pratham conducted the first household-level Annual Status of Education Report (ASER) survey in late 2005.⁷⁶ The results suggested that over 90% of children ages 5 to 16 were enrolled in school, but 50% could not read at a grade two level. Meanwhile, the government commissioned a third-party assessment of total enrollment by a private company, which reported the

same enrollment figure. Pratham claimed that if the government could accept the accuracy of its impressive enrollment figures, it should also accept data on the dire state of learning outcomes for children enrolled in school. Given the accuracy and timeliness of the ASER survey data, the government decided to use the report to inform its Annual Economic Survey. From 2005 to 2012 one page of the Finance Ministry's annual report featured ASER data. In 2012, the Ministry of Education began emphasizing learning outcomes and urged state governments to begin their own

⁷⁶ For further information on the ASER Center, please see its [web site here](#).



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

measurement processes, many of which were modeled on the ASER survey.

The ASER Centre, established as an independent unit, is “the largest household survey of children conducted in India by citizens’ groups, carried out by more than 25,000 volunteers and covering over 700,000 children in 15,000 villages each year.”⁷⁷ As of 2017, the ASER survey is conducted in 595 of 613 rural districts in 29 states and four union territories and is constantly refined through rigorous field tests.

ASER has played an important role in scaling Pratham’s work, including TaRL, for three key reasons. First, ASER (with significant influence from Pratham) placed learning outcomes at the center of the national debate on education. As one expert noted:

“The reason we’re focusing on learning in India is because ASER made the problem in learning visible. Everyone knew people weren’t learning, but they put a vocabulary around it. A vocabulary everyone understood.”

Second, the ASER survey and the organization built around it highlight and reinforce Pratham’s organizational commitment and capacity to produce and use evidence.



Former ASER Director Rukmini Banerji, now CEO of Pratham, has a PhD in Education and has advised the government of India and others on using data to inform education policy. She has been a co-chair for the Global Learning Metrics Task Force, convened by the Brookings Institution and UNESCO.⁷⁸

The ASER Centre reinforces Pratham’s ability to collect and use data because the two organizations are intertwined and interdependent. Moreover, ASER data on

⁷⁷ *Ibid.*

⁷⁸ “Rukmini Banerji,” International Growth Center, 2018. [Access here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

children's learning levels are one of the most important forms of evidence Pratham uses to engage government or community partners and determine appropriate interventions.

Finally, ASER has generated demand from policymakers for interventions to address low learning levels in their states. ASER illuminated the scale of the problem and built demand for solutions like Pratham's. Many states have also implemented their own solutions thanks to ASER's reporting:

"There was a lot of action and programs, which states take to improve their metric of ASER. They [the ASER Centre] are in a very powerful position. So how they frame the problem will in fact, whether or not policymakers acknowledge it, guide the actions a lot of policymakers take. I know a lot of states, they would say they were sending out packages to different districts, saying please get these children to be able to recognize letters, words, because they are saying, look we need to get these forms up on ASER."

Pratham's leaders promoted a culture of learning, which they further institutionalized through the establishment of the Measurement, Monitoring, and Evaluation team.

TaRL evolved and achieved scale over the course of over two decades. Pratham's leaders promoted a culture of learning, which they further institutionalized through the establishment of the Measurement, Monitoring, and Evaluation team. This internal commitment to testing and improving Pratham's programs, along with partnerships with external evaluation organizations like J-PAL, resulted in a methodical, iterative approach to program design with demonstrated outcomes. At the same time, increased public awareness about learning levels, thanks in large part to ASER, created new opportunities to scale TaRL directly and in partnership with governments.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part II: **Key Themes + Insights**

This section discusses the Evidence in Practice themes as they pertain to Teaching at the Right Level 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.

Evidence producers, policymakers, and implementers agreed that, while RCTs have played an important role in refining and legitimizing TaRL, they are a “slice” among many types of evidence required for program design and implementation. Other types of evidence have proven equally useful depending on the purpose and stage of the

policy-making process. Process monitoring and qualitative data, as well as RCTs, helped identify critical program components. The results were often used to get a foot in the door with funders and policymakers, though the extent of evidence needed to convince policymakers varied. In some cases, policymakers with academic backgrounds reviewed the RCTs. In other cases, J-PAL explained the relative rigor and objectivity of different kinds of evidence in order to persuade policymakers to try TaRL. In some cases, policymakers felt that a stamp of approval from a well-respected, elite institution like J-PAL indicated that there was sufficient evidence behind the model for a pilot to be politically acceptable.

In every case, evidence of Pratham’s ability to operate at scale was at least as important as the academic evidence. Field trips to observe children learning through the TaRL approach

were often most effective to secure buy-in from policymakers, funders, and government implementers. In addition, policymakers commissioned third-party assessments of pilot programs to determine whether and how to scale.

In part, the open, explicit, and shared acknowledgement of the acceptability of diverse types of evidence enabled partner organizations, such as Pratham, J-PAL, funders, and government actors, to be flexible and develop trust.

Ultimately, there are times when evidence may be compelling but at odds with political promises or priorities. Thus, different forms of evidence have led to an increasingly nuanced, though still unresolved, debates on the best approaches to improving educational outcomes. The story of Pratham illustrates that there is always a tension in the policy process: between what is most effective and what is most efficient, between what is better in the long term and what is politically feasible in the short term. The quality of evidence that Pratham and its partners have created has allowed for a much deeper debate on what the next iteration of that policy development process looks like, with each successive cycle of debate based on increasingly rigorous and complete data.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Incentive 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.

Pratham's commitment to helping all children achieve a foundation in literacy and numeracy has translated into a focus on organizational learning to improve outcomes. Pratham's incentives have thus aligned well with evidence producers and funders with similar goals and outlooks. For example, J-PAL and Pratham have enjoyed aligned incentives throughout their partnership. J-PAL could test large-scale pedagogical interventions with an effective implementer, while Pratham could build its evidence base and credibility

“... we're in favor of Chinese restaurants. We have a menu and you can take that menu, you can change it a bit here and there, but when you set up your own restaurant people will say, 'This is a Chinese restaurant.'”

collaborating with a highly respected evidence producer. However, Pratham recognized that buy-in from everyone, from communities to senior policymakers, who have vastly different incentives, was necessary for the systemic change it sought. When possible, Pratham created shared incentives by co-designing interventions with other stakeholders, investing time to build relationships, and holding stakeholders accountable through monitoring and measurement.

As noted above, policymakers at times have incentives not to use specific evidence. Political cycles and short tenures in administrative positions drive investment in programs with short-term results. Evidence supporting a policy intervention may also reduce policymakers' discretion to adopt a more politically or personally expedient policy. Adopting a new program, even if backed by evidence, can be politically risky if it fails due to poor design or implementation. To account for and counteract such potentially misaligned incentives, both Pratham and J-PAL have found it useful to present policymakers with a menu of options as opposed to a prescribed program. This is possible with TaRL because the essential program components have been identified and can generally be financed by reallocating existing education budget line items. Policymakers from multiple states noted that Pratham's willingness to adapt the program based on their input was a key reason they decided to partner.

By co-designing the approach with policymakers, Pratham aims to increase government “ownership” so that the TaRL approach becomes aligned with government incentives, as well as engrained and resilient to political cycles:

“In fact, our theory is that we are not about to set up a McDonald's chain, so you insist that your meat patty and your



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love



saucers should be, your exact prescription of Golden Fries should be available everywhere. No, we're in favor of Chinese restaurants. We have a menu and you can take that menu, you can change it a bit here and there, but when you set up your own restaurant, people will say, 'This is a Chinese restaurant,' because it looks like that. ... Every government is opening a Chinese restaurant, which looks a little bit like what we have, but its flavor is different. [...] That is how actually ideas start spreading."

Thus, Pratham understands that providing a menu of options and allowing policymakers

to choose is more effective than pushing a prescribed initiative. Pratham has determined which program components are essential vs. which can be adapted based on context. This is possible because the essential components of TaRL are simple and well-defined: group children by level, teach them at that level, and ensure teachers are effectively motivated.

Pratham has also used monitoring and measurement to align stakeholder incentives. For example, as noted above, the MME team would present policymakers from two Districts with comparison reports on common indicators to spur competition between them.

Timing Misalignments

The different and often discordant timeframes within which researchers, policymakers, and implementers operate often hobble efforts to coordinate, let alone collaborate, on evidence-informed approaches. Electoral cycles and political windows differ from NGO funding cycles and from academic publishing rhythms. Yet each actor is bound by the timeframes of her formal stakeholder group.

Pratham has developed patience and the agility to take advantage of and create shifts in the education discourse. However, timing misalignment has posed a significant challenge to scaling TaRL through government. For example, turnover within government can make long term partnership difficult or impossible. In several cases, key champions in a state left their positions after the government had signed an MOU with Pratham. Afterwards, there was insufficient political will to carry out the program as planned. Pratham and J-PAL have learned to identify and cultivate champions across administrative levels for redundancy and to endure political transitions.

However, structural timing misalignments remain. Budget cycles are rarely aligned with program implementation timelines: while a state-wide scaling up would take several



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

years, budgets are generally approved on an annual basis. Delays or shortfalls in approved budgets have at times damaged the implementation of TaRL. It is challenging to find a solution to this timing misalignment because policymakers' due diligence processes take time and their priorities change in ways that are difficult to anticipate.

Despite these challenges, many governments are open, and even enthusiastic, about using evidence to inform education policy and programs. Pratham and J-PAL have learned to avoid states where there is significant turnover, lack of trust of NGOs, a history of covering up poor program results, or entrenched programs that would be difficult to replace.

Trust, Respect, and Buy-in Among Stakeholders

The cross-stakeholder collaborations required for evidence-informed policies and practices are often difficult to initiate, develop, and sustain. Particularly when institutional incentives are lacking, personal trust, respect, and buy-in between individuals across stakeholder groups become critical to fostering the effective flow of evidence into practice.

Trust and respect developed between stakeholders in large part due to relationships



Pratham formed with researchers, policymakers, and funders. J-PAL has served as a trusted partner through its role as an evidence producer that published and presented Pratham's RCT results and, more recently, through its policy team that was established in 2010 to help scale rigorously tested interventions. TaRL is one of a handful of programs that the policy team promotes and is one of the few that has been tested in multiple contexts. J-PAL's policy team approaches policymakers with whom it has stronger relationships than Pratham and presents the evidence behind TaRL in various forums.

Pratham's commitment to scale required a model that was cost-effective, and simple enough to implement in extremely diverse contexts. As a result, even when an RCT showed significant gains in learning outcomes, Pratham insisted on adjusting and testing the model until the most impactful, cost-effective components were identified. Understanding these essential features enabled Pratham to be flexible when adapting TaRL to a new context in partnership with other stakeholders. Pratham, J-PAL, funders, and government staff at different administrative levels were all involved in this process of translating evidence into practice. Pratham launched each new adaptation or expansion of TaRL by convening the relevant stakeholders to jointly explore how best to improve learning outcomes for students. The partners overcame common design and implementation challenges due to an integrated, collaborative approach to translation that leveraged stakeholders' complementary resources, networks, and insights to promote and operationalize TaRL.

Pratham invested in long term relationships to promote government adoption of its evidence-informed programs. Pratham simultaneously built relationships with communities and local policymakers through direct programs, and with decision-makers at the highest levels of government through advocacy and



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

advisory efforts. Pratham's leaders also served on government committees and cultivated relationships with policymakers who went on to hold senior positions in central and state governments.

Through its experience working with government at all levels, Pratham has learned that partnerships will not move forward unless champions with convening power have been cultivated. Champions are critical at all phases of a partnership and at all administrative levels. One policymaker explained that with the Education Secretary and Minister as champions of a partnership with Pratham, it was easy to make budget allocation decisions and to move from pilot to scale.

Although Pratham is one of India's largest NGOs and works across the country, its staff recognizes that "government is the only one that can change things. Pratham can't change things alone." The organization is eager to build the capacity of government partners so that more children can benefit from its evidence-informed methodology. As Pratham scales this effort, it is also learning strategies to overcome common challenges of working through government.

Policymakers reported that, given the complexity and inescapable trade-offs of their decisions, Pratham's flexibility was especially

“It was much more in a ‘we have a shared goal and we’re all thinking about the best way to reach it’ sort of way than in a directive ‘we’re the funders, it’s our money, and we think this is what you should do, so do it.’”

FUNDER

important. Pratham leaders share the belief that being flexible about program design and implementation based on the unique needs of a given context is not only necessary, but it also benefits Pratham's work.

Long term relationships also helped Pratham align incentives with funders. Over time, funders developed confidence in Pratham and became comfortable supporting Pratham's experimentation to improve its already effective model. However, even

with time there were funders who preferred scaling “proven” models over experimentation. Pratham recognized this dichotomy among funders and cultivated a “healthy mix” of donors (with different underlying incentives) to fulfill these two goals. Several individuals at funding institutions played a significant role as champions in their willingness to fund experimentation. Several such individuals moved from one foundation to another and continued their relationships with, and support for, Pratham.



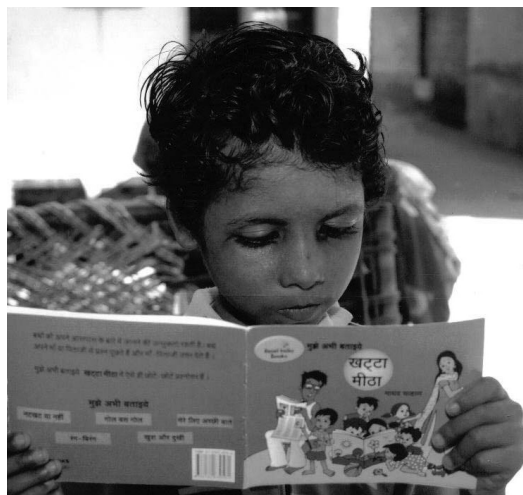
Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Devoting Time and Resources to Integrating Evidence

Few organizations provide incentives or carve out explicit time for managers to explore emerging evidence in their field. Even fewer assign staff to find relevant evidence and translate it into accessible formats for the organization. As a result, the role of preparing and sharing evidence that is timely, useful, and relevant for practitioners is sometimes explicitly played by formal intermediaries (e.g., certain think-tanks). More frequently, an actor who holds a formal role within another stakeholder group spontaneously takes on the (additional) responsibility for trying to integrate evidence, with no actor formally responsible for the process. Discovering and integrating evidence requires time, energy and funding.

Externally generated evidence was used by all stakeholders involved in this case. Pratham and the ASER Centre used academic literature and research conducted by think tanks to inform assessment tools. Policymakers used independent contractors to conduct third-party assessments of pilot programs to determine whether to scale. Funders used RCT results to make decisions about grant-making strategy to improve learning outcomes. J-PAL used government data to adapt TaRL for a new context.

However, several interviewees pointed out that many policymakers lack the capacity to use data, even when their capacity to collect data has strengthened over time. In this case study, externally-generated data was presented and explained to policymakers by both Pratham and J-PAL. Pratham's MME unit designed reports for policymakers based on their perceived level of comfort interpreting data, their decision-making authority, and their willingness to change. J-PAL's policy team met with senior policymakers to explain different types of evidence and their relative usefulness for policymaking. In some cases, this advocacy resulted in the government adopting a TaRL program.



Operationalizing Evidence Captured Internally

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.

Pratham launched its MME unit in 2013 to improve coordination in measurement, monitoring, and evaluation processes, and to be able to more easily share and interpret data across programs and sites.

The MME unit is equipped to create useful tools for monitoring and translating operational data because of its collaboration with the ASER Centre and other Pratham teams. Pratham now has strong internal capacity and resources to collect and analyze program-generated evidence. When Pratham has lacked internal capacity to collect or interpret data, it has been



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

remarkably good at building relationships with organizations with complementary capacity and resources.

Learning from Failure

Potential consequences for risk-taking and experimentation with innovative approaches are generally seen as negative and dissuade the exploration of novel, evidence-informed interventions. Fear of failure can further hinder the incorporation of novel evidence into practice, even when stakeholders recognize the value and applicability of the evidence.

Observers and staff emphasized Pratham's culture of commitment to experimentation and an unusual openness to being evaluated. Pratham tends to hire people who are not tied to a specific education ideology and who have a strong interest in data and impact. There is a desire among staff across all levels of the organization to constantly improve and expand Pratham's impact. In this environment, experimentation is encouraged and failure is accepted as an opportunity to learn.

Pratham has required funding not directly tied to program implementation to nurture its data-driven culture and to test and refine its approach. Two elements seem to enable a successful, long term relationship between

Pratham and a flexible funder. One is a shared mission and commitment to learning. As one funder explains:

"It was much more in a 'we have a shared goal and we're all thinking about the best way to reach it' sort of way than in a directive 'we're the funders, it's our money, and we think this is what you should do, so do it.'"

The second ingredient is open communication, in which both parties learn from one another. Two-way communication about successes and failures invites potential solutions. As one senior staff member at Pratham puts it:

"We set very high standards for ourselves and we are not scared to fail. That's also a cultural thing in Pratham. If you do fail, it's okay, as long as you acknowledge it and you want to make a change, it's fine. Being open with ... our partner funders on those aspects has also been very beneficial for us because we've been able to learn from them."

"Internally as an organization and externally as a major player on the Indian scene, Pratham learned that it is important to be flexible and nimble, to seize and to create opportunities, and to continue to push the learning agenda on every available forum."



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Conclusion

Pratham, through TaRL and other programs, has had a significant impact on the education sector in India. Through its unrelenting commitment to the generation and use of evidence, Pratham has systematically influenced the public debate on remedial education, first insisting that grade level is not synonymous with learning level, then pushing for measurement of learning levels, and then partnering to implement solutions to improve learning levels. Pratham's approach to evidence-informed policy and practice is rooted in the long term relationships it has cultivated with policymakers, researchers, funders, and communities.

As a simple and cost-effective program, TaRL is well-suited to scale through multiple channels. Pratham's relationships, along with its organizational commitment to learn over time and achieve better outcomes, have enabled Pratham and its partners to improve learning outcomes for millions of children in India. Moreover, many governments and other organizations, in multiple settings and contexts around the world, have taken up TaRL or programs modeled after TaRL,⁷⁹ achieving similarly positive results.

Pratham's commitment to learning is summarized well by Rukmini Banerji and Madhav Chavan:⁸⁰

“The major lesson for Pratham was also that the journey of transformation of communities and of systems is a long one, with continuous learnings at every step. Internally as an organization and externally as a major player on the Indian scene, Pratham learned that it is important to be flexible and nimble, to seize and to create opportunities, and to continue to push the learning agenda on every available forum. At the same time, it is also essential to play the game simultaneously on both fields—the micro dynamics at the ground level and the macro dynamics of national and state priorities, policies and plans. [...] Finally, simple measurement is used to track progress of individuals and groups towards the goals and to guide corrections throughout the process.”

⁷⁹ See the case study in this series on the *Teacher Community Assistance Initiative in Ghana*.

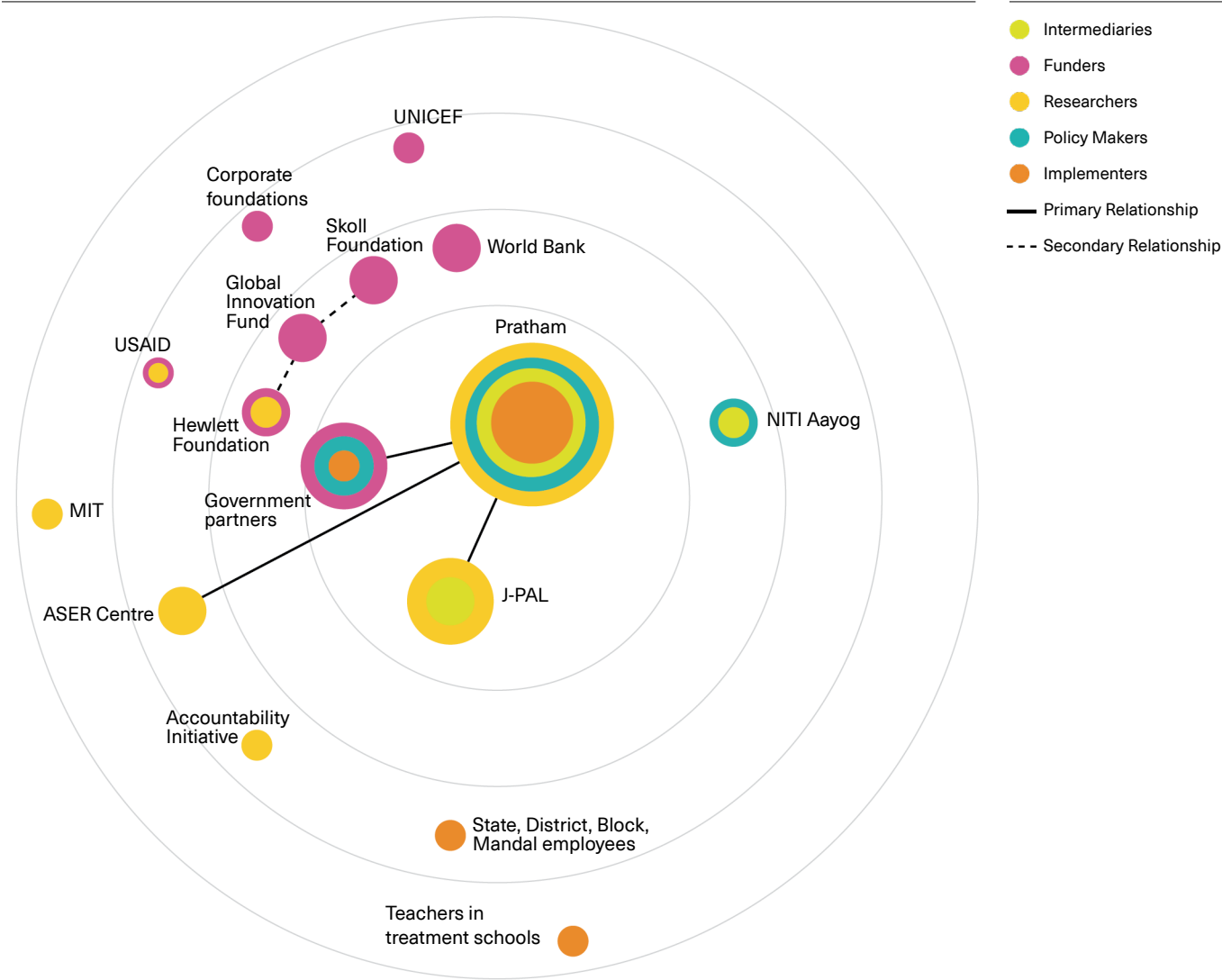
⁸⁰ R. Banerji and M. Chavan, “Improving literacy and math instruction at scale in India's primary schools: The case of Pratham's Read India program,” *Journal of Educational Change* 17, no. 4: 453-475. [Access here](#).



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Stakeholder Map

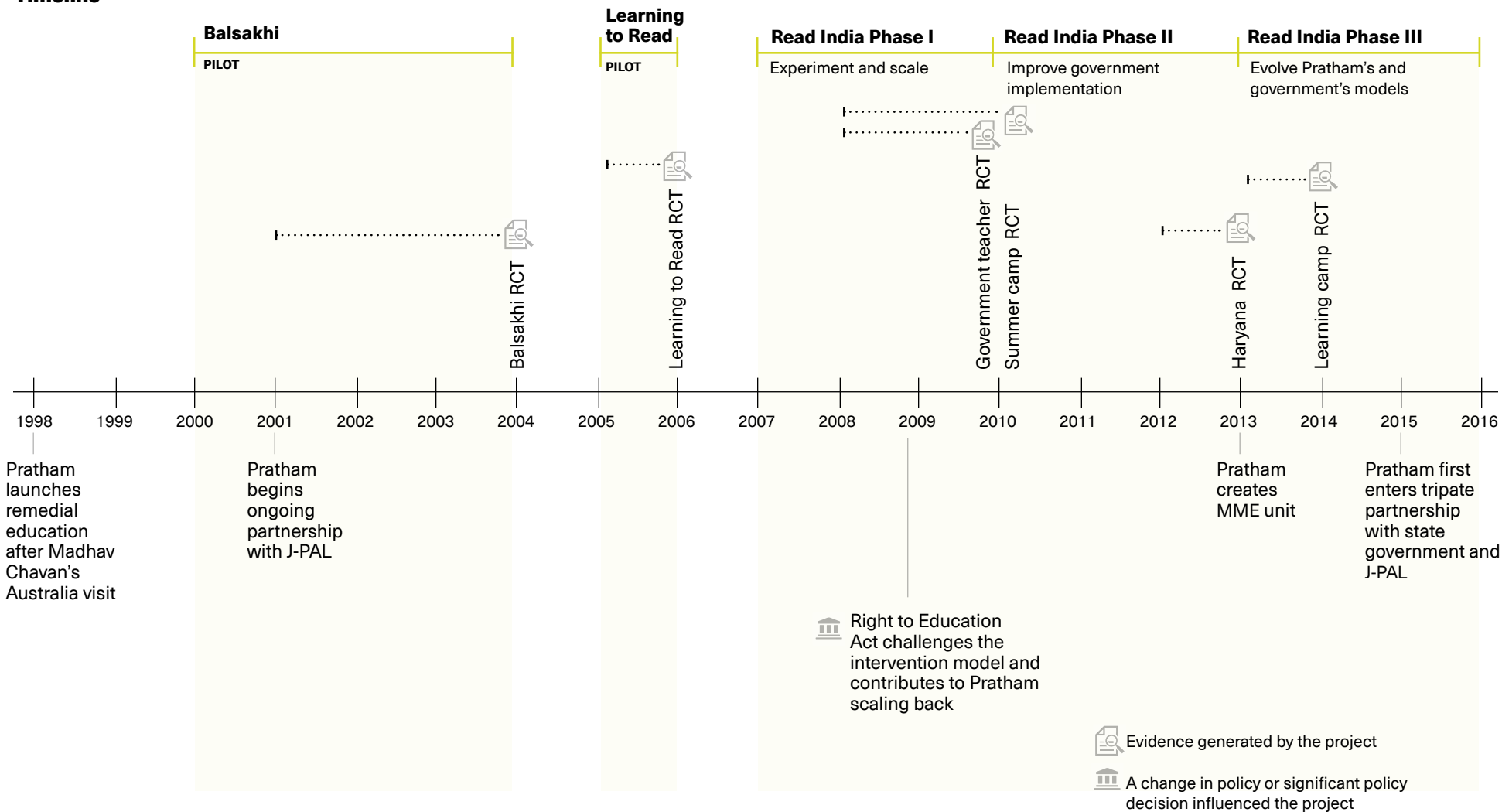
This stakeholder map is a visual representation of the major stakeholders involved with this project. The importance of each of the actors is defined by their relative size, and their proximity to the center of the project. Their role is defined by the color; multiple colors indicate multiple roles. 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.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Timeline

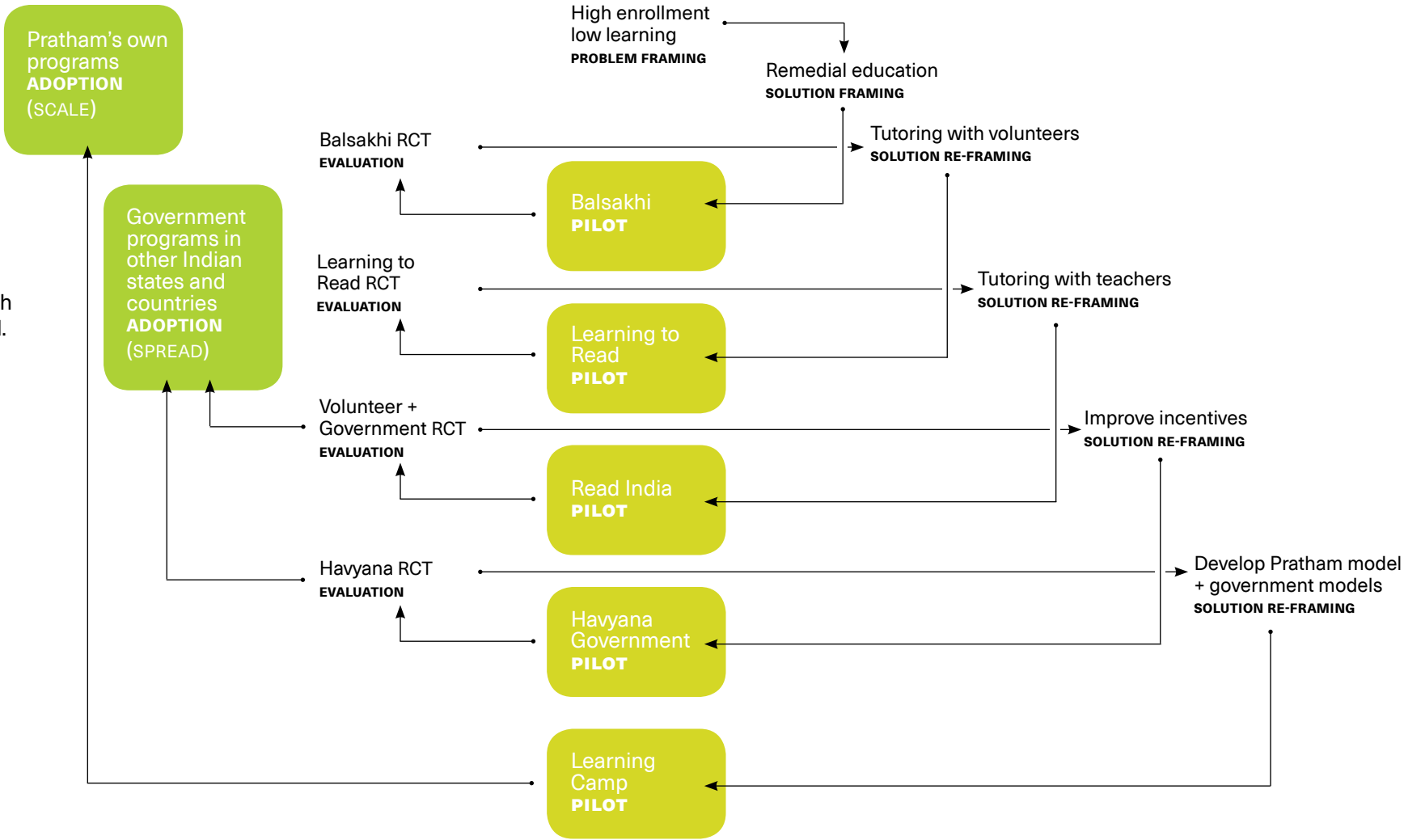




Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Process Diagram

An iterative approach to creating and incorporating evidence allowed Teaching at the Right Level approaches to both **scale** and **spread**.





Appendix I: Summary of RCTs

Between 2001 and 2014, Pratham and J-PAL conducted the following randomized controlled trials (RCTs) to test and refine the components of TaRL.

Year	State	Description of Evaluation*	Outcome	Resulting Question
2001-2004	Gujarat, Mahara-shtra	Compared balsakhi program, in which paid local tutor taught Pratham's remedial curriculum to students in grades 2-4, to Pratham's Computer Assisted Learning (CAL) program	Positive learning gains in both programs, although balsakhi program was more cost-effective than CAL	Could similar learning gains be achieved with volunteer tutors who could scale the program inexpensively?
2005-2006	Uttar Pradesh	Compared Pratham's "Learning to Read" model, in which community-based volunteers conducted learning camps for children ages 7-14 outside of school hours, with providing information to communities about their right to education	Significant positive learning gains when "Learning to Read" model was implemented with fidelity to the model; information alone had no impact	Could the model be implemented by government teachers to overcome challenges of high volunteer attrition and low uptake by eligible students?
2008-2010	Bihar	Tested three models as part of Pratham's Read India program: (1) month-long summer camp implemented by government teachers i for grades 3-5; (2) dedicated time during the school day for the Read India curriculum taught by government teachers for grades 1-5; and (3) volunteer implemented learning camps for grades 1-5 [as in Uttar Pradesh]	Disappointing results when implemented by government teachers during the school year due to low teacher compliance; significant learning gains from summer camp implemented by government teachers and learning camp implemented by Pratham volunteers	See below.
2008-2010	Uttara-khand	Tested two models as part of Pratham's Read India program: (1) dedicated time during the school day for the Read India curriculum taught by government teachers for grades 1-5; and (2) volunteer support for teachers implementing the curriculum during school hours for grades 1-5	Disappointing results when implemented by government teachers during the school year due to low compliance when teachers felt the program competed with other priorities and was not part of their job description	Based on results from Bihar and Uttarakhand, could government teachers be better incentivized to implement the model with fidelity?
2012-2013	Haryana	Evaluated government implementation of TaRL during a dedicated hour in school for students grouped by level in grades 3-5; cluster resource coordinators trained, monitored, and mentored teachers	Cluster resource coordinators played an essential role in generating strong compliance among teachers that resulted in significant learning gains for students	Illustrates model for scaling through government (indirectly)
2013-2014	Uttar Pradesh	Evaluated implementation of TaRL learning camps by Pratham staff during one hour of the school day in short bursts of 10-20 days, four times per year for grades 3-5	Results demonstrated "the relative ease with which apparently daunting learning gaps can be closed." ¹	Could Pratham also scale using its own resources? Illustrates model for scaling through Pratham staff and volunteers (directly)



**Evidence in
Practice**

Full Report

Print

Case Studies

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Page **118** of **220**

INDIA | WATER & SANITATION

Aqua+

Marketing of a new water purification product illustrates how evidence on behavioral patterns can be as crucial as evidence on scientific effectiveness.

By Erika Drazen





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part I: **The Aqua+ Story**

Antenna and WATA Technology

Antenna Foundation, a Swiss non-profit organization founded in 1989, is engaged in scientific research devoted to the development and dissemination of technologies to improve the lives of vulnerable, base of pyramid (BoP) households.⁸¹ In 2006, Antenna designed a sodium hypochlorite solution to disinfect water called “WATA.” Antenna wanted to test the technology in the field in as many varied contexts as possible, to confirm that it was technically viable. Additionally, a technically viable product would also require an effective distribution and business model to achieve impact. Antenna wanted to learn best practices on how to distribute the WATA technology to BoP customers, yet it did not have the resources or networks in developing countries to do so.

In 2010, Antenna received a grant from the Swiss Agency for Development and Cooperation (SDC) to partner with small and medium enterprises in developing countries. Antenna would give them the WATA technology, teach them how to use it, and give them distribution rights so they could develop and distribute products on their own. Antenna partnered with businesses in eleven countries in Africa and Asia, each with its own product design and distribution model. For example, some partners sold purified water in jugs, while others sold small bottles of droplets. Some chose to place their product in stores, while others chose to deliver directly

to homes. Antenna offered individualized support to each local partner, and collected evidence regarding the opportunities and challenges for each business and distribution model. Each partner gathered and used their own evidence as well, taking into account the local context.

Antenna acted as an intermediary to facilitate learning and interaction among the diverse business partners, and contribute to international learning regarding water sanitation and distribution to BoP customers. For instance, Antenna hosted workshops every year where grantees could learn

from each other. Additionally, with the International Rescue Committee (IRC), Antenna compiled a toolkit based on data from its local partners about monitoring and evaluation, microfinancing, and enabling environments for water purification sales.⁸² Antenna routinely published fact sheets so that other organizations could learn about the work of its grantees. At the time the research for this case was conducted, Antenna was in Phase II of funding from the SDC, which runs through 2018.

History of Development Alternatives Group and Water Sanitation

Development Alternatives (DA) Group, founded in 1982, is an Indian social enterprise dedicated to sustainable development. Similar to Antenna, it researches and designs

Development Alternatives Group



Product testing
+ development



Market research
+ business
incubation



Product sales
+ distribution

⁸¹ Base of the pyramid (BoP) refers to the approximately 4 billion people globally who live on less than US\$2 per day. Social enterprises seek to develop products and services to generate positive impact for BoP households by adopting a market-based, financially sustainable business model.

⁸² “Scaling-Up Safe Water,” Antenna Foundation. [Access here.](#)



Evidence in Practice

Full Report

Print

Case Studies

Page 120 of 220

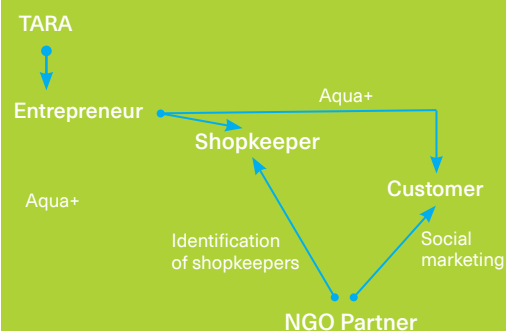
Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

life-enhancing technologies to distribute to BoP customers. To do this, there are three different branches of the organization. Development Alternatives is the research and development arm, with its own lab to develop technology and products related to water, energy, habitat, and waste management. Once a technology has been developed, it is transferred to TARA (Technology and Action for Rural Advancement), which incubates decentralized, small and micro-enterprise business models to sell the products that DA develops to BoP markets. The third arm of DA Group, ultimately charged with marketing the WATA technology, is TARALife Sustainability Solutions, a company under the DA Group responsible for the promotion, manufacture, marketing, and sales of quality-of-life products for BoP households. TARALife was founded in 2014 with the explicit purpose of marketing Aqua+, together with a basket of products deemed to be aspirational for BoP customers. DA Group has operated with this kind of tri-partite structure of developing, testing, and spinning off independent social enterprises for each of its products, including TARA Machines (for manufacturing rural construction materials), TARA Urja (for developing solar micro-grids), and TARA Livelihood Academy (for offering vocational training services).

Business Model: TARA as Production Center



Business Model: Technology Transfer with Entrepreneur Selling Aqua +



DA Group has worked in water and sanitation since the 1990s. Dr. Vijaya Lakshmi, Vice President of Development Alternatives and a PhD in chemistry, was hired in 1990 to develop environmental quality testing kits. Dr. Lakshmi also serves as an expert advisor for the Indian federal government's Department of Science and Technology (DST). She developed a water quality testing kit, which was externally validated and used by UNICEF and the World Bank in the 1990s and early 2000s. The success of these kits gave DA an opportunity to complement them, when appropriate, with a water purification solution. The group researched alternatives—including boiling water, sand filters, chlorine tablets, and other solutions—to purify water effectively, sustainably, and at scale.

The DA Group Decision to Partner with Antenna (2010-2011)

Urs Heierli, who had previously worked for SDC, worked in India as a consultant to both DA Group and Antenna. When Antenna was seeking partners in India, Heierli recommended the DA Group.⁸³ Antenna was an attractive partner for DA Group both because it trusted Heierli's judgment and because the WATA technology fit well into

⁸³ Heierli also recommended that Antenna partner with SpringHealth, which became one of Antenna's grantees for the WATA technology.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

its portfolio. When deciding to develop a new technology, DA's development process typically builds on existing evidence first. Dr. Lakshmi assesses novel findings from different networks (e.g. about fluoride, arsenic, etc.) where researchers share knowledge and evidence. DA selects promising prospects from these networks, as well as journals, meetings, and outside collaborations, and then internally replicates evidence extensively before choosing to incorporate a technology into its portfolio. During this process, DA is concerned not only with the effectiveness of a technology, but also with its underlying potential to produce and distribute a product while making a profit.

Accordingly, before becoming Antenna's distribution partner in India, DA Group tested the WATA technology in its own laboratories to ensure that it was an effective solution for water purification. It found the technology to be both effective and attractive given its small size and ease of use.

Understanding the Needs of the Market (2010-2011)

Upon agreeing to partner with Antenna, DA transferred responsibility to test business models for the WATA technology to TARA. Heierli gave TARA a report to help its staff understand the WATA landscape: while there were pros and cons to different product

and distribution approaches taken in the past, it was clear that there was no universal solution and that it was critical to understand the local context. Another report, by Hystra online, highlighted potential water purification solutions for BoP customers. Aside from these studies and the ongoing learnings from Antenna grantee events, TARA did not systematically seek out other external evidence to inform its foray into the consumer products segment of the rural market, which would factor into subsequent difficulties faced in distribution and sales.

To explore how to convert the WATA technology into a viable and marketable product, TARA contracted EarthSafe Enterprises, a consulting, market research, and training firm. This was consistent with TARA's business model, which depends on keeping lean operational teams while outsourcing other specialized tasks, like market research. EarthSafe and TARA conducted a study in the slums of Delhi, both because of their proximity to TARA's headquarters and their apparent similarity to other target markets.

The study examined drivers of supply and demand through speaking with key opinion leaders in the community. The study consisted of over 200 one-on-one interviews with individuals from both the demand and





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Once TARA chose the business model, it hired a leading market research company ... The study aimed to help TARA understand consumer behavior and the barriers and incentives to purchasing the Aqua+ bottles of drops.

supply sides. On the supply side, the study included doctors, chemists, and shop keepers in an effort to understand potential distribution strategies. On the demand side, the study heard from many potential customers, who said they would prefer to buy the proposed product from a chemist or grocery store. For other programs in their portfolio, TARA typically identified entrepreneurs and then trained and supported them with technology to turn them into distribution partners. Given the nature of the water purification product, however, TARA also explored strategies to distribute directly to customers. From the initial study, TARA and EarthSafe proposed three potential business models.

The first model envisioned TARA as a production center, where TARA would manufacture and distribute bottles of drops and reach the target market through local shopkeepers. The second was a franchise model, where a network of local entrepreneurs would be selected to manufacture and distribute bottles of the solution. In the third model, entrepreneurs would bottle WATA-treated water and sell it directly. The three models relied on local partners (usually an NGO) to help identify potential entrepreneurs and to conduct marketing to increase demand from potential customers. Ultimately, TARA decided to follow the first model—more closely aligned with

its traditional way of working—where it would manufacture and sell bottles of purifying solution to local distributors. Since customers would ingest the product, quality control to ensure customer health and safety was a top priority.

The Creation of Aqua+ (2010-2011)

Once TARA chose the business model, it hired IPSOS, a leading market research company it had worked with before, to conduct a market study. The six-month IPSOS study consisted of interviews and focus group discussions centered on Delhi slums. The study aimed to help TARA understand consumer behavior and the barriers and incentives to purchasing the Aqua+ bottles of drops. Combined with insights from two prominent reports (Hystra's "Access to Safe Water for the Base of the Pyramid" and Marketing Safe Water Systems's report on point-of-use purification), the IPSOS work persuaded the TARA team that the Aqua+ product should be marketed to households (vs. individuals), and the volume per bottle should reflect this.

A second, more complex issue was the periodicity of purchases implied by the volume of a bottle. A small volume would require frequent purchases, which might be affordable but cumbersome. Alternatively, a larger volume would be more convenient



but would allow less opportunity to create new purchasing habits, making it more likely for a household to discontinue use after a first purchase. After consideration, TARA determined that an average family of five should purchase the product once a month, which translated to a prototype 60mL bottle—enough to purify 500 liters of water. The Aqua+ bottle was priced at 20 Indian rupees (then USD 0.33) and had a shelf life of 50 days.

Selecting Distribution Partners (2011-2012)

Early in the process TARA determined that it would not invest in creating its own distribution network, as TARA already had a network of partner organizations that were well-established in target communities. TARA identified a set of criteria to select partners, including NGOs, commercial distributors, and social enterprises. One criterion was that the partner should already have presence in and regular access to a minimum number of households. Second, it should have credibility and a positive reputation in the community. Third, it should have prior experience selling commercial products, not just implementing development programs. Fourth, it should use Hindi as its operating language, so TARA could communicate with them (this automatically limited the geographic scope). As will be shown below and in contrast to its



experiences with other products, it has been difficult for TARA to find reliable, long-term partners for Aqua+.

Innovating on the Aqua+ Product (2012)

As a result of customer feedback on the trial product, TARA and DA made several important product adjustments. First, DA and Antenna worked to extend the shelf life of the solution from 50 days to six months. This would ensure that the drops would remain

effective after the necessarily long journey from production to households through distribution partners. Customer and partner feedback also led to changes in product pricing, packaging, and promotion. Clearer messaging was put on the bottle label, and the final product released in 2013 was priced at 42 Indian rupees instead of the initial 20 rupees, to allow for necessary intermediary mark-ups throughout the supply chain.

Improving the Business Model (2012-2014)

As TARA and its partners continued the distribution of Aqua+, sales numbers were far below expectations. In retrospect, this low level of sales was attributable to two primary factors: operational issues (explored here) and the underlying business strategy (discussed below).

A TARA study revealed that consumers generally liked the product upon their first purchase, but did not buy it again because they could not reliably find it. While TARA received some repeat orders, most partner organizations were not replacing their Aqua+ stock after initial purchases, because they perceived ongoing demand to be low—partly as a result of the long sales cycles—and feared accumulating unsold (perishable) stock. With low inventories that limited product availability, customers quickly



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

reverted to their regular water drinking habits without voicing their interest in the product to sellers. Distribution partners interpreted this lack of expressed interest as a confirmation of low underlying demand for Aqua+.

To seek an increase in sales TARA tasked its existing distribution partners with identifying, coordinating, and managing a network of micro-franchisees, who would sell directly to households. Simultaneously, TARA decided to expand distribution beyond Delhi through new partners and its existing distribution network. By 2014, TARA had distribution partners in the states of Kerala, Tamil Nadu, and Uttar Pradesh. However, these operational changes still did not increase sales as TARA had hoped.

At the heart of the disappointing sales was an incomplete understanding of consumer habits, mindsets, and needs. Before making large strategic commitments, TARA had not routinely or systematically collected information about its target users, except for a few punctuated efforts in limited locations (Delhi slums). In line with this, there was not an open flow of information between TARA's headquarters and the locations where Aqua+ was being sold. Data collection relied upon infrequent and limited visits from TARA staff to the field that, due to personnel and time constraints, could not happen often.

These operational changes still did not increase sales as TARA had hoped. At the heart of the disappointing sales was an incomplete understanding of consumer habits, mindsets, and needs.

Additionally, TARA collected data on paper, which further complicated and lengthened the process of analysis. Once TARA realized that the lack of timely operational and user information was a determining factor in its disappointing sales, it started to develop more immediate, easier to analyze feedback systems.

Further Market Research into Behavior Change (2013-2014)

While increasing repeat sales continued to be an objective of the TARA team, TARA was also having difficulty increasing its base of *new* BoP consumers, particularly in rural areas. TARA believed that potential customers did not see a need to buy the product because they were accustomed to drinking dirty water and did not know that the water was unsafe. Therefore, in 2013, TARA hired IPSOS to conduct a second study, this time specifically on the “use of marketing principles and techniques to advance a social cause, idea, or behavior.” From this study—conducted with the same methodology as the previous one—TARA and IPSOS developed a marketing toolkit to educate potential customers on water safety, which TARA began distributing to its partner network in 2014. TARA experimented with several social marketing methods, including classroom education and village meetings to spread the word about water sanitation and Aqua+. These campaigns sought to communicate that the water families were drinking was unhealthy and discussed different options for treating water, including the proper use of Aqua+. But awareness campaigns did not translate to better sales. One participant from a water and sanitation think tank said:



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

There are a lot of NGOs and governments really trying to convince ... citizens to drink safe water because it's good for their health. But now we see that ... maybe you should try to convince them with other messages, like for instance that it's more convenient and it saves time, and it actually saves money because you have to go to the doctor less, and it saves medical costs.

Like other NGOs across the world, TARA was acting on the assumption that users drank untreated water mainly out of ignorance about its health risks and not based on deeper motivations and tradeoffs. TARA acted on the assumption that customers would buy Aqua+ if they were educated on its health benefits, but did not consider that drinking habits might be based on other economic considerations, underlying values, or patterns of behavior. After visiting five of Antenna's SME partners in India, Nepal and Cambodia, IRC staff said of customers at SpringHealth:

They didn't mention health as one of the biggest factors that drives them to become a customer. They said that it was more about convenience, that the drinking water is delivered at their doorstep. So I think that's quite interesting, because you see that a lot of NGOs really focus on the health aspect,



you know, they really want to create demand for clean drinking water by conveying the message to citizens that they should drink clean water, otherwise they'd become ill ... But now you see that actually, that they are willing to pay, not because it's good for their health, but because it's just more convenient. They don't have to walk I don't know how many kilometers to access water that they don't know is safe. They just get it to their doorstep.

SpringHealth was successfully distributing ten-liter bottles of purified water to customers' doorsteps. SpringHealth had learned, through

a process of iteration, that customers valued the home delivery model over a village kiosk model because of its convenience, as well as the fact that it increased the customers' perceived status and reputation in the community. Yet, TARA continued to emphasize health education (hoping for behavior change) rather than focus on convenience, availability, and economic value and shift to a behavioral model of product demand.

DA executives believed that the lessons from these other experiences were at odds with the distribution and business models they had committed to early in the process. TARA had made a decision to distribute a "point-of use" water-purifying solution—and not bottled water—through local marketing channels to ensure the highest quality and health standards throughout the supply chain and to avoid the distribution costs of delivering and retrieving large jars. TARA concluded that SpringHealth's success hinged on providing bottled water, which it believed carried unacceptable quality or health risks. But SpringHealth had made its choices based on similar concerns and considerations. Most important, its success with customers had less to do with the structural characteristics of its product and much more to do with its understanding of, and framing of the product around, user habits, motivations, tradeoffs, and concerns.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

TARA focused on efforts to educate potential customers on the evident health benefits of regular use of Aqua+, under the premise that users stuck to their drinking habits solely out of ignorance of health implications. SpringHealth, in contrast, examined the underlying attitudes that determined usage and purchasing patterns, and found that the health argument was not nearly as critical for customers as a product's convenience and alignment with their many—often binding—existing constraints. Even though DA Group noted that other products and distribution models were seeming to have greater success, its focus on quality and supply chain concerns prevented DA from identifying the underlying assumptions that were limiting sales of Aqua+.

Throughout the Antenna grant, TARA employees attended conferences on safe drinking water. Companies presented on the solutions they had developed, and TARA would gain insight into other players in the market. While these learning opportunities were not mandatory, TARA staff felt that they got a lot of value from attending conferences as participants and speakers. At the same time, TARA discounted many of the insights from other participants, as the organization felt its own business model decisions made these lessons irrelevant.

Even though DA Group noted that other products and distribution models were seeming to have greater success, its focus on quality and supply chain concerns prevented DA from identifying the underlying assumptions that were limiting sales of Aqua+.

Decision to Spin Off TARAlife (2014)

As discussed previously, DA develops and tests new solutions, which TARA then incubates and develops into sustainable business models. New companies are then created to scale up the solutions once there is a proof of concept. Although Aqua+ did not have steady sales, TARA decided to spin off Aqua+ into a new company, TARAlife, hoping to establish a new business model focused on business-to-consumer sales. To date, TARAlife has continued to face challenges in selling Aqua+ as a stand-alone product.

Narrowing Geographical Focus (2015–Present)

Initially, TARA had distributed Aqua+ to whatever channel partners were interested in it. This led to enormous geographic dispersion, which added to the complexity of gathering timely and consistent operational data. TARA could not familiarize itself with its potential customers and build trust with them, or understand their (varying) specific needs. A TARA staff member, when reflecting on what TARA could have done differently, said:



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

If I could go back and revisit the process, customer engagement is most important. The engagement that TARA had with the end consumer was a bit weak, because we worked through channel partners... We overestimated the connection the partners had with end consumers... The biggest difficulty in selling products was that we were not as engaged in the last-mile part of process as we should have been ...

SpringHealth, another of Antenna's WATA grantees/distributors, had developed strong relationships with customers because it had offices in the locations where it was distributing the WATA technology. TARA decided to move toward a distribution strategy focused on more concentrated geographies, in order to build trust with communities and to establish systems to collect more detailed data on its operations and sales.

About this shift in geographic focus, the project manager for Aqua+ has said:

I've tried to do more focused stuff in a much more intense manner in a particular area to ensure that we see success, we can track it better, we can understand the results better.



By narrowing the geography in which TARA operates, the organization aimed to reduce its dependence on channel partners and establish relationships to truly understand its customers' habits and needs. TARA could build trust and brand recognition, as well as gather better evidence.

Signaling a major shift for TARA, as of April 2017 a range of new products, sourced from suppliers outside the DA Group, has been added to TARALife's suite of offerings. With investor funding, this expanded consumer

products portfolio is now managed by new leadership, which is strategically aligned to goals of the DA Group but operates independently of TARA.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part II: **Key Themes + Insights**

This section discusses the Evidence in Practice themes as they pertain to Aqua+ 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.

Definitions, creation, and use of evidence vary greatly within the Aqua+ case. Within the DA Group, evidence had different definitions depending on the perspectives of the different entities (DA, TARA and TARAlife). Specifically, there are differences in the “burden of proof” for what counts as legitimate evidence and exactly how much evidence is needed to inform decision making.

For DA, a group of “hard” scientists (chemists, physicists, etc.) who design and develop core technologies, there is a high burden of proof for evidence. DA staff constantly have feelers out in the scientific community, reading papers, attending conferences, and talking with colleagues. While the staff had studies from Antenna that demonstrated that the WATA technology worked, they still conducted systematic, rigorous experiments on the underlying technology to ensure that it met the highest standards of reliability and efficacy. In the case of Aqua+, the question was quite clear: the solution either kills bacteria effectively and reliably or it does not. This made the definition and interpretation of evidence relatively easy.

In contrast, TARA’s role was to design and evaluate a business model to sustainably deliver the product to consumers, as well as a marketing and engagement model to change users’ habits and turn them into

repeat consumers. This made the definition, identification, and interpretation of reliable, relevant, and timely evidence much harder. A critical element of this was that TARA framed the central problem as a “supply” question—as an operations and marketing challenge—without an explicit “demand” question—examining issues of user habits, motivations, and tradeoffs. As a result, TARA identified specific moments when evidence would be needed (e.g., prior to making structural, business model decisions), but collected only a narrow set of evidence through an external market research firm and through its own periodic consumer surveys. TARA did not recognize the crucial importance of continuously gathering and incorporating evidence from its end users, outside organizations, or researchers in the field of water sanitation or behavioral economics. One TARA employee said in hindsight:

I would say that Aqua+ has been perhaps the most challenging. Because the technology was only 20, 25% of the effort. 75% of the effort and understanding is how to connect with the customer.



TARA's capacity to invest in systematic customer research was also constrained by its limited funding for this kind of evidence generation. As TARA's CEO, Shrashtant Patara, noted:

I see this as a larger problem. Development organizations and social businesses underestimate the amount of funding required for data collection and analysis that would lead to robust 'evidence building.' Donors and investors are just not willing to part with the required sums of money. They want us to 'implement.'

The focus on operational aspects led TARA to first only look at alternative business models and marketing strategies, while neglecting a deeper examination of behavioral aspects of consumers' decision making. Furthermore, it led the organization to overlook important evidence regarding user insights and behaviors from outside sources because of visible differences in the products and business models. The staff developed alternative business models, selected one, and made adjustments over time. But there was not a comparable effort on the behavioral side, which academic research and the experience of other Antenna partners had shown to be of critical importance. The initial framing of evidence and its accompanying organizational decisions meant that TARA did

The initial framing of evidence and its accompanying organizational decisions meant that TARA did not develop a systematic way of gathering and evaluating evidence on what lay behind lackluster sales.

not develop a systematic way of gathering and evaluating evidence on what lay behind lackluster sales. It developed a different definition of what constituted acceptable and useful evidence, how much evidence was needed in order to make decisions, and the most useful sources of such evidence. It did not help that TARA's initial definition of the challenges led to organizational and operational decisions that structurally hindered its ability to collect useful data. For example, TARA's distribution strategy, which

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

focused on finding partners in many (often remote) locations, meant that TARA remained distant from the field and thus lacked the funding and capacity to reliably and routinely collect data on customers.

This also explains why TARA was exposed to external evidence (for example, at the Antenna partner meetings where actors discussed their experience selling WATA technologies), but did not recognize its value and thus did not integrate it with its internally developed information. The information was seen as inconsistent with TARA's model and not immediately relevant. Even when developing evidence internally, research was only conducted on the existing (operational) situation, and there was little exploration of alternative definitions of the problem.

For instance, TARA learned about SpringHealth's relative success in delivering jugs of purified water. Although the long-term success of SpringHealth's model is yet to be confirmed, the critical insight it offered was how the distribution was framed to fit within user habits, motivations, and needs. However, since TARA had chosen not to distribute purified water due to quality concerns, the organization overlooked the insights embedded within what seemed to be a difference in approach to product design. One TARA employee said:



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Whatever information we have, there is no information that we keep referring to or going back to for making decisions with respect to Aqua+. It is always viewed with respect to whatever the mandate is ... And the mandate has been to come up with the most efficient delivery process. But if you talk about specifically for social marketing, all the information is collected in a very reactive way.

There is a recognition that long-term, sustained research—including on what could lead to behavioral change among potential customers—would benefit decision making within TARA. Unfortunately, Antenna did not hire IRC to conduct this more systematic research until 2014. Had such an effort occurred throughout the entire process, more learning could have been applied to the design of TARA's underlying business model. Additionally, within TARA, staff began to understand that more frequent data collection was necessary to understand and respond to user needs, but the organization lacked the systems to do it cost-effectively. As of 2017, TARA was developing an online app that will allow them to collect more data, more quickly, and at a lower cost. This will remove their current, paper-based process and make timely and ongoing analysis easier. Additionally, this app could be used not only by TARA staff members, but also by



local micro-franchisees, to provide feedback to TARA staff on how different distribution strategies are working. By engaging these micro-franchisees more in the data collection process, TARA can create a shorter feedback loop for itself as well as ensure increased micro-franchisee understanding and buy-in.

Devoting Time and Resources to Integrating Evidence

Few organizations provide incentives or carve out explicit time for managers to explore emerging evidence in their field. Even fewer assign staff to find relevant evidence and translate it into accessible formats for the organization. As a result, the role of preparing and sharing evidence that is timely, useful, and relevant for practitioners is sometimes explicitly played by formal intermediaries (e.g., certain think-tanks). More frequently, an actor who holds a formal role within another stakeholder group spontaneously takes on the (additional) responsibility for trying to integrate evidence, with no actor formally responsible for the process. Discovering and integrating evidence requires time, energy and funding.

DA places a high priority on its researchers serving as evidence producers and verifiers, conducting scientific experiments. Therefore, they systematically incorporate evidence into their work. DA staff also keep updated on current trends with the research in their field. DA staff, like Dr. Lakshmi, go above and beyond their job descriptions to serve the scientific community.

Within TARA, organizational priorities do not systematically encourage either external evidence examination or internal evidence



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

While Antenna and DA Group are open to admitting and discussing the evidence of prior failures, TARA staff remarked that they lack the time or mandate to think about the bigger questions unveiled by prior missteps.

generation and incorporation. This may be because TARA is a social enterprise focused on launching and selling products, constantly under pressure to rapidly respond to setbacks. More often than not, these exigencies do not allow time for examination of other kinds of evidence or of alternative problem definitions, such as the kind of (external) behavioral economics research that was available when Aqua+ was being developed, or the creation and incorporation of (internal) timely, relevant information on user habits.

While Antenna and DA Group are open to admitting and discussing the evidence of prior failures, TARA staff remarked that they lack the time or mandate to think about the bigger questions unveiled by prior missteps.

While there is space to make mistakes, DA Group cannot adequately learn from them as it rarely devotes time and resources to analyze them appropriately. As noted earlier, funders' priorities often exacerbate this tendency. As TARA's CEO noted:

Governments, donors, and sometimes investors, are willing to put funds (sometimes quite large) into the hands of evaluators and consulting firms for these functions. They are extremely wary however, to support 'evidence gathering/building' capacity within social businesses and have not (with one notable exception in our work) been willing to have 'embedded' monitoring, learning and evaluation partners working together with project teams.

This absence of timely and relevant internal evidence has limited DA Group's ability to adjust and improve the design, marketing, and sales of Aqua+. However, as of 2017 TARA has expanded the team working on Aqua+ and, as noted earlier, is building in more capacity for data collection and analysis.

Developing Trust and Relationships

The cross-stakeholder collaborations required for evidence-informed policies and practices are often difficult to initiate, develop, and sustain. Particularly when institutional incentives are lacking, personal trust, respect, and buy-in between individuals across stakeholder groups become critical to fostering the effective flow of evidence into practice.

DA Group initially received funding from Antenna because of the relationship it had built with Urs Heierli, who also had relationships with SDC and Antenna. Heierli's relationships with both organizations were built over a long period of time working for Swiss organizations in the Water, Sanitation and Health sector in India. These relationships provided the foundation for DA's interest and led to its rigorous testing to confirm the scientific evidence underlying the WATA technology.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

While building relationships can help implementers like DA Group learn about existing evidence, key stakeholders also need to examine what kinds of evidence may be lacking and build into the project design mechanisms to collect and share that evidence. Antenna was active in building partnerships among its grantees and sought to connect grantees with others in the broader field of water purification. However, these efforts can only inform practice if actors invest the time and resources to absorb and incorporate emerging evidence.

TARA's narrow definition of its main challenges—and by extension of the evidence that should inform them—meant that TARA missed opportunities to use its existing networks more effectively to identify and incorporate useful evidence. For example, TARA could have developed a much more integrated relationship with Antenna to define an agreed upon definition of goals, evidence to be collected – including lessons from the different environments and organizations—which could have greatly informed the development of Aqua+ from its origin. One DA Group executive mused:

I think perhaps now in retrospect, we could have benefited from a very strong advisory group. As I said, we

have [had] to stumble and get people from the distribution, same products, etc. I think if I was to do it again, my recommendation would be to basically have a very strong, not research advisory group, this is not about research. It is a ... whatever you do for aspirational products. So psychologists, market experts, communication expert, product expert, people who just ask very deep questions.

Learning from Failure

Potential consequences for risk-taking and experimentation with innovative approaches are generally seen as negative and dissuade the exploration of novel, evidence-informed interventions. Fear of failure can further hinder the incorporation of novel evidence into practice, even when stakeholders recognize the value and applicability of the evidence.

Antenna created an enabling environment for its grantees to experiment with flexibility on use and timeline of its funding. Moreover, it was clear that funding from Antenna would eventually cease, so the ultimate goal of the DA Group was to create a self-sustaining social enterprise around the Aqua+ product. There were clear incentives to learn what was and was not working to create a viable business model.

TARA's narrow definition of its main challenges—and by extension of the evidence that should inform them—meant that TARA missed opportunities to use its existing networks more effectively to identify and incorporate useful evidence.



Evidence in Practice

Full Report

Print

Case Studies

Page 133 of 220

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The team at TARA did actively discuss its temporary setbacks and its “certain amount of limited success”. It approached them as learning experiences to make directional shifts. Yet, while TARA adjusted elements of its business model when results fell below expectations, there was not an in-depth analysis of all available—internal and external—evidence, including a consideration of alternative definitions of the underlying problem leading to low sales. While there was a culture to accept failure and adjust from it, there was not the priority, time, capacity or funding to analyze failures systematically. This created certain blind-spots that inadvertently constrained TARA to its initial business model choices and, especially, its initial definition of the challenges to be solved.

Conclusion

DA Group seems to have recognized the limitations in its not giving sufficient attention to cultivation and absorption of evidence that falls outside of its very strong technical expertise. Recently TARA has been examining more thoroughly its potential customers’ underlying motivations and purchasing behavior, and factoring these into its Aqua+ strategy. As one DA Group employee stated:

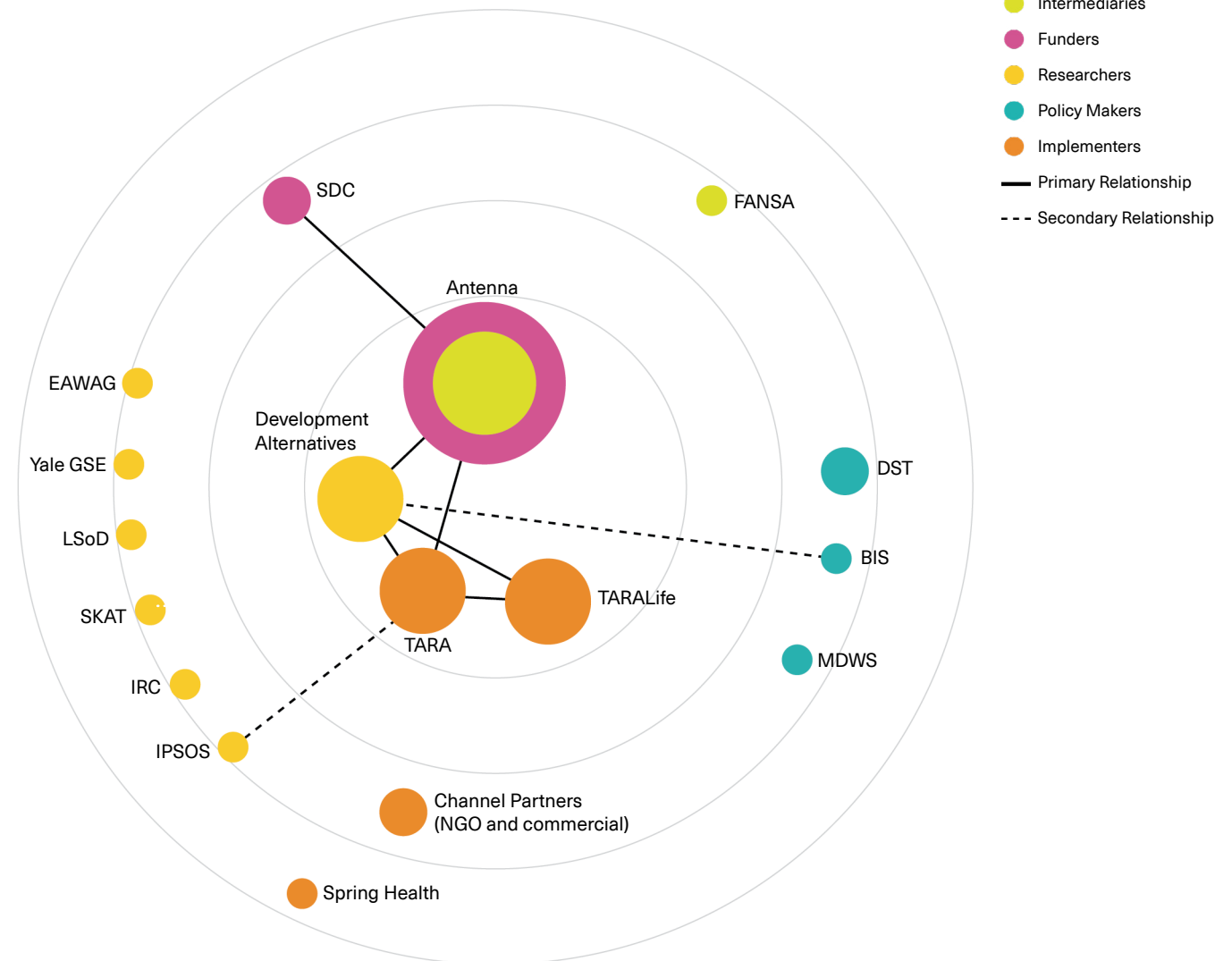


I think after a huge amount of effort [and] investigation, I would say that it was down to one single point. It’s not the cost, it’s not the price of the product. It’s not that people are not willing to buy it. It’s not that people were only experimenting with it. The fact of the matter is that it is not easily available to them when they want it. ...So basically it’s a product which must be easily available to the consumer, virtually at their doorstep, and therefore you need simple mechanism whereby you keep checking with them, keep reminding them, and slowly going ahead encouraging them. But I think how, what those companies have learned and what we have learned, I think in rapid time this has to be top of the mind. Top of the mind means it has to be very easily available. People are letting us believe that in the initial stage people will not go out to buy the product, so it has to be available, shown to them, and made available from them, and that is something which now is let’s say transformed into our business model.

TARA’s emerging strategy for marketing Aqua+ to its base of the pyramid target clients reflects this crucial insight.



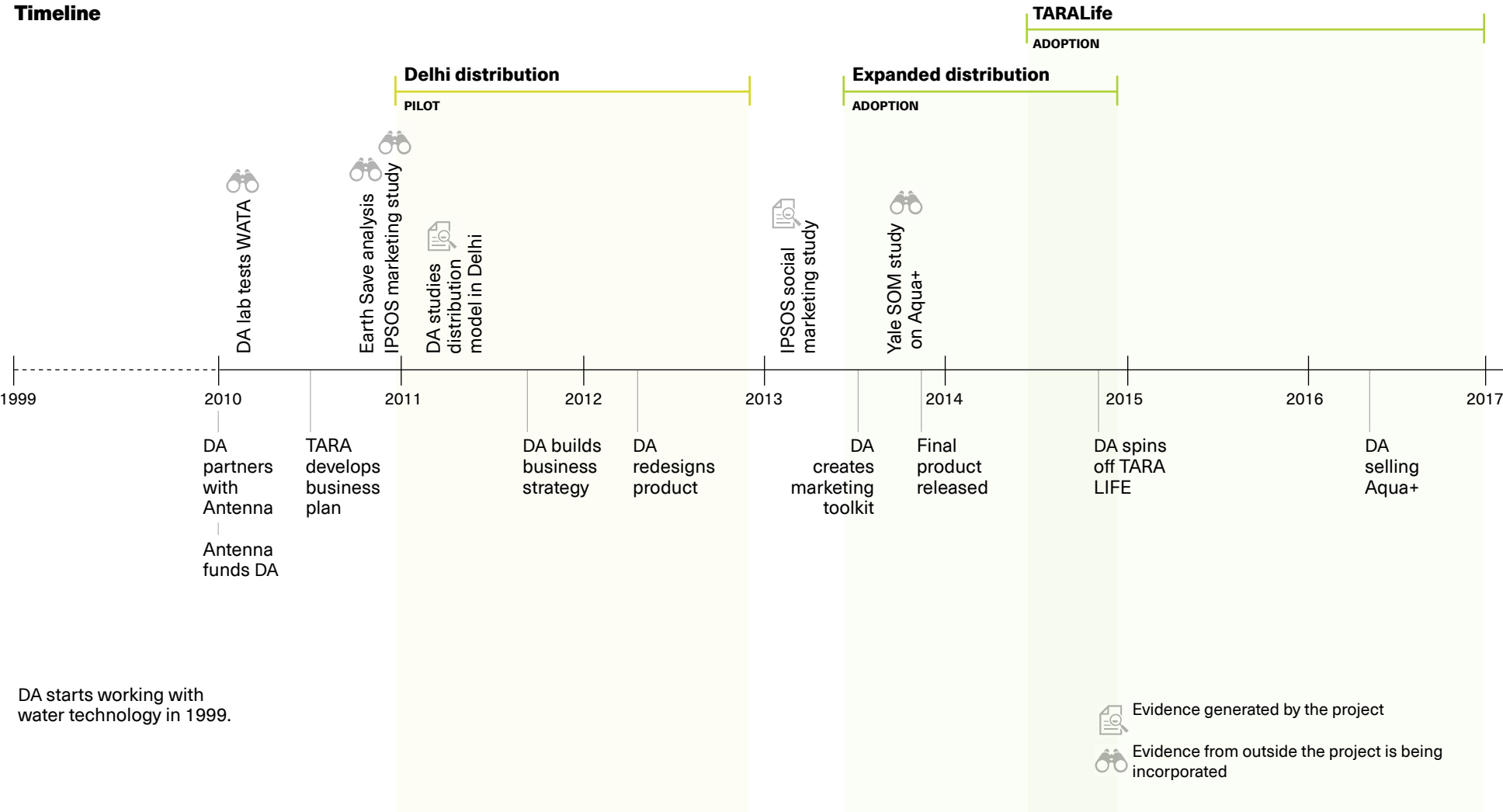
This stakeholder map is a visual representation of the major stakeholders involved with this project. The importance of each of the actors is defined by their relative size, and their proximity to the center of the project. Their role is defined by the color; multiple colors indicate multiple roles. 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.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Timeline

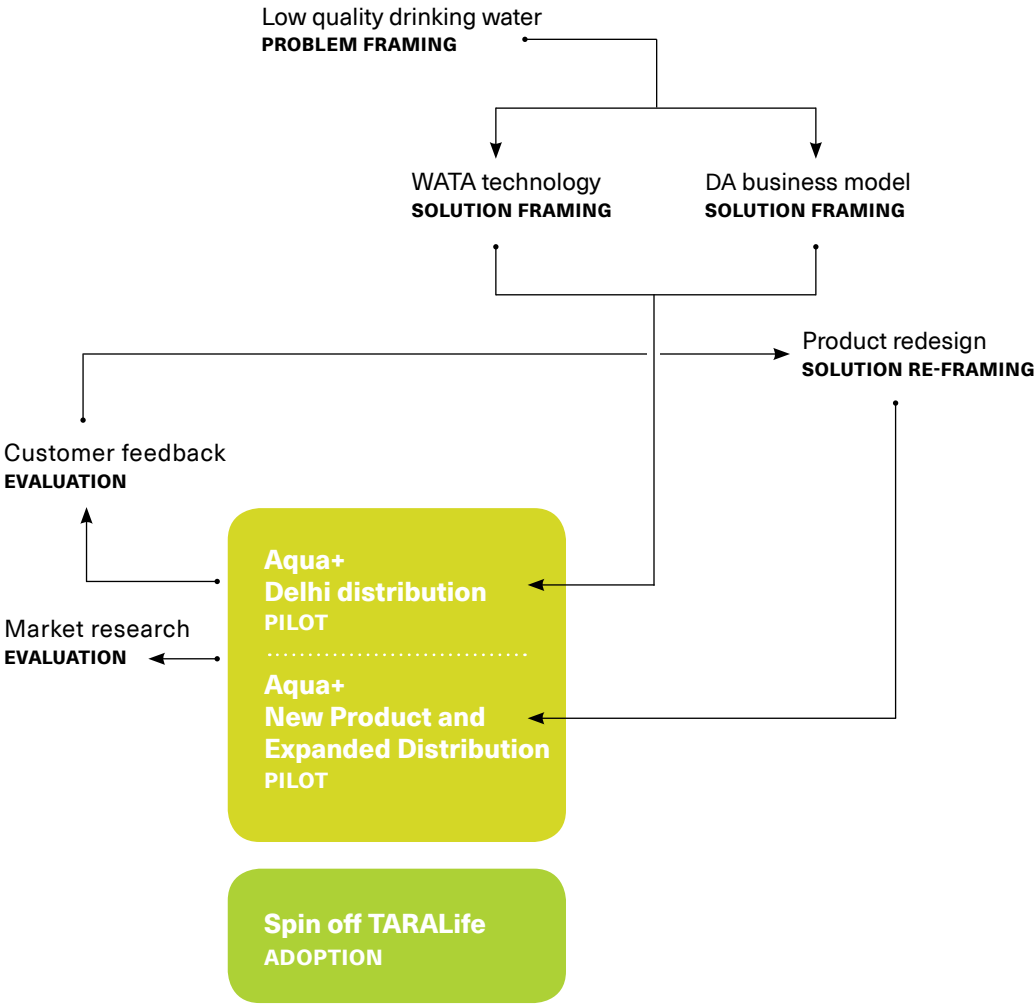




Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Process Diagram

TARALife still struggles to **scale** AQUA+ by matching the solution to market need.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Appendix: List of Stakeholders

(listed alphabetically, bold are actors)

Antenna Technologies/Foundation: A Swiss incubator of innovations for disadvantaged populations; created WATA technology

Channel Partners (NGO/Commercial): intermediaries in villages who buy Aqua+ to sell to consumers for a profit/for humanitarian reasons

Consumers: End-users of the Aqua+ product, usually base-of-pyramid populations

Development Alternatives Group (DA Group): Indian social enterprise devoted to sustainable development, eliminating poverty, and regenerating the environmental resource base through methods that are highly scalable – partner with Antenna to develop WATA technology into selling Aqua+

Development Alternatives (DA): The think tank and research branch of DA Group

Technology and Action for Rural Advancement (TARA): The incubator and market research arm of DA Group

TARAlife Sustainability Solutions Pvt. Ltd. : The commercial Business-to-Consumer arm of DA group

Freshwater Action Network – South Asia (FANSA): NGO that aims to strengthen the engagement of civil society organizations in policy-making and development initiatives to achieve the international targets on water and sanitation

Indian Federal Government:

Department of Science and Technology (DST): a department within the Indian federal government aimed at promoting science and tech

Bureau of Indian Standards (BIS): National Standard Body of India for activities of standardization, marking and quality certification of goods

Ministry of Drinking Water and Sanitation (MDWS): Ministry to accelerate sanitation of and access to clean drinking water

IPSOS: Large and successful survey-based market research company in India

International Rescue Committee (IRC): A large international nonprofit that helps people whose lives and livelihoods are shattered by conflict and disaster to survive and recover



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

SKAT Consulting/Foundation: Swiss-based consulting company working with partners in developing countries on international development issues; Foundation for knowledge sharing and learning

SpringHealth: NGO similar to TARA developing Antenna's WATA technology to bring to market in India

Swiss Agency for Development and Cooperation (SDC): Switzerland's international cooperation agency within the Federal Department of Foreign Affairs; funded Antenna who in turn funded DA Group

Swiss Federal Institute of Aquatic Science and Technology (EAWAG): Research organization that's part of ETH group concerned with concepts and technologies for dealing sustainably with water bodies and with water as a resource; bridges science and 'real world'

Yale School of Management: Students worked with DA Group on a short term project related to distribution of Aqua+



**Evidence in
Practice**

Full Report

Print

Case Studies

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Page 139 of 220

MEXICO | ECONOMIC DEVELOPMENT

Progresa | Oportunidades

A pioneering national program incentivizing school attendance and healthcare through conditional cash transfers.

By Victor Cervantes and María del Mar Gutiérrez





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part I: **The Progresa Story**

In 1994, presidential elections were held in Mexico. That year, Mexico experienced its highest level of income inequality in the past 30 years.⁸⁵ Shortly after the winner of the elections, Dr. Ernesto Zedillo, assumed office, a profound economic crisis struck the Mexican economy. The crisis had an especially large impact on the poor. The extreme poverty rate grew from 21.4 percent to 37.4 percent between 1994 and 1996.⁸⁶

At that time, most poverty alleviation programs consisted of food subsidies. In 1996, roughly two thirds of the subsidies were untargeted, generalized price subsidies and more than three fourths of the subsidies went to urban areas, leaving rural areas—where extreme poverty was more prevalent—underserved.⁸⁷ Food subsidies were expensive, did not reach the poor, and were routinely abused as tools of political manipulation—for example, exchanging subsidies for votes. The crisis prompted the urgency of supporting the poor more effectively within the added constraints of a reduced budget. A body of academic literature had been generated before 1995 advocating for a paradigm shift in anti-poverty policy. Building on this literature, Progresa (“Prosperous”) marked a notable departure from existing social policy in Mexico, developing a groundbreaking program based on targeted, conditional cash transfers. Progresa (later renamed Oportunidades, or “Opportunities”) illustrates several key ways

in which barriers that tend to inhibit the integration of evidence into practice can be overcome. First, it demonstrates the ways in which embedding evidence into the fabric of a program from the beginning can align the (often conflicting) incentives of different stakeholders and overcome issues related to discordant decision-making timelines. Second, it shows that developing deep relationships with, and engaging a diverse group of stakeholders can be critical to developing trust—necessary to implement a rigorous evidence-based program.

⁸⁵ Esquivel, Gerardo, Nora Lustig, and John Scott, (2010). “A decade of falling inequality in Mexico: market forces or state action?” Declining inequality in Latin America: A decade of progress.

⁸⁶ Cortés, Fernando, and Rosa María Rubalcava, (2012) *El PROGRESA como Respuesta a la Crisis de 1994*. International Policy Centre for Inclusive Growth, No. 169.

The Story of Progresa: The Program and its Objectives

The program's main objective was to disrupt the intergenerational transmission of poverty by encouraging and supporting household investments in child education. In particular, the program combined health, education, and nutrition elements into a single intervention, creating reinforcing effects between the three. In essence, the program consisted of a cash transfer to eligible households conditioned on continued and demonstrable school attendance by the family's children, coupled with routine health and nutrition checkups for all household members. During the checkups, nutritional supplements were provided for children under the age of five and pregnant or nursing women.⁸⁸ The cash transfer was given to the female head of the household, seeking to improve the allocation of resources within the family.



Dr. Ernesto Zedillo

⁸⁷ Levy, Santiago. “Progress against poverty.” Washington DC: Brookings Institute (2006).

⁸⁸ *Ibid.*



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Progresa marked a dramatic departure from existing social policy in Mexico. It was the first program in Mexico to explicitly combine the three elements of healthcare, nutrition, and education. Progresa was targeted to specific families according to their needs and characteristics, which were clearly articulated in advance through clear rules that were defined by rigorous statistical models. That meant not only that the beneficiaries were likely to be those most in need of the subsidy, but also that once the criteria were established, eligibility would not be manipulated by policymakers, implementers, or beneficiaries. This shielded the program from political abuse, as discussed below.

While the program's main objective was to limit the inter-generational transfer of poverty, the intervention also alleviated current poverty. Initially, the larger component of the cash transfer was calibrated to offset the household opportunity costs of keeping children in school, according to their age and gender (accordingly, the amount transferred varied by family size and composition). As the program evolved, the alleviation of current poverty became an increasingly important objective. Based on its impact, the program has been the most effective instrument in Mexican social policy for redistributing income to the poor.⁸⁹



Initial Design, Pilot Project, + Poverty Measurement

After assuming Mexico's Presidency, Mr. Zedillo⁹⁰ appointed Dr. Santiago Levy to be the Deputy Minister of Finance, whose portfolio would include preparing the budget and approving funding for federal programs. Dr. José Gómez de León was appointed to chair the National Council of Population (CONAPO). At that time, Levy and Gómez de León were public officials with considerable academic backgrounds. Levy had obtained a Ph. D. in Economics in 1980 and had been a faculty member at Boston University. His research focused on macroeconomics, international

trade, and development. In the early 90's, he was hired as a consultant for the World Bank to evaluate anti-poverty programs in Mexico. In his work, he questioned the effectiveness of traditional programs, such as the tortilla subsidy.⁹¹ In 1992, he joined the Mexican government as the first President of the Federal Commission for Competition. Gómez de León started his career at the Mexican Central Bank, where he and Mr. Zedillo met. He obtained a Ph.D. in Demography and was a faculty member at El Colegio de Mexico, researching the intergenerational transmission of poverty before joining the Mexican government as the Chief Advisor to Zedillo at the Ministry of Budget and

⁸⁹ Scott et al., 2010

⁹⁰ In 1974, President Zedillo pursued a Master's degree and a PhD at Yale University. He worked at the Mexican Central Bank and served in two Ministries (Finance and Education).

⁹¹ One of the programs started in 1986 as a generalized subsidy in urban areas that gave the right to buy tortillas at a subsidized price. In 1990, Fidelist emerged, a new directed scheme that granted the right to obtain a free kilo of tortilla every day. These two programs coexisted until October of 1998, when the generalized subsidy was eliminated. In November of 2000 Fidelist was phased out. [Levy, Santiago; Evelyn Rodríguez. *Sin Herencia de Pobreza*. Banco Interamericano de Desarrollo – Planeta. 2005. Pages 119-121.]



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

“Within Progressa, we never thought about the development of the structural aspects of the program without always thinking of the operational aspects as well.”

RESEARCHER

Programming. He later served as Director of the National Institute of Public Health. Gómez de León was very close to Zedillo, as they were longtime collaborators.

In late 1994, Zedillo, based on a set of recent publications that extensively and rigorously analyzed the shortcomings of existing paradigms for poverty alleviation, asked Gómez de León to design a program to fight poverty based on conditional cash transfers. Early on, Gómez de León and his team⁹² proposed a rough design of the program, based on the evidence available to them, including a rich body of literature. The team paid particular attention to recent work by the World Bank discussing the diverse qualities,

problems, and benefits that different types of policies had shown, and different alternatives for fighting poverty.

Gómez de León and his team had been researchers, and they collected additional quantitative data to inform the design of the program. In particular, a critical initial challenge was identifying the target population, as Mexico lacked an official measure of poverty, and existing data was insufficient to target at the household level. In 1995, using their previous work on marginalization indexes at the municipal level, they created tools to estimate poverty at a more granular level. In particular, the team began to roll out a national survey that

would allow the federal government to create a more rigorous measurement of household-level poverty to effectively target eligible program recipients. The survey sampled communities where poverty was expected to be more prevalent according to the general marginalization indexes. The instrument, named “ENCASEH” (Survey of Socioeconomic Characteristics of the Households) sought to collect granular, household-level baseline data on health, education, and consumption habits. These data were designed to allow for the precise targeting of an intervention and would become instrumental for the randomized controlled trial that later served as the initial pilot for the program. The team also gathered information on the school drop-out rate by gender. The survey was designed to estimate the size of the subsidy required to offset a family’s opportunity costs of keeping children in school, a cost that increased with a child’s age. Additionally, information on the local supply of education and health services had to be gathered, as it was necessary to ensure local capacity to absorb the increased demand for health services that would be generated by the program.

⁹² Mr. Gómez de León’s team included Rodolfo Tuiran, Daniel Hernández, and Monica Orozco.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

In parallel, the Deputy Ministry of Finance, Mr. Levy, and his team⁹³ had been working independently on the design of a similar program, seeking to address the inter-generational transmission of poverty. An initial design of the program—one that captured the essence of what became the final design—was presented in March of 1995. Levy and his team were strong proponents of the program, had control over the budget, and understood the broader political and financial situation of the Federal government. To test the initial program design, Levy's team conducted a pilot of 31,000 households in the state of Campeche along with the Ministry of Social Development during the first half of 1996. The pilot was designed to generate evidence to analyze the conditional cash transfer scheme. The name of the pilot was “Basic Food Basket Program for the Wellbeing of the Family”. The pilot was externally evaluated by the Autonomous Technology Institute of Mexico (ITAM) and by the National Institute of Nutrition Salvador Zubirán.

The initial pilot showed that, as hypothesized by the program's designers, directing the subsidies towards women enhanced the effects of the program, as it led to a more judicious use of the transfers within the households and did not create feared secondary consequences, such as domestic violence. The pilot thus demonstrated



One of the programs started in 1986 as a generalized subsidy in urban areas that provided the right to buy tortillas at a subsidized price. 1990 saw the emergence of Fidelist—A new directed scheme that granted the right to obtain a free kilo of tortilla every day. These two programs coexisted until October of 1998, when the generalized subsidy was finally eliminated.

that a targeted cash transfer was much more efficient and effective than traditional untargeted subsidies. It also revealed that building some sort of conditionality into the program, as the program designers had long believed, was important to further amplify positive effects by, for example, linking cash transfers to useful behavioral changes within the household. Finally, the pilot showed that household members preferred cash to in-kind transfers and that they valued the healthcare component.⁹⁴

After two of his trusted collaborators independently developed similar program proposals, President Zedillo instructed them to join efforts and became particularly invested in the success of the unified program, turning it into his flagship social policy. The teams divided the next phase of work: Levy and his team created technical papers documenting the Campeche pilot to support the claims made; Gómez de León's

⁹³ Mr. Levy's team included Evelyn Rodriguez, Antonio Alvarado and Enrique Davila.

⁹⁴ Levy, 2006



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

team continued collecting rigorous, granular data to target households for the next iteration of the program. Several elements of the pilot in Campeche were modified in further refinements of the program, including the addition of a conditionality on child education, which was not present in the pilot, and providing the transfers in cash instead of doing it through electronic cards. The electronic cards were dropped as the media and program detractors had criticized them, making them politically toxic. The pilot had enabled the team to convince the President and most of the Cabinet that the program was both viable and effective, which supported them in the face of stiff political opposition: based on longstanding tradition, the Ministry of Social Development opposed targeted or conditioned cash transfers and strongly favored generalized subsidies.

Levy was respected among academics, but he also demonstrated managerial and political skills. In parallel, Gómez de León was in charge of demographic information and policy, and had the skills and a team capable of gathering, creating and analyzing a wide range of data. Both Levy and Gómez de León were technocrats but showed sensitivity towards political and operational constraints. They were technically strong and committed to sound policymaking. They both had strong personalities, but did not let their

Against government tradition, Zedillo decided to wait for a thorough and tested design, instead of launching—with typical fanfare—an ambitious program at the beginning of his administration.

egos come in the way of the program. Rather, contemporaries describe them as recognizing each other's complementary skills and establishing a productive partnership. Gómez de León is regarded as leading the implementation and having a great deal of influence on program design. Mr. Levy, also influential in program design, is regarded as an agile leader who oversaw the necessary political negotiations. President Zedillo's leadership was also a necessary condition for the success of the program,

as he was directly involved with Gómez de León and Levy reporting directly to him. The accompaniment, encouragement, and ownership by Zedillo was key. The leadership that Gómez de León, Levy, and Zedillo showed, as well as their background as researchers who had become policymakers, proved to be very powerful.

Against government tradition, Zedillo decided to wait for a thorough and tested design, instead of launching—with typical fanfare—an ambitious program at the beginning of his administration. In August of 1996, two years into an administration under enormous political strain from the 1995 crisis, but not until the results of the pilot project had been presented, the President authorized an expanded intervention named Program of Nutrition, Health and Education (PASE).

Launching the Program

Before August of 1996, the participation of other government agencies in the program had been minimal. At that point, given the decision to incorporate a behavioral conditionality into the program—school attendance and visits to physicians—the involvement of the Ministry of Health and the Ministry of Education became crucial. While in the pilot, for example, the health interventions were monitored by a group of



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

doctors hired specifically for this purpose, a larger roll-out would require using the public health system. Given its federal structure, health and education services in Mexico are decentralized and managed by each state government. Thus, certifying the fulfillment of the behavioral commitments on education and healthcare by the beneficiaries required coordination with state governments, including negotiations on how to launch and monitor operations in each state. This was not a trivial issue, as the new program design identified target households ex-ante based on technical factors, removing the local government's long-held control over the distribution of government subsidies.

Levy took the leadership in engaging these two key ministries as well as state governments. Zedillo provided top-level support (and often pressure) to create political goodwill among other government stakeholders. An inter-ministry group, with the Chair of CONAPO and the Deputy Ministers of Finance, Education, and Health, was created to refine the operational aspects of the program. Zedillo's involvement was particularly important to enable the coordination between ministries and with state governments. Although the path to implementation was not without resistance,

government stakeholders showed the will to collaborate and commit to the program as a result of the leadership, support, and pressure from President Zedillo.

During that time, weekly discussions were held among the inter-ministry group⁹⁵ on the operating and technical aspects of the program. Questions ranged from the amount of the transfers to the specific formula of child nutritional supplements. Gómez de León and his team built a database with the geo-location of the communities and households intended for the program in order to draw the routes to go to the field and disburse the cash. Also, negotiations with the local authorities to help with the logistics started. This level of inter-governmental coordination proved instrumental to the success of the program. Mr. Levy and his team played a central role in conducting negotiations and securing federal funding for the program, most of which came from shifting funds previously used for the generalized subsidies. The pushback from the proponents of generalized subsidies was relentless, partly because this had been the prevailing paradigm, but also because of the room this approach provided for political manipulation. Such opposition only increased the team's commitment to rigorously and transparently evaluate the new program's impact.



⁹⁵ Group with the Deputy Ministers of Finance, Education, and Health, and the Chair of CONAPO.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The program was ready to launch in December of 1996—and its funding was approved in the Federal budget of 1997. However, Mr. Zedillo’s administration decided against launching before the midterm elections to be held in July of 1997, to avoid suspicion of political use of the program and again departing from traditional electoral norms in Mexico. The name of the program was changed to Program of Education, Health, and Nutrition (Progresa). It was the first social program to carefully and transparently define operating rules and principles. The rules were intended to increase transparency and shield the program from political interference. (Based on this program’s experience, explicit rules of operation subsequently became a standard practice for all social programs in Mexico.)

Implementing the Program

Zedillo delegated program implementation to Gómez de León and created the National Coordination of Progresa as a decentralized body inside the Ministry of Social Development (which was the ministry that had traditionally managed the generalized subsidies and thus was the most ardent opponent of Progresa). To maintain appropriate inter-ministry coordination, an oversight board was created, with permanent seats for the Ministers of Health, Education, Finance, and Social Development. In June of

1997, Gómez de León and his team moved from CONAPO to the National Coordination of Progresa. At that time, Progresa’s National Coordination was a young and small organization of no more than 30 people. The team was led by Gómez de León, and received support from Levy and his team.

The roll-out of the program, which had to be gradual given resource and operational constraints, was designed as a randomized control trial (RCT) to create rigorous evidence for subsequent program refinement and growth. Out of the group of potential beneficiary communities, a total of 300,000 households were selected as the treatment group and the rest were assigned to the control group, producing an experimental design that allowed for the rigorous testing of several hypotheses about the program’s effects.

This experimental approach had its critics. There was a vocal group, including some in the press, who thought it was outrageous for the government to conduct experiments with social programs, as they thought it was unethical to arbitrarily deny access to some households. Two factors made the randomized control trial socially and politically viable. First, the randomization was made with the community, not the household, as the treatment unit. If a household was

in the program, its neighbors would also be in the program. That eased suspicions that some households would be arbitrarily (or politically) excluded from the program. It also simplified program operations. While choosing the community as the unit instead of the household reduced the inference power of the RCT, the program was big enough to cover a large number of communities and draw statistically reliable conclusions. Second, all the households in the control group would eventually added be to the program. More importantly, a large number of household would have to be excluded in the beginning, even in the absence of the RCT, for purely operational and financial considerations. Random assignment to the treatment or the control groups was, in fact, an arguably fairer decision than one made based on political or operational considerations.

Evaluating the Program

During 1997, Levy and Gómez de León had explored alternatives for a rigorous, transparent, and external evaluation of Progresa. Although there were costs associated with having a robust evaluation, the benefits far outweighed them. Progresa was a departure from traditional policy in several respects. First, it was targeted and based on objective, verifiable rules of operation. Second, the program was



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

explicitly forbidden from being used for political objectives, and beneficiaries were systematically informed of this. Third, the program explicitly forbade adding beneficiaries during political campaign seasons. Finally, the program was organizationally shielded from political influence: program officers were, by mandate, selected based solely on their technical merit and were non-partisan. Furthermore, no career politician was allowed to participate. As a result, there was enormous political pressure to change or scrap the program. Only evidence of its effectiveness, validated by a legitimate, external institution could provide political cover. The evaluation would thus have to be performed by an institution that was viewed as independent from the Mexican government or other political actors in Mexico, and that had reputable evaluation experience.

The program designers did not want a visible organization, such as the Inter-American Development Bank (IDB) or the World Bank, to be involved in the initial evaluation in order to avoid suspicions of foreign financial institutions influencing the program. Nora Lustig, who was then a researcher at the Brookings Institution, had mentioned the International Food Policy Research Institute (IFPRI) to Levy in 1995. Additionally, some



economists at the World Bank encouraged Evelyn Rodriguez, who worked with Levy, to consider IFPRI during a visit to Washington D.C. in 1997. Levy and Daniel Hernández Franco, who was part of Gómez de León's team, were aware of the work by IFPRI on the effect of nutrition programs on poverty. Levy and Gómez de León decided to hire IFPRI to conduct the evaluation, but they thought that

it would be desirable to have some senior participants beyond IFPRI staff to give it greater visibility and credibility.

Gómez de León asked Lustig, who had been recently appointed at a senior position at the IDB, to organize a small workshop, and gave her a set of names of people whom the team wanted to get involved. The workshop took



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love



place in December of 1997 at the IDB, with IFPRI representatives and some potential senior external advisors, including Jere Behrman from University of Pennsylvania, Paul Schultz from Yale, and James Heckman from University of Chicago (who could not attend but invited Petra Todd instead). Paul Gertler, from UC Berkeley, later became part of the team, as he learned from Hernández about the project and asked to be involved. As the program was to be expanded very rapidly from the initial experimental rollout, data collection and evaluation was integrated into program operations from the beginning. The IFPRI team, the four external advisors, and the Progresa staff worked out the details of designing the questionnaires used in the evaluation. The contract with IFPRI lasted for three years, after which it was thought that the credibility of the evaluation would have been established and a strong domestic institution could continue the evaluation. There was a contract with the National Institute of Public Health⁹⁶ (INSP) to continue the evaluation.

The IFPRI evaluation found that Progresa had positive effects on the majority of the intermediate targets. There was no evidence of the effects on cognitive skills, but what was clear was that the behavioral conditions in the program had significantly impacted household habits around nutrition and child education.

The original program provided cash transfers to mothers who could demonstrate continued enrollment of their children in school up to 9th grade. The evaluation, in particular the work by Paul Schultz, found that the conditional cash transfer had the strongest impact on the enrollment in secondary school (grades 7 to 9). Based on this, the leadership of the program decided to extend program eligibility for households with children enrolled up to 12th grade. In contrast, the impact evaluation showed that the nutritional supplements had a negligible impact on anemia, due to a mistake in the iron formula that caused low absorption. Some operational evaluations showed that this was compounded by a lack of household education and monitoring, which translated to low intake of the nutritional supplement, and that appropriate “consumption declined as indicators of poverty increased.”⁹⁷ Even though the evaluations detected the problem as early as 2001, the iron formula was not changed until 2005. This was partly because, while the supplements did not have a measurable impact, they certainly posed no harm and were popular among beneficiaries, so they helped improve compliance with other program conditions.

⁹⁶ Mr. Gómez de León was the director of the INSP before being invited to lead the CONAPO.

⁹⁷ Levy, 2006, p. 55



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

In addition, there were other, more critical, adjustments to the program that were prioritized, such as expanding the program to urban populations (discussed below).

After launching, the National Coordination of Progresa committed to a fast scale-up, with the support of Levy and Zedillo. Beneficiary households increased five-fold by the end of 1998 and by the end of the year 2000 the program reached almost 2.5 million households.

The New Federal Administration: 2000-2006

In April of 2000, José Gómez de León passed away and Daniel Hernández assumed the leadership of the National Coordination by presidential appointment. In July of 2000, Mr. Vicente Fox won the Presidential election. That was the first time in more than seven decades that a political party other than the Institutional Revolutionary Party (PRI) had won a presidential election.

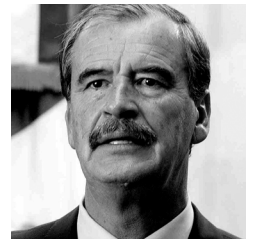
Levy shared the program's impact evidence with Mr. Fox as President Elect. Also, Hernández talked to Fox and Vázquez Mota, who was appointed as the Social Development Secretary, about keeping and strengthening the program. The body of evidence that had been produced was key

As the program was to be expanded very rapidly from the initial experimental rollout, data collection and evaluation was integrated into program operations from the beginning.

during that process. The President Elect asked specifically about the results and rigor of the evaluations. Because of the program's performance evidence and IDB's and the World Bank's positive opinions of the program, President Fox decided to maintain the program. Under the new administration, the program's core design and components were maintained, with some adjustments, including changing its name to "Oportunidades" in

2002. President Fox decided to increase the coverage from 3 million families to 5 million in the first years of his administration. That required extending the coverage to urban areas, as almost total coverage of the poor in rural areas had been already achieved.

The first director of Progresa to be appointed by the new administration, in January of 2001, was Vicente Arredondo. He was replaced by Rogelio Gómez Hermosillo in August of 2001. Mr. Gómez Hermosillo came from the NGO sector and did not have experience with conditional cash transfers. However, he maintained key people in operations who had worked from the beginning and reappointed others who had left during Mr. Arredondo's brief tenure. He later became an expert on conditional cash transfer programs. The program was rigorously audited by the Fox administration under the suspicion that the program could have been misused or that money could have been diverted from its intended purposes. This placed a burden on the operational teams and created stress for the public officials who had been in the program during the previous presidential term. The audit found that no misuse had taken place, which strengthened the legitimacy of the program and the Fox administration's respect for what had been done.



Vicente Fox



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The National Coordination faced the challenge to develop the technical tools to increase the scale and scope of the program. During the initial years of Fox's tenure, Progresa/Oportunidades had clear operational rules and was well established in the budgeting process, so the Fox administration simply kept the program running based on its previous design and specifications. In order to scale the program, however, Mr. Gómez Hermosillo and his team had to redesign the targeting systems and to recalibrate the parameters of the cash transfer, as the mechanisms designed for rural areas were not directly applicable to urban areas. They spent 2001 and 2002 on this endeavor, and in 2003, the strategy for the urban expansion was finalized. Unlike the rollout in rural communities, there was no pilot or RCT for the urban expansion. The operation of the transfers, control systems, planning systems, communication with families, and co-responsibility review systems were improved, but important opportunities to strengthen evaluation were missed, including designing an evaluation for the urban rollout. At the end of Mr. Fox's term in 2006, the program reached 5 million households.

During Ms. Vázquez Mota's tenure as Secretary of Social Development under President Fox, and based on the Progresa/

Oportunidades experience with rigorous evaluation, the National Council of Evaluation (CONEVAL) was created with the mandate to evaluate the government's social programs. The systems for monitoring and evaluation fostered by CONEVAL took important lessons from Progresa, and in fact all the members of CONEVAL's first advisory board had been evaluators of that program.

In 2001, the Office of Public Debt at the Ministry of Finance brought a team from the IDB to explore giving a loan to Mexico, tied to the maintenance of Progresa. The program was fully funded through the federal budget, so it did not require additional resources. Rather, the loan was negotiated as a strategic protection so the program would continue under its core operating rules and guardrails. In fact, the loan agreement gave the federal government ample discretion on the use of the loan proceeds. Changes in core elements of the Progresa program, however, would trigger the immediate payment of the principal plus some termination fees. The goal of the loan, then, was to shield the program from political cycles or temptations at manipulation. In 2002, the IDB granted a US \$1 billion loan to the Mexican government, which was renewed in 2005 for US \$1.2 billion. While this was not sizable compared to the program's, let alone to Mexico's, overall

budget, the loan was large relative to the IDB's lending history. Rather, the partnership with the IDB, which included technical assistance and monitoring, as well as contractual obligations that contained automatic repayment triggers following alterations to the program's operating rules, provided firm technical and political guardrails.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part II: **Evidence in Practice: Key Themes + Insights**

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

Evidence Definition, Creation, + Use

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.

Definition of Evidence

As described in the case study narrative, Progresa was conceived based on the most rigorous academic evidence available at the time. Designers centered their work on a rich body of literature, paying particular attention to the work of the World Bank on social policy and on the discussion of in-kind vs. monetary transfers. As Progresa developed, the continued common thread was the

importance and use of rigorous evidence both to make programmatic decisions and to provide strong political support, rooted in the randomized control trial (RCT) that was embedded in the rollout of the program. The conceptualization of evidence, the definition of program objectives, and the philosophy for policy design brought in by Progresa were diametrically opposed to the traditional incentives and philosophies for policymaking in Mexico. While Progresa was always defined by the demonstrable impact it could have on its beneficiaries, policymaking in Mexico had traditionally focused solely on political dimensions and implications. Progresa, therefore, brought in a radically different definition both of what category of evidence was most relevant for policymaking and what actually constituted reliable, actionable evidence.

This new approach to integrating evidence into policymaking was so influential that it not only informed the evolution of Progresa but also led to the creation of the National

Council of Evaluation (CONEVAL), which now brings in rigorous methodologies to the evaluation of all social programs in Mexico, illustrating how the effective integration of evidence into practice can translate to new structures for routinely integrating evidence into practice in other areas.

Evidence Creation

Progresa/Oportunidades was a departure from the traditional design of social policy in Mexico. The case narrative shows the important role evidence plays: from committing to generating it, to collecting and using evidence; while also being flexible enough to take the risk of a new idea for which the evidence was not all in hand. The conditional cash transfer model had not been empirically tested at the time. What had been demonstrated, however, was that generalized and in-kind subsidies—which had dictated the paradigm of social policy in Mexico up until that moment—were grossly ineffective. The decision to launch the program came before there was a randomized controlled trial to prove the effectiveness of the model. The team, partly because of its background in academia and also because of the humbling wave of evidence that had just been produced regarding existing social programs, knew



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

the importance of designing policy based on evidence, and the importance of generating new evidence both to evaluate progress and to inform future policy. The Progresa team thus made it a priority not only to research the best available evidence, but also to generate new data and evidence when necessary to inform their policy decisions.

The decision to pilot and evaluate before fully launching the program was crucial, allowing policymakers to use the data to understand the impacts of the program and the components that needed to be improved. Also, precisely because of the rigor and transparency with which they were generated, these data provided critical political strength to carry on with the rollout and scale-up of the program against enormous political resistance. The information produced was essential not only for conceptual policymaking, but for successfully translating the theory of change into a feasible and actionable large-scale program.

Another key factor the case study illuminates is the importance of not limiting the use of evidence to the design phase, but integrating the generation and use of evidence into every phase of the program so its quality and effectiveness can be assessed and adjusted based on the findings of the ongoing

The Progresa team made it a priority not only to research the best available evidence, but also to generate new data and evidence when necessary to inform their policy decisions.

evidence created by the program. It is also worth considering that, throughout the design and rollout, the program retained significant flexibility to incorporate lessons from any new findings that emerged from program operations. Evidence and practice were thus integrated fully throughout design and ongoing implementation—with a natural evolution on both the scale and type of decisions that were made and the most relevant evidence to be generated. Put differently, the best lesson to learn from Progresa as a policy success is the process that led to its creation much more than the specific characteristics of its conditional cash

transfer policy. Ironically, this lesson was often lost, even on those with experience with the program who decided to scale it into urban areas (where it proved less successful).

Uses of Evidence

The paradigm shift sought by Progresa implied not only a different conceptualization of social policy, but also and more importantly, the loss of a powerful and effective tool for political mobilization. Accordingly, one of the key motivations to document and produce evidence from the very beginning was to provide powerful elements to face and overcome relentless political pushback



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

from different actors, including from the proponents of generalized subsidies, members of Congress, and influential members of the governing political party.

Another key lesson from Progresa is how a program that generates and adapts to rigorous ongoing evidence can increase its own probability of survival across political cycles. The year 2000 marked a watershed moment in Mexico's democracy. For the first time in 70 years the PRI lost the presidential elections to the National Action Party (PAN). As mentioned above, designers and implementers of Progresa shared the program's impact evidence with Vicente Fox as President Elect and Vazquez Mota, who was appointed as the Social Development Secretary. The program's positive results, the rigor of the evaluations, the fact that the evaluations were performed by external trustworthy institutions, and the large body of evidence that was produced and widely disseminated were key to what seemed impossible: maintaining the flagship social program of the previous administration and its now maligned party.

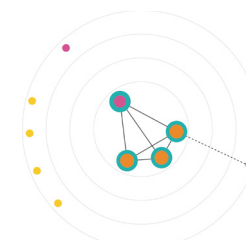
The evidence generated by Progresa also shaped policy making in other countries around the world that adapted the conditional cash transfer program. In 2009, the World

Bank published a volume on conditional cash transfers, counting around 30 programs similar to Progresa around the world. Informing policymaking was a two-way street for the program and its implementers, where the delegations the World Bank sent helped the Mexican team learn the strengths and weaknesses of the program through the questions and discussions with the international visitors.

Incentive Alignment Among Actors

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.

As researchers who had become policymakers, President Zedillo and the team of designers and implementers shared the vision of integrating rigorous evidence into policy and practice. Although they had all been researchers, they also had plenty of experience as public officials, and were well aware of the political landscape and operational realities. They did not produce original research to inform the program, but rather used the best available research to design the first pilot of the program, and then generated the evidence needed to evaluate its implementation and design the program at scale. No decision was made without gathering data and documenting it. Second, the idea of a robust external evaluation was of crucial importance. The program not only generated quantitative and qualitative information, but also made it public. Levy and Gómez the León were pragmatic about the quality of the information, especially the need for rigor and objectivity. This resulted in spending time and resources finding the best available external partners, as well as engraining evidence generation deep into the program's structural design. Finally, the commitment to ongoing learning within the design and implementation of Progresa can be seen throughout the decisions made and the organizational structure for operating the program. Learning from the evidence



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



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

generated was central to Progresa, providing an operational structure with the capacity to make decisions and incorporate emerging evidence so the program could have more impact.

The Role of Timing in the Incorporation of Evidence into Practice

The different and often discordant timeframes within which researchers, policymakers, and implementers operate often hobble efforts to coordinate, let alone collaborate, on evidence-informed approaches. Electoral cycles and political windows differ from NGO funding cycles and from academic publishing rhythms. Yet each actor is bound by the timeframes of her formal stakeholder group.

The three main actors behind the design of the program (Zedillo, Levy and Gómez de León) had a good understanding of the time horizons of research. Also, their public policy experience provided them with the knowledge of the “traditional” timeframe of public policy design and implementation both in Mexico and in global best practices. Due to this shared experience, they were able to go against the political pressure of quick timing to maximize public “splash” with constituents, sacrificing the launch of the program at the beginning Zedillo’s presidential tenure, and develop and test a solid program. Put differently, there is a clear and substantial



political trade-off in the integration of evidence as demonstrated by Progresa. From a traditional policymaking perspective, the rollout of Progresa was unusually slow. The program was officially launched two years into the Zedillo administration—and even then the program’s initial scale was modest by the standards of its predecessors. This was an administration under enormous political pressure to have a visible policy accomplishment in the wake of the 1995 financial crisis. The design team nonetheless chose a gradual, evidence-informed design and rollout of its flagship social program.

While Progresa ultimately became Mexico’s most successful social policy, it is important not to lose sight of the enormous trade-offs involved in its design.

Need to Devote Exclusive Time and Resources to Learn About and Operationalize Evidence

Few organizations provide incentives or carve out explicit time for managers to devote to learning about emerging evidence in their field of endeavor, and even fewer have staff explicitly devoted to learning about evidence and translating it into forms relevant for the organization. Even organizations with strong



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

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. Discovering, incorporating and translating evidence requires time, energy and funding.

President Zedillo, Santiago Levy and José Gómez de León were well positioned to become champions of evidence uptake. Progresa helps us understand five successful paths to guide evidence uptake. These paths are: 1) recognizing the crucial value of gathering, analyzing, and integrating emerging evidence, 2) commitment of resources to constant learning, 3) structural flexibility so a program can adapt to learning, 4) transparency, objectivity, and rigor in evidence generation, including through external partnerships, and 5) building decision structures for ongoing collaboration and negotiation among the key players. First, the designers of Progresa incorporated a long-term focus on demonstrable impacts for program beneficiaries, which translated to an understanding that gathering, analyzing, and evaluating its impact was not a luxury, but a necessity. Second, throughout the case

The designers of Progresa incorporated a long-term focus on demonstrable impacts for program beneficiaries, which translated to an understanding that gathering, analyzing, and evaluating its impact was not a luxury, but a necessity.

narrative we observe a constant commitment to ongoing learning, from designing better questionnaires, to obtaining better information about the needs of intended beneficiaries through the pilot, to constantly evaluating results to learn and modify the program and its theory of change based on emerging evidence. Third, the realization that, although there were costs and enormous

tradeoffs associated with the commitment to a robust evaluation, the benefits far outweighed these costs. Every evaluation was done by reputable external actors that ensured the objectivity, transparency and rigor of the program. This delayed the launch of the program for the Zedillo administration—much in need of an early policy success—but created a powerful political platform, built on legitimate evidence, that armored the program against opposition throughout its incorporation and allowed it to survive through four different presidential terms and two transitions between opposing political parties.

Finally, weekly discussions were held among the key players, in this case an inter-ministry group that met to discuss the operation and technical aspects of the program, which illustrates the strong commitment to fostering ongoing collaboration among the key players. Progresa integrated school attendance and visits to physicians, which meant the participation of other government agencies like the Ministry of Health and the Ministry of Education. Cooperation was crucial to the success of the program and fostering open dialogue and inclusion could make or break it.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Importance of Building Trust and Forging Relationships Among Stakeholders

The cross-stakeholder collaborations required for evidence-informed policies and practices are often difficult to initiate, develop, and sustain. Particularly when institutional incentives are lacking, personal trust, respect, and buy-in between individuals across stakeholder groups become critical to fostering the effective flow of evidence into practice.

Progresa's narrative highlights two very important relationships based on trust, without which success would have been impossible. The first component of forged relationships was the trust among Zedillo, Gómez de León, and Levy. President Zedillo appointed them based on the personal affinity and trust he had for them, as he had an academic background himself, but not explicitly because he was trying to staff the government with researchers. However, that affinity put the right people in the right places. Gómez de León and Levy trusted each other based on their shared goals, their common understanding of the overarching objective (demonstrable impact for beneficiaries), and the professional respect they had for each other. This trust was also infused into their teams.

Once all the design details were fleshed out, the inter-ministry group (with the Deputy Ministers of Finance, Education, and Health, and the Chair of CONAPO) was created to finalize the operational aspects of the program. This shows the commitment to ongoing collaboration, trust-building and negotiation among the key players from early on. A key element to enable the weekly discussions to develop the operating and technical aspects was President Zedillo's involvement to enable the coordination between ministries. Government stakeholders showed the will to collaborate and commit to the program as a result of the leadership, support, and pressure from Mr. Zedillo. Fostering inter-governmental coordination through strong leadership was key to the success of the program.

The Role of Funders | Shielding the Program

Funders have levers to support evidence-informed policy and practice. Progresa sheds light on a very interesting angle, using an international loan agreement to maintain the continuity of the program under its core operating rules based on rigorous evidence. The goal of the loan agreement was to shield the program from political cycles or temptations to manipulate it for electoral

purposes. The guardrails created by the IDB loan indeed protected the core of the program and strengthened its structural commitment to evidence integration.

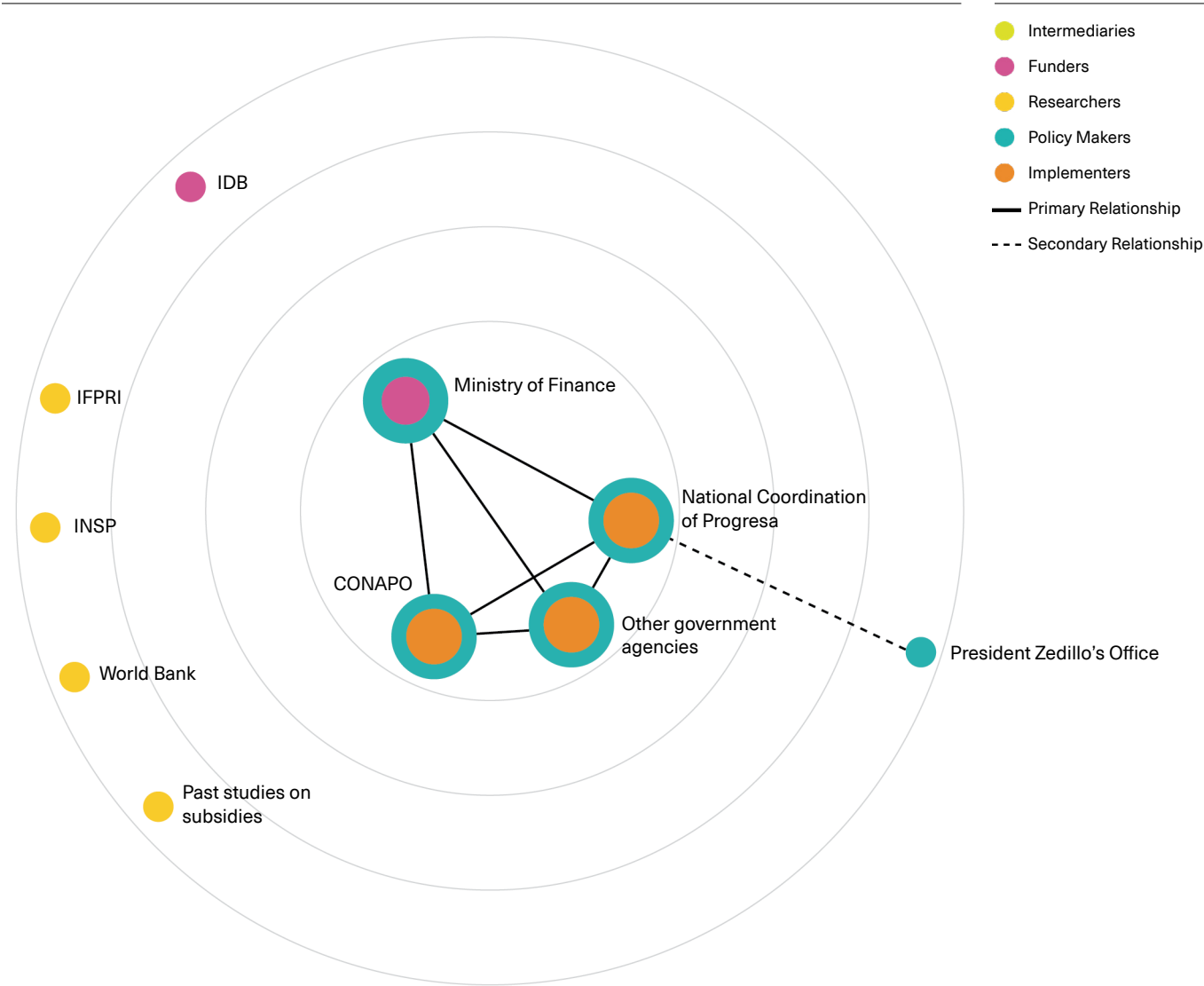
Conclusion

Progresa established a new standard for social programs in Mexico. Its insistence on generating and using rigorous evidence to inform its design and ongoing implementation challenged the prevailing political norms, in terms of both content of social programs (previously centered on unsuccessful price subsidies) and process (where programs tended to be subject to political manipulation). Progresa's evidence-driven approach led to the creation of an extremely successful program on multiple levels: Progresa has reached massive scale in Mexico; the concept of conditional cash transfer programs has spread to dozens of countries around the world; and the establishment of CONEVAL, to assess the effectiveness of all the country's social programs based on rigorous evaluation of evidence, has transformed the structure of how Mexico integrates evidence-based practice into public policy.



Stakeholder Map

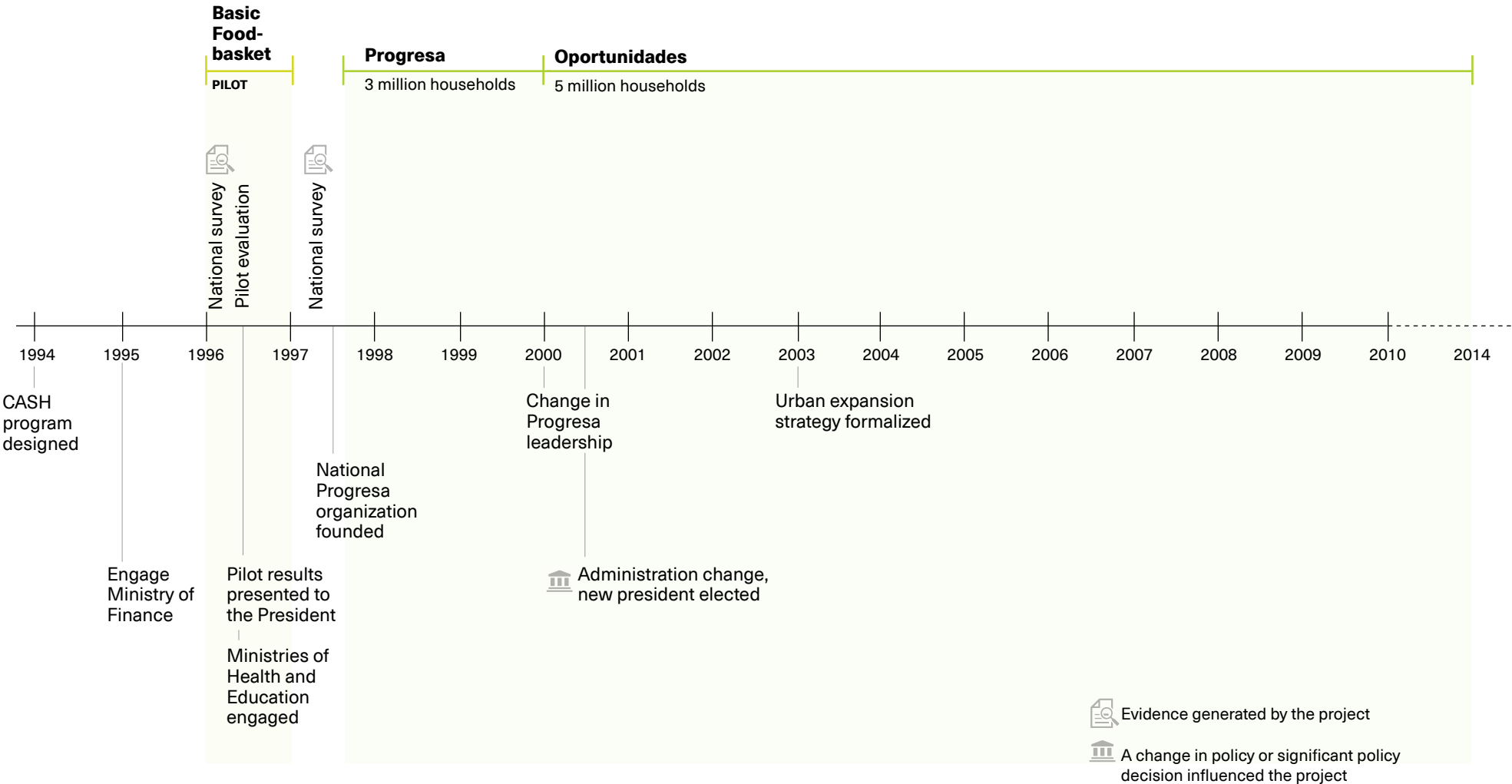
This stakeholder map is a visual representation of the major stakeholders involved with this project. The importance of each of the actors is defined by their relative size, and their proximity to the center of the project. Their role is defined by the color; multiple colors indicate multiple roles. 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.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Timeline

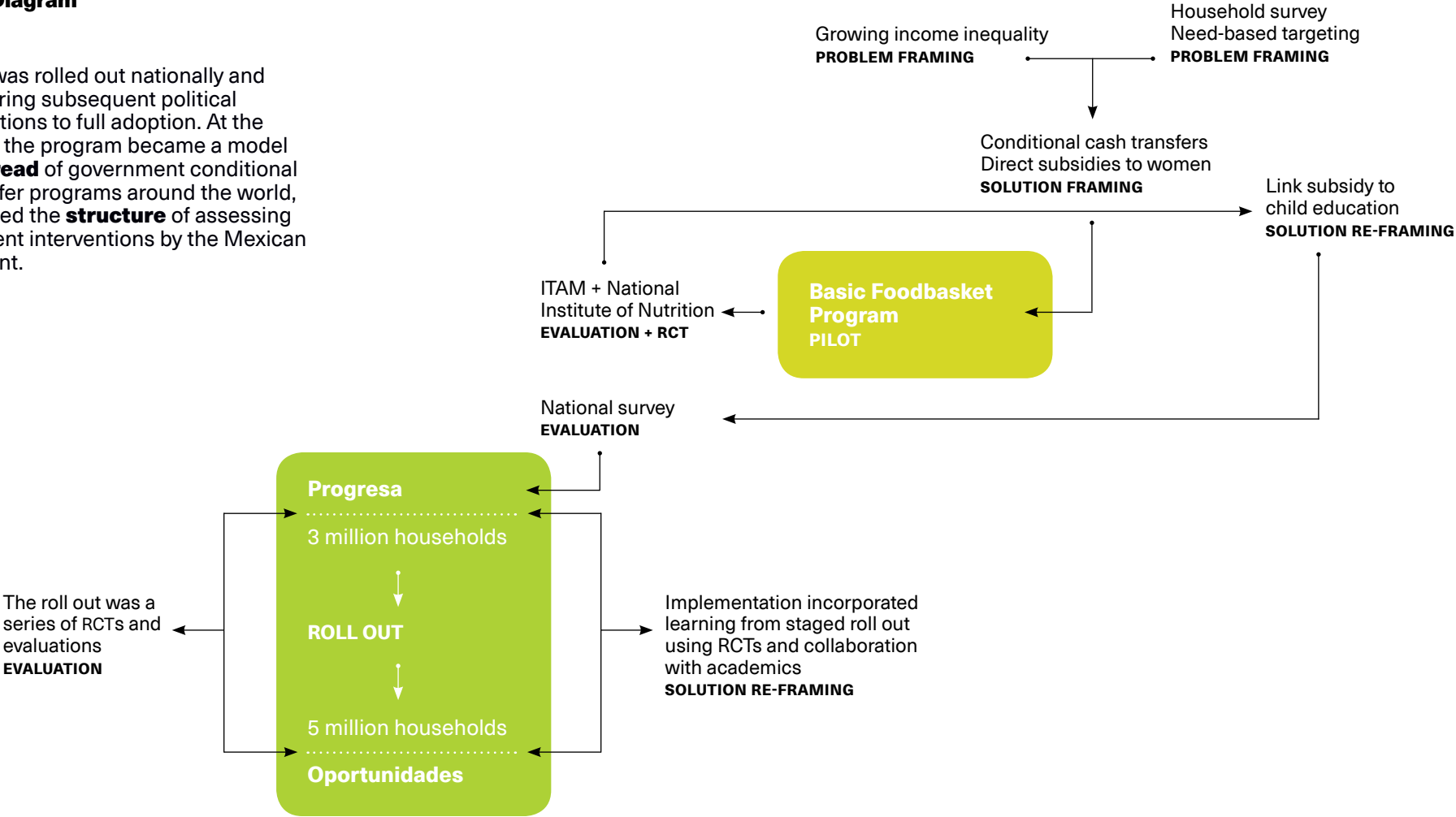




Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Process Diagram

Progresa was rolled out nationally and **scaled** during subsequent political administrations to full adoption. At the same time the program became a model for the **spread** of government conditional cash transfer programs around the world, and changed the **structure** of assessing development interventions by the Mexican Government.





**Evidence in
Practice**

Full Report

Print

Case Studies

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Page **160** of **220**

MEXICO | EMPLOYMENT

Programa Primer Empleo

A national employment generation program with good intentions and a desire for rapid impact was implemented too quickly to integrate existing and emerging evidence, creating insurmountable barriers,

By María del Mar Gutiérrez





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part I: **The PPE Story**

Programa Primer Empleo (PPE) launched during the first months of 2007 and was intended to create incentives for employers to generate new permanent jobs. It consisted of a government subsidy to reduce the costs of hiring, provided through a retroactive, partial reimbursement in the mandatory social security fees paid by employers who hire new workers.⁹⁸ The program was designed by the incoming Presidential administration of Felipe Calderón, funded through the federal budget, and implemented by the Mexican Social Security Institute (the Instituto Mexicano del Seguro Social, or IMSS). The program was voluntary, but contained guidelines and restrictions that defined the eligibility of employers and employees. The decree established that the program would start operations in March of 2007 and end on November 30, 2012. In the 2007 federal budget, the program was assigned \$3 billion Mexican pesos (\$274.5 million USD 2007 equivalent).⁹⁹

The story of Programa Primer Empleo illustrates how political and social pressures can lead policymakers to seek program results and visibility at a much faster speed and much larger scale than the available data—and the uncertainty surrounding initial assumptions—would otherwise dictate. The design, development, and management of Programa Primer Empleo also shows the importance of defining the problem to be addressed by a program in consultation with all the relevant stakeholders, which can only occur in a framework of trust among all the

actors involved. Also, the case highlights the barriers to evidence uptake. Finally, the program's implementation involved challenging choices and presents learning opportunities regarding all these topics.

The Design Process | Context: 2006 -2007

After a bitterly contested campaign season and election, Mr. Felipe Calderón was elected president by a narrow margin of 0.56 percent.¹⁰⁰ The election was followed by a long, controversial post-electoral process that created an environment of enormous

political pressure. Mr. Calderón, who ran on a platform of job creation (he was the “candidate of employment”), needed quick and visible results, so the transition team started working on an ambitious plan to encourage job creation, which later became Programa Primer Empleo. This process began during the transition period that followed the elections in July 2006 and ramped up after President Calderon's inauguration on December 1, 2006. The program's creation was officially announced in January 2007, less than two months after the new administration began. Programa Primer Empleo became a flagship program of the administration, addressing the “more and better employment” campaign pledge.

⁹⁸ Objective published in the official decree creating the program. Federal Official Gazette. January 23, 2007

⁹⁹ Average 2007 exchange rate \$1.00 USD = \$10.9282 MXN. [Available here.](#)

¹⁰⁰ “Official Results, Mexico 2006 Presidential Elections,” Journal of Democracy. [Access here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The story of Programa Primer Empleo illustrates how political and social pressures can lead policymakers to seek program results and visibility at a much faster speed and much larger scale than the available data would otherwise dictate.

Evidence Used

In 2003, a theoretical economics paper titled “A Model of the Mexican Labor Market with and without Social Security, 2003”, modeled the elasticity of employer fees paid to IMSS and their effect on the uninsured (informal) labor market. The study argued that a 6% reduction in the fees would translate to as many as 300,000 jobs in the formal, private, non-agricultural sector, mostly through the formalization of informal jobs.¹⁰¹ The authors sought to generate knowledge, “especially because of the importance of the non-affiliated labor sector in Mexico and in many other countries and the lack of

macroeconomic models to finance social security around the world that include labor market specifications in which a large group of workers are not covered”.¹⁰² The paper specifically focused on the transition of existing jobs between the formal and informal labor markets (not on the creation of new jobs), but the authors nonetheless believed it provided a useful point of reference for labor market dynamics and elasticities in Mexico. The direct implications of the paper for the specific, job creation objectives of PPE were much less clear.

Targeting + Objective

In line with the theoretical paper, the President’s team believed that IMSS fees were high enough to constitute a real constraint on the creation of new, formal jobs.¹⁰³ Programa Primer Empleo thus sought to encourage formal job creation by lowering this cost to employers. The official objective of the program, stated in its creation decree was: “to support employers to generate new permanent jobs, through granting a subsidy that would be applied to the fees paid by the employers for hiring additional new workers and registering them at IMSS.”¹⁰⁴ The program’s founding document does not explicitly address a target population, but the

101 The model takes a redistribution of work without coverage to have coverage in the supply, but without more hours of work being offered or more people who were not working being incorporated into the labor market. Since that is not considered, the 300,000 would not be new jobs, the jobs would be going from the informal sector to the formal one.

102 Nora Garro Bordonaro, Jorge Meléndez Barrón, Eduardo Rodríguez- Oreggia, Un modelo del mercado laboral mexicano con trabajo con o sin seguro social (IMSS) (Mexico: Universidad Iberoamericana, A.C, 2005), 5-7

103 Design Evaluation. CONEVAL, page 9. October, 2007.

104 Objective published in the official decree creating the program. Federal Official Gazette. January 23, 2007



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

name of the program and its eligibility requirements indicate that it was targeted specifically for recent graduates and people entering the formal job market for the first time. Also, the National Development Plan for 2007-2012 mentioned the program as part of the strategy to create new jobs for youth entering the formal work force.

Determining the Rules of Operation

The program design began as a simple and straightforward initiative, but policymakers soon worried about two potential secondary consequences. One was the fiscal impact, as any subsidy implies the loss of income or an expense for the government. The second had to do with the duration of the subsidy and its potential to generate long-term market distortions. These two factors, among several others, disproportionally influenced the design process. As described below, each program design decision was driven by commendable intentions and solid reasoning, but some of them nevertheless—and unintentionally—turned into barriers and restrictions that negatively affected the performance of the program.

To be eligible for the program, a candidate had to be entered into the IMSS registry as a permanent employee and had to meet the following criteria:

- *The employee had never been previously registered at the Mexican Social Security Institute (IMSS).* Programa Primer Empleo conditioned eligibility to target young people who were joining the workforce for the first time or people who had worked in the informal sector but could transition into the formal sector. This rule also sought to avoid a replacement effect.

- *An employee's enrollment in IMSS had to be in addition to the maximum permanent number of workers already registered by the hiring company.* The program aimed to create new employment, so this rule granted eligibility only to companies that increased their net workforce.
- *The employer could request the payment of the subsidy only after the tenth month of continuous full employment of the newly registered employee.* Through this rule, the program aimed to encourage full-time, long-term jobs and avoid promoting temporary jobs.
- *To participate, the company was required not to have fiscal debts with IMSS or the Mexican government.* Since it was a public program, beneficiary companies had to be up to date with their fiscal obligations.
- *The maximum duration of the subsidy would be twelve months.* The subsidy would vary with a worker's base salary, ranging from 10% to 100% of the IMSS fee. The duration was chosen in order to provide a large enough incentive for companies to hire new workers, while also limiting any harmful, long-term effects on public finances and avoiding the creation of perverse incentives, such as market distortions through the gradual incorporation of the expectation of subsidies into an organization's business model. The intention was that, a) once organizations trained and then observed the quality of a new employee for a year, they would choose to retain him or her; b) once workers were brought to the formal labor market, they would prefer to remain; and c) the subsidies would be short-term but would trigger a long-term impact.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

With these eligibility requirements and historical data, IMSS officials worked with Mexican academics to estimate, through simulations based on the projected behavior of the labor market, various scenarios that had different outcomes. Another consideration was the 800,000 employers registered at IMSS, so up to that many employers could have been eligible for the program. It remained to be seen how many of them would actually seek the subsidies in practice. For a number of reasons explored below, however, the program never lived up to its expected results.

Trust

As mentioned above, what began as a straightforward initiative with good intentions soon concerned well-intentioned policymakers with unwanted or unanticipated impacts. Each identified concern was addressed through adding complexity and rigidity to the eligibility requirements, seeking to anticipate and avoid “gaming” of the system. As a result, policymakers unwittingly developed a program built, at its core, on distrust of the same employers that had been planned as vehicles to reach the intended end-beneficiaries.

This mistrust by government officials the private sector was, not surprisingly, reciprocated. The private sector often

What began as a straightforward initiative with good intentions soon concerned well-intentioned policymakers with unwanted or unanticipated impacts.

does not trust government programs because it believes their scope is rarely clear or transparently explained, much less communicated adequately. Accordingly, Programa Primer Empleo was received with skepticism. First, the program was fraught with complicated requirements and commitments. Second, the monetary incentive had strings attached, with added costs and risks, like the need to be up-to-date on all taxes and government fees. Employers had three main concerns: (1) the risk of being subject to additional audits, (2) the costs of artificially long contracts with workers, and (3) the uncertainty and administrative burden of claiming the subsidies, ex-post, from IMSS.

By law, IMSS is one of three Mexican institutions with legal auditing and sanctioning capacities. Companies thus feared that the program would provide an opportunity for IMSS to audit them. In addition, Mexican labor law was known for its rigidity. The program granted a partial, temporary subsidy, while companies absorbed the long-term financial and legal risks of recruiting individuals with no experience in the formal labor market. According to the program's rules, employers were only allowed to claim the subsidy once the new employee had completed ten months of employment. As the president of Durango's chapter of the Mexican Employers' Association (COPARMEX)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

mentioned in a newspaper article, “when a new employee without experience enters the workforce of any business, he or she requires training time; however, in this process there is always desertion, approximately in the fifth or sixth month”¹⁰⁵ and under this program, the employer did not reap any benefit unless the new employee stayed at least ten months.

Including the Private Sector in the Conversation

Government officials involved during the design phase of the program had felt no need to consult with the private sector, because they believed PPE was a simple, straightforward idea. Time was also limited, since they wanted the program to start very early in the Calderón administration to make a strong political statement. The design process entailed robust discussions among a group of highly skilled and experienced policymakers, so designers were thoroughly convinced by the logic behind the program. At the time, policymakers were convinced that they had listened to outside voices, but the other stakeholders experienced the program as a unilateral, executive decision by the Federal Government.

When asked, key representatives from the private sector said they were not invited to be part of the design process and only learned about the program after it was

Government officials involved during the design phase of the program had felt no need to consult with the private sector, because they believed PPE was a simple, straightforward idea.

launched. Many mentioned that the Federal Government had been the only one involved in the design phase, and not seeking the business perspective had been one of the reasons why the program did not succeed. Private sector representatives believed that PPE was far too complex and would have required substantial modifications to become attractive. One of them explained that COPARMEX had attempted to intervene and influence the program design, but since the organization played no official role within the IMSS Technical Council,¹⁰⁶ its engagement had been limited to indirect political action and marketing efforts.

Program Implementation + Monitoring

IMSS was responsible for operating the program, even though it lacked the legal mandate to promote employment or provide fiscal incentives. The rationale was that, as the institution tasked with keeping track of the formal labor market and running the country’s employment database, it held the most timely and accurate employment information, so it was best positioned to oversee PPE. It also made administrative sense, as PPE operated as a government subsidy to employers who would be reimbursed for fees already paid to IMSS. At the same time, IMSS could not finance the program, as it is a tripartite body, financed by employer fees, worker contributions, and transfers from the Federal Government. These resources are used to cover workers’ pensions, provide them with medical services, and expand the medical infrastructure, and by law cannot be used for other types of programs or activities. As a result, the Ministry of Finance (SHCP) provided the necessary resources and became a central actor in the program.

¹⁰⁵ Citlalli Zoé Sánchez, *Necesario, modificar Primer Empleo: IP* Published: Wednesday August 8, 2007. *El Siglo de Durango*. Accessed August 31, 2017. [Access here](#).

¹⁰⁶ IMSS Technical Council is integrated by representatives of the employer, worker and government sector. In terms of IMSS Internal Regulation, the Technical Council is the legal representative and manager.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Within IMSS, two units were responsible for the program. Formally, the Inspection and Collection Unit was designated as responsible for the program, but the Planning and Evaluation Coordination department¹⁰⁷ was in charge of day-to-day operations. At the same time, PPE had no formal organizational structure, only individuals assigned to it from the two aforementioned units.¹⁰⁸ The program had national coverage, with no emphasis on specific geographical areas, so it was available throughout Mexico to companies that met the eligibility requirements. The program wanted to avoid generating additional administrative burdens for employers registering for PPE, so it was designed to run on an online platform. The system was developed so that all the registration and verification was done through the online system, avoiding extra administrative costs for the employers and reducing the program's operating costs, which would hopefully generate greater incentives for employers to register for the program, and make the benefits available to as many workers as possible in an efficient and effective way.

Given its dependence on the voluntary registration of employers, the program required effective communication to all eligible companies. Communication efforts in 2007 included printed materials, a phone

campaign, a media campaign, direct mailings, and dissemination throughout the IMSS network (see appendix 1 for details). While there was no systematic effort to measure the effectiveness of these communication campaigns, at the time PPE's implementers believed the efforts had been sufficient. They acknowledged later that it might have been good to have maintained or intensified the level of awareness-building of the program.

PPE was funded by federal resources and, by law, was subject to a series of rules and controls, including audits by the Auditor General of the Federation and a formal Design Evaluation by a Federal entity. In addition, given its tradition of careful data collection and analysis, IMSS systematically collected and analyzed program results. Despite the data and the evaluations of the program's operations, there was no systematic analysis of its intended impacts or of the reasons behind PPE's lack of acceptance among eligible companies. There were several structural reasons why this was the case. First, IMSS does not have the legal mandate to carry out impact evaluations. Second, federal money is usually earmarked with very specific restrictions on its use, which was the case for PPE. Finally, the team for PPE was clearly focused on ensuring that implementation happened ethically and with strict observance of the program's rules.



¹⁰⁷ The Planning and Evaluation Coordination is a separate administrative unit which is not part of the Inspection and Collection Unit

¹⁰⁸ Design Evaluation. CONEVAL, page 58. October, 2007.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Design Evaluation

A few years earlier, in 2004, new legislation created the National Monitoring and Evaluation System and the National Council for Social Development Policy Evaluation (CONEVAL)—the latter with the mandate to measure poverty and assess social development policy.¹⁰⁹ Following this mandate, CONEVAL commissioned a design evaluation of Programa Primer Empleo, focusing on archival data from the Inspection and Collection Unit at IMSS. The evaluation also included a review of existing literature and of other relevant international experiences. It identified several positive features of the program, as well as a series of structural design flaws. On the positive side, the evaluation found that “the program had minimal operating costs, due to the highly-automated process which resulted in few operational field activities.”¹¹⁰

The evaluation’s final report, however, also identified critical shortcomings:

“Programa Primer Empleo performed well below expected results due to an inadequate program design, resting on the assumption that a subsidy to the social security fee was a determining factor to stimulate demand of formal and permanent jobs for new entry-level workers who traditionally did not access permanent hiring schemes despite the inertial dynamics of the labor market.”¹¹¹

“The weakness showed by the operation of Programa Primer Empleo was in line with international experience, which conclusively points out that strategies that use a salary subsidy scheme as the only component are not effective in order to integrate a vulnerable population into permanent and formal jobs. The programs that promote a combined strategy that include subsidies to the employer, job training for vulnerable groups, and support and information services while searching for employment are more likely to succeed. Also, for combined strategies to be more likely to succeed, the training received by workers must be linked to productive needs and the benefits of the program should be targeted to the vulnerable population. Even if the subsidy provided by Programa Primer Empleo had been higher, the Program would have not worked properly because the international evidence is conclusive that such programs, operating solely through a subsidy to the employer, have not worked.”¹¹²

“There are other factors that explain the low level of permanent employment in the formal sector, such as: 1) the rigidity of labor legislation; 2) the low valuation of workers regarding social security, especially young people; 3) the high information costs faced by a vulnerable segment of the population; 4) incentives for circumvention and avoidance of employer contributions; and 5) the low educational levels of the vulnerable segments of the population.”¹¹³

¹⁰⁹ Created by the Social Development Law, see: *¿Quiénes somos? – Creación del CONEVAL*. CONEVAL [Access here](#).

¹¹⁰ Design Evaluation. CONEVAL, pages 77-79. October, 2007.

¹¹¹ *Ibid*, 6-7.

¹¹² *Ibid*, 20.

¹¹³ *Ibid*, 37.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The evaluation delves into the structural impediments to youth employment because of their apparent absence in the design of PPE, which contributed to the program's lackluster results.¹¹⁴ Because of the PPE's structural shortcomings, the evaluation issued the recommendation to terminate the program due to:

1) "The unviability of the Program under its current design, and 2) the relevance of modifying the PPE towards a comprehensive strategy when these efforts already occur at the Ministry of Labor and Social Welfare."¹¹⁵

Should the administration retain the program, the evaluation proposed several modifications to address some of the program's flaws, as seen in Table 1.

Table 1. Proposed Modifications to PPE (2007) ¹¹⁶	
ISSUE	PROPOSED RECOMMENDATION
Potential and target populations are not well identified or adequately quantified by the Program.	The potential population should be defined as vulnerable groups that, in the absence of the Program, would not obtain permanent employment in the formal sector. The target population should be workers with low educational levels, of productive age, with little or no work experience, from low socio-economic strata and with a higher propensity to work in the informal sector, who would not otherwise join the formal job market.
The purpose of Programa Primer Empleo is to encourage the demand for vulnerable workers beyond the normal inertia of the labor market. This is not fulfilled, so the Program is subsidizing workers who were already likely to enter the formal labor force.	
The Program does not allow portability. An employee who changes jobs after registration in the program but within the eligibility period automatically loses eligibility.	Allow portability. The program should not generate additional rigidities in the labor market. Workers, especially those with lower incomes and educational levels, have greater mobility and turnover, so the program should allow them to change jobs while retaining eligibility for the Program.
The subsidy component of the worker-employer fee is not adequate, because it is based on a theoretical assumption about the dynamics of the informal labor market, not about new job creation. Furthermore, unidimensional programs, such as the provision of economic incentives to employers, have proven insufficient in previous international experiences and in the evidence of PPE itself.	This issue forces us to rethink the pertinence of keeping Programa Primer Empleo in operation due to the infeasibility of the Program under its current design, and the relevance of modifying PPE towards a multi-dimensional, comprehensive strategy that includes training and job placement (the Ministry of Labor and Social Welfare already has programs that seek to follow these principles).
OTHER ISSUES AND RECOMMENDATIONS	
The information about employers is not enough to know their needs thoroughly.	It would be important to carry out a survey to understand the employers' perspective, opinions, and suggestions. This would provide knowledge on the perceptions companies have on youth employment, training, and employment demand.
A survey of workers is also required to know their perceptions on employment, education, and job offers.	This study would serve to understand workers' disposition on types of employment and training programs.

¹¹⁴ Ibid.

¹¹⁵ Ibid, 86.

¹¹⁶ Ibid, 77-79, 86.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Regarding the paper “A Model of the Mexican Labor Market with and without Social Security, 2003,” which served as the primary evidence and inspiration for the program, the evaluation states that the study was a useful theoretical analysis of labor market dynamics that, nevertheless, had no empirical validation. In fact, the early results of PPE demonstrated that the reduction of employer’s fees was not enough to increase the number of new, permanent jobs in the formal sector above the normal market inertia.¹¹⁷

Finally, the evaluation concluded that previous international experiences had not been adequately incorporated. In particular, the evaluation mentions two types of international experiences focused on the employment of vulnerable workers that seemed relevant to PPE:

- 1) Programs that provide a subsidy to the employer - providing economic support to hire the worker. The two programs mentioned in the evaluation were: The New Jobs Tax Credit¹¹⁸ and the Targeted Jobs Tax Credit.¹¹⁹ Research on both programs concluded that neither had been effective.
- 2) Programs that promote a multidimensional strategy providing support for employers (subsidy) and employees (training and job search support). The two programs mentioned in the evaluation were: The Youth Incentive Entitlement Pilot Project¹²⁰ and the AFDC Homemaker-Home Health Aide Demonstrations,¹²¹ with less conclusive evidence.

The CONEVAL evaluation thus concluded that the evidence on unidimensional programs, like PPE, was unequivocal: they are not sufficient to generate new jobs or insert vulnerable groups into formal and permanent jobs.

Civil Service Middle Management, the Federal Government, and the Private Sector

The team charged with the execution of PPE had been appointed to IMSS during the presidential transition, but many had had long careers as civil servants. Because of their experience in the public sector, they were skeptical of the program and the assumptions it was built upon. Regardless, it was their task, and not that of program designers, to ensure that PPE was implemented in strict accordance with its operational rules. As program results trickled in (IMSS is a rigorous collector of labor market data), the fears of the implementation team were confirmed: numbers were far below expectations and with no signs of promise.

The CONEVAL evaluation was a turning point, because it not only corroborated the internal analysis that IMSS had performed on the operation and evolution of the program, but also was an objective report from an *external* actor, indicating that the program was not having the desired results because of structural flaws in its design (and not because of ineffective implementation). Implementers agreed with the CONEVAL assessment, but also felt that employers probably feared IMSS as a fiscal control entity: The program stated that companies had to be up to date with all their fiscal obligations with the Mexican government and open to random, unannounced audits from IMSS.

¹¹⁷ *Ibid*, 13.

¹¹⁸ *New Jobs Tax Credit: USA – 1977.*

¹¹⁹ *Targeted Jobs Tax Credit: USA – 1978.*

¹²⁰ *Youth Incentive Entitlement Pilot Project: USA – 1978 to 1980.*

¹²¹ *AFDC Homemaker-Home Health Aide Demonstrations: USA—1983 to 1986*



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

As much as implementers agreed with them, most of the critical recommendations provided by CONEVAL fell beyond the institutional mandate and jurisdiction of IMSS. Implementers were legally constrained by their official roles and by program rules, so even as they attempted to integrate the new evidence, they were only able to tweak the program at the margins. This was not only insufficient, but also made it appear as if IMSS had resisted or ignored what the evidence was suggesting was needed. IMSS implementers had to keep the program running according to the federal budget and the program rules. Addressing the main recommendations, in contrast, would require involvement from higher levels in the federal government and a fundamental shift in strategy, far above the mandate or purview of IMSS.

In time, CONEVAL understood that those who operated the program were not the ones who made decisions on whether to terminate it or even on its structural design. To the evaluator's surprise, those responsible for implementing the program were transparent, open, and receptive. It was clear that there were good intentions behind all the decisions, even those that kept the core of the program design intact. As one of the evaluators noted: "I was surprised by that, but at the same time I was not. It is difficult for someone within a

government to have an evaluation and say, 'Ah, I will totally change it,' especially when it's a flagship program."

After CONEVAL's evaluation and in line with the implementers' intuition, the program underwent two major modifications to ease the requirements for companies and beneficiaries. In general terms, starting in December 2007, the rules were slightly simplified:

2007 | 2008 ¹²²

- Targeted population expanded: individuals who had not been employed for more than nine consecutive months registered with one employer as a permanent worker remained became eligible.
- The waiting period to claim the subsidy was shortened from ten to four months after the employee's registration.
- The registration period for employers and workers was extended to August 31, 2011 from February 28, 2011.
- The condition of being up to date with all fiscal obligations was made more flexible: the requirement for timely completion of fiscal obligations was kept, but only at the moment of claiming the subsidy and not at the time of registration.

After the first round of modifications was put into place, a document developed by the Mexican Employers' Association (COPARMEX) noted several program shortcomings, including: so many requirements that most companies would not consider it as an option; restricting the subsidy to hiring of people who had never had a formal job, and so precluding companies from hiring staff with previous experience; not taking into account the reasons behind any fiscal liabilities companies might have; and having to maintain a steady workforce, with no margin for industry dynamics or natural turnover.

The program continued to perform below its expected results, so in September 2011 the Federal Government consulted the private sector on how to restructure and relaunch PPE.¹²³ COPARMEX proposed seven modifications: (1) extend the duration of the program, (2) eliminate the requirement for full fiscal compliance, (3) through a decree, confirm that there is no need to be up to date

¹²² ASF. *Informe del Resultado de la Fiscalización Superior de la Cuenta Pública 2012. Programa Primer Empleo performance Audit (12-1-00GYR-07-0379) DS-095 page, 2. 2012 and Comisión Representativa ante Organismos de Seguridad Social: Programa Primer Empleo. December, 2011.*

¹²³ *El Empresario. Access here.*



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

with Federal Government liabilities, (4) reduce the period to begin collecting the subsidy to two months, (5) IMSS should deposit 100% of the subsidy in the month following an employer's request, (6) the option to choose between receiving the subsidy in a bank account or a 10-day credit note, and (7) design a new communications campaign, as the first set of modifications were not well known by target companies and beneficiaries.

On November 14, 2011, the Federal Government issued a new set of guidelines for Programa Primer Empleo, which incorporated some of these recommendations:

2011¹²⁴

- The employer registration period was extended from August 2011 to September 2012.
- Worker eligibility was made more flexible, allowing people with limited work experience (and not only individuals with no experience) to register for the program.
- The waiting period to receive the subsidy payment was shortened from four to three months of full employment of the registered employee.



As noted before, since the beginning of the Calderón administration, the federal government had intended Programa Primer Empleo to make a big splash. When the program was launched, the signing of the creation decree was hosted by Mr. Calderón himself with the presence of union leaders, business leaders, legislators, and the media.

During his speech, he highlighted that with this program, his government was fulfilling one of the most important commitments

made to the Mexican people during his campaign. That same day, Mr. Calderón inaugurated the National Chamber of Industry's (CANACINTRA) 2007 National Industrial Convention where he announced the creation of Programa Primer Empleo.

¹²⁴ ASF. *Informe del Resultado de la Fiscalización Superior de la Cuenta Pública 2012. Programa Primer Empleo performance Audit (12-1-00GYR-07-0379) DS-095 page, 2. 2012 and Comisión Representativa ante Organismos de Seguridad Social: Programa Primer Empleo. December, 2011.*



By 2009, the President had stopped talking about the program and during the beginning of 2010 the director at IMSS said during an interview for a newspaper that “Programa Primer Empleo was not in the institute’s agenda as a priority item.”¹²⁵ After its first year, President Calderón’s administration became absorbed by the security agenda and PPE lost prominence. PPE was maintained because canceling it would have given the government negative attention, especially as people were critical about the focus on security at the expense of economic issues. By 2008 the program’s resources were redirected and it only kept enough funds to meet the commitments with the registered companies.

Results

Programa Primer Empleo officially came to an end with the Calderon administration, in November of 2012. At this point, the program had met 13%¹²⁶ of the employee registration target and 3.2%¹²⁷ of the employers’ enrollment target. The results were far below the expectations which targeted a goal of 738,133 employees registered between 2007 and 2012¹²⁸ and, in the eyes of an evidence producer, “the results were not aligned with the theory of change,” the magnitude of the expected results was vastly lower than projected.

Figure 1. Programa Primer Empleo Results

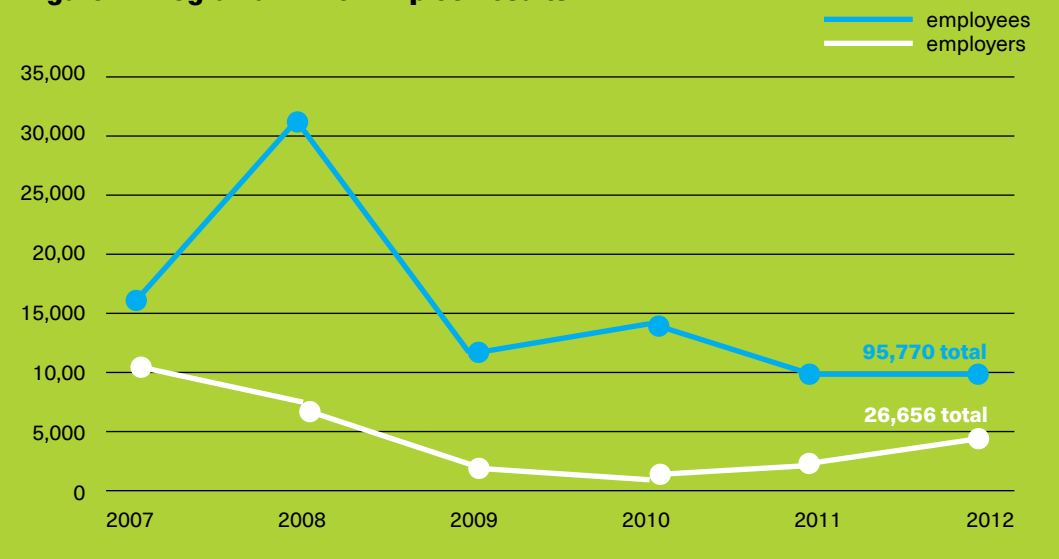


Figure 1 shows the evolution of the program throughout its history. As shown, there was an interesting spike in the number of registered employees in 2008, the same year the first major modifications were implemented to make the program more flexible. According to one of the implementers, the spike and the evolution of the statistics suggests that the first set of modifications had a positive impact, but it was not

¹²⁵ Martínez, María del Pilar. “Programa Primer Empleo, al archivo muerto del IMSS” *El Economista*. Published March 3, 2010. [Available here](#).

¹²⁶ ASF. *Informe del Resultado de la Fiscalización Superior de la Cuenta Pública 2012. Programa Primer Empleo performance Audit (12-1-00GYR-07-0379) DS-095 page, 5.* 2012

¹²⁷ *Ibid.*

¹²⁸ *Ibid.*



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

sustained for all the reasons described above. In addition—and in reaction to the 2008 financial crisis—the Mexican economy and labor market contracted in 2008, further affecting PPE results. The incentives provided by Programa Primer Empleo could have helped soften the impact of the recession, but once again certain design features interacted with emerging realities in unexpected ways. The subsidy was contingent on increasing the historic maximum workforce registered by a given employer, but few companies were able to increase absolute employment in a context of low growth. So even if companies hired new employees during those difficult times, they were not eligible for PPE if their entire workforce remained below their historic maximum. While the logic behind that rule was to encourage the creation of new jobs and prevent employee replacement, it actually played against the program's potential to mitigate the economic crisis. It was yet another way in which the program's design process and rigid structure hindered its own potential.

While the logic behind that rule was to encourage the creation of new jobs and prevent employee replacement, it actually played against the program's potential to mitigate the economic crisis.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part II: **Evidence in Practice : Key Themes + Insights**

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

The Role of Timing in the Incorporation of Evidence into Practice

The different and often discordant timeframes within which researchers, policymakers, and implementers operate often hobble efforts to coordinate, let alone collaborate, on evidence-informed approaches. Electoral cycles and political windows differ from NGO funding cycles and from academic publishing rhythms. Yet each actor is bound by the timeframes of her formal stakeholder group.

Programa Primer Empleo illuminates how constraints, especially campaign promises and post-electoral pressures, play a major role in facilitating—or constraining—the incorporation of evidence into policy and practice. When designing and implementing a major government program, policy makers have incentives to start new programs and fulfill campaign promises at the beginning of their political cycle.

Within the five-month transition period before the Calderon administration began, Programa Primer Empleo was designed in order to be launched at the beginning of the administration, as a fulfillment of one of President Calderon's most important campaign promises, in an environment of social fragility and political pressure. This situation illustrates the struggle between following a truly evidence-informed design process (with sufficient time to generate and/or consider evidence, implement pilot programs, and modify the design accordingly) and the exigencies of a public policy implemented with a fast-paced timeline in order to generate a quick and big “splash”; in the case of PPE, this approach affected not only the design and launch of the program but also its ongoing implementation.

Another key lesson that can be drawn from PPE is the importance of designing public policies that retain flexibility for learning. This naturally has to be balanced with the need

for strict operating and accountability rules in the use of public funds. But as other cases have shown, operating rules can include a mandate for adjustments based on learning. They can also include a structure that allows all relevant stakeholders to remain engaged throughout the design and implementation process, to promote the integration of all relevant evidence. In the case of PPE, the sequential (and disaggregated) role played by different stakeholders meant that the program missed critical perspectives in its design and was not adaptive to useful insights about its performance.

For Programa Primer Empleo, program designers did not have much evidence to support the assumptions behind the core of the program, but the pressures to start at a major scale and have a big splash pressed the program forward. Alternatively, the launch of the program could have been seen as an opportunity to pilot the intervention and test all critical hypotheses, in order to gain insight on what was needed to achieve a larger impact.

As time progressed, and especially towards the second set of modifications, the time constraint of the program's duration and with Calderon's tenure coming to an end, any major modifications were difficult to implement, due to: 1) the structure and rules of the program



that gave little flexibility and 2) the incentive politicians have to garner credit from their initiatives. Although the program did not have the expected results, politically it was very difficult to cut the program. Policymakers were thus subject to a series of pressures and restrictions that limited their ability to integrate the results of a potentially helpful evaluation. PPE's implementers literally could not change any structural features of the program because of its rules, which lacked the mandate to collect or adapt to evidence. The political environment made it difficult for those higher in the administration to intervene. Primer Empleo helps us understand the many factors leading to the realization that evaluations do not necessarily change public policy or program design, especially when the programs are important to the political strategy.

Policymakers were thus subject to a series of pressures and restrictions that limited their ability to integrate the results of a potentially helpful evaluation.

Evidence Definition, Creation + Use

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.

Programa Primer Empleo's design was based on a theoretical economic model developed a few years earlier. The model identified potentially critical leverage points in stimulating employment, but did not consider many underlying assumptions or identify

conditions that might occur in the field that would be conducive or resistant to its proving successful. Further, the program's rules and structures of operation, while designed with good intentions, proved to be unappealing to the target audience of employers. While the rationale behind them was correct in conceptual terms, the complexities of the field proved them to be inaccurate.

PPE illustrates many of the risks that arise in well-intentioned public policies when a comprehensive, evidence-informed approach is not followed in its design and roll out. Rather, PPE reflects the political imperative to launch full-scale, apparently “watertight” programs. PPE shows the importance of clearly identifying (a) the assumptions underlying an initiative, (b) the types of



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

evidence that would best shed light on those assumptions, (c) the actors who could best provide such evidence, and (d) how the initiative could adapt to such new evidence. This is especially true with programs that seek to modify behavior at scale, where inaccurate assumptions about actors' incentives can lead to unintended effects and a drain on resources.

While PPE at its core had the ambition and the potential to be a flagship program for the Calderon administration, the way it was designed and executed essentially turned it into a single (and expensive) test of a larger, unexamined hypothesis—which proved to be false. There were lost opportunities to engage key stakeholders as genuine partners, including potential employers, academic institutions, and think tanks in the design process, and implementation staff in using the evaluations' recommendations to strengthen the program. Its rigid design left committed, well-informed implementers unable to act upon emerging evidence. The narrative of the case highlights how government officials often lack the mandate, resources, and time to make use of the evidence generated by rigorous evaluations that can lead to program improvements.

Need to Devote Exclusive Time and Resources to Learn About and Operationalize Evidence

Few organizations provide incentives or carve out explicit time for managers to devote to learning about emerging evidence in their field of endeavor, and even fewer have staff explicitly devoted to learning about evidence and translating it into forms relevant for the organization. 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. Discovering, incorporating and translating evidence requires time, energy and funding.

As described in the case study narrative, the evaluation found that Programa Primer Empleo did not use all of the international evidence available at the time, such as the New Jobs Tax Credit and the Targeted Jobs Tax Credit from the U.S. This emphasizes the importance of not only explicitly considering different types of evidence upfront, but also factoring in the resources, bandwidth, research capacity, and time needed to ensure that relevant, timely and useful information is





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

made readily available to program designers and implementers. The critical analyses from the periodic evaluations of the program were also not fully taken into account. For the Primer Empleo story, considering all these factors could have helped develop a good intention to its fullest potential without investing valuable resources in a weakly conceived intervention—and ultimately to the discrediting of a flagship program.

Importance of Building Trust and Forging Relationships Among Stakeholders

The cross-stakeholder collaborations required for evidence-informed policies and practices are often difficult to initiate, develop, and sustain. Particularly when institutional incentives are lacking, personal trust, respect, and buy-in between individuals across stakeholder groups become critical to fostering the effective flow of evidence into practice.

Programa Primer Empleo illustrates the importance of stakeholder engagement to make sure policymakers, intended partners, and beneficiaries have the same understanding of the problem and potential solutions. The program's objective focused mainly on supporting employers to create permanent formal jobs, yet employers did not participate in the design of the program. Not having agreement on what would be

Another key lesson that can be drawn from PPE is the importance of designing public policies that retain flexibility for learning.

appropriate incentives or targets created confusion among the relevant stakeholders, each having a different understanding of what the program aimed to address and how to achieve its goals. PPE illustrates the importance of effective communication between key stakeholders, both upfront and ongoing, as fundamental for the successful design and implementation of a program. The case study research suggests that lack of trust blinded both sides, where policymakers complicated the program's operational rules (e.g., by requiring a company not to have fiscal debts with IMSS and by creating tight restrictions on employee eligibility) out of concern that employers could bend the rules and abuse the program. In turn, employers doubted the appropriateness of, and felt

threatened by, the rigidity of the rules—a natural reaction to rules that clearly were built on mistrust. Because of the program's apparent simplicity, together with the rushed timing and the fear of being taken advantage of by employers, policymakers did not include the private sector in the design process. As noted earlier, this process prevented the main stakeholders from being mutually invested in the program, or developing strong, productive relationships.

Even after the two rounds of program modifications had addressed some of their concerns, private sector employers were still not attracted to enroll. This illustrates that both sides did not understand the problem and each other's perspective accurately and could have benefited from more in-depth studies and collaborative participation in identifying how best to generate new employment opportunities.

Failure to Learn from Failure

Potential consequences for risk-taking and experimentation with innovative approaches are generally seen as negative and dissuade the exploration of novel, evidence-informed interventions. Fear of failure can further hinder the incorporation of novel evidence into practice, even when stakeholders recognize the value and applicability of the evidence.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Programa Primer Empleo had been operating for three years and was known not to be achieving its expected results. The evaluation identified international evidence that corroborated the lack of effectiveness of unidimensional tax incentives to promote employment. Nevertheless in October 2010 a new effort to incentivize employment was carried out. The Senate¹²⁹ approved amendments to the income tax law, and the Law to Promote First Employment (*Ley del Fomento al Primer Empleo*). The main difference between Programa Primer Empleo and the new Law to Promote First Employment was that the incentive in the new law consisted of an additional deduction in the employer's income tax from each employee hired and registered at IMSS for the first time.

The two programs at their core were very similar, and the new legislation did nothing to address the underlying flaws that international evidence illustrated and the lessons learned from the design and operation of Programa Primer Empleo, even though these were widely known and had been identified in PPE's evaluation.

Conclusion

The case of Program Primer Empleo illustrates several of our research project's broader themes, especially the role of timing

Although the lessons learned by the design and implementation of PPE were difficult ones, some of our informants recognized the value that PPE had on how these lessons contributed to a richer knowledge about how to support the design of public policy in the future.

in the incorporation of evidence into practice; the role of evidence definition, creation and use; the need to devote time and resources to learn about available evidence; the importance of building trust among stakeholders; and the failure to learn from failure.

Although the lessons learned by the design and implementation of Programa Primer Empleo were difficult ones, some of our informants recognized the value that PPE had on how these lessons contributed to a richer knowledge about how to support the design of public policy in the future. A high ranking official recognized Programa Primer Empleo as a failure from the public policy standpoint, but pointed out how PPE has

also served as an experimental exercise to understand the complexities of policy design and how the government could better design and implement social programs, so that they would be addressing the issues in a rigorous and productive way.¹³⁰

129 México, D.F., a 26 de octubre de 2010.- Dip. Jorge Carlos Ramírez Marín, presidente; Sen. Manlio Fabio Beltrones Rivera, presidente; Dip. Cora Cecilia Pinedo Alonso, secretaria; and Sen. Adrián Rivera Pérez, secretario

130 Miranda González, Sergio and Salgado Vega, María del Carmen. La nueva Ley Federal del Trabajo en México formaliza la flexibilidad laboral. Revista Trimestral de Análisis de Coyuntura Económica. April-June 2013. Vol. VI Number 2. [Available here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Stakeholder Map

Programa Primer Empleo was mostly a government program, so the stakeholder map describes the relationships between government agencies.

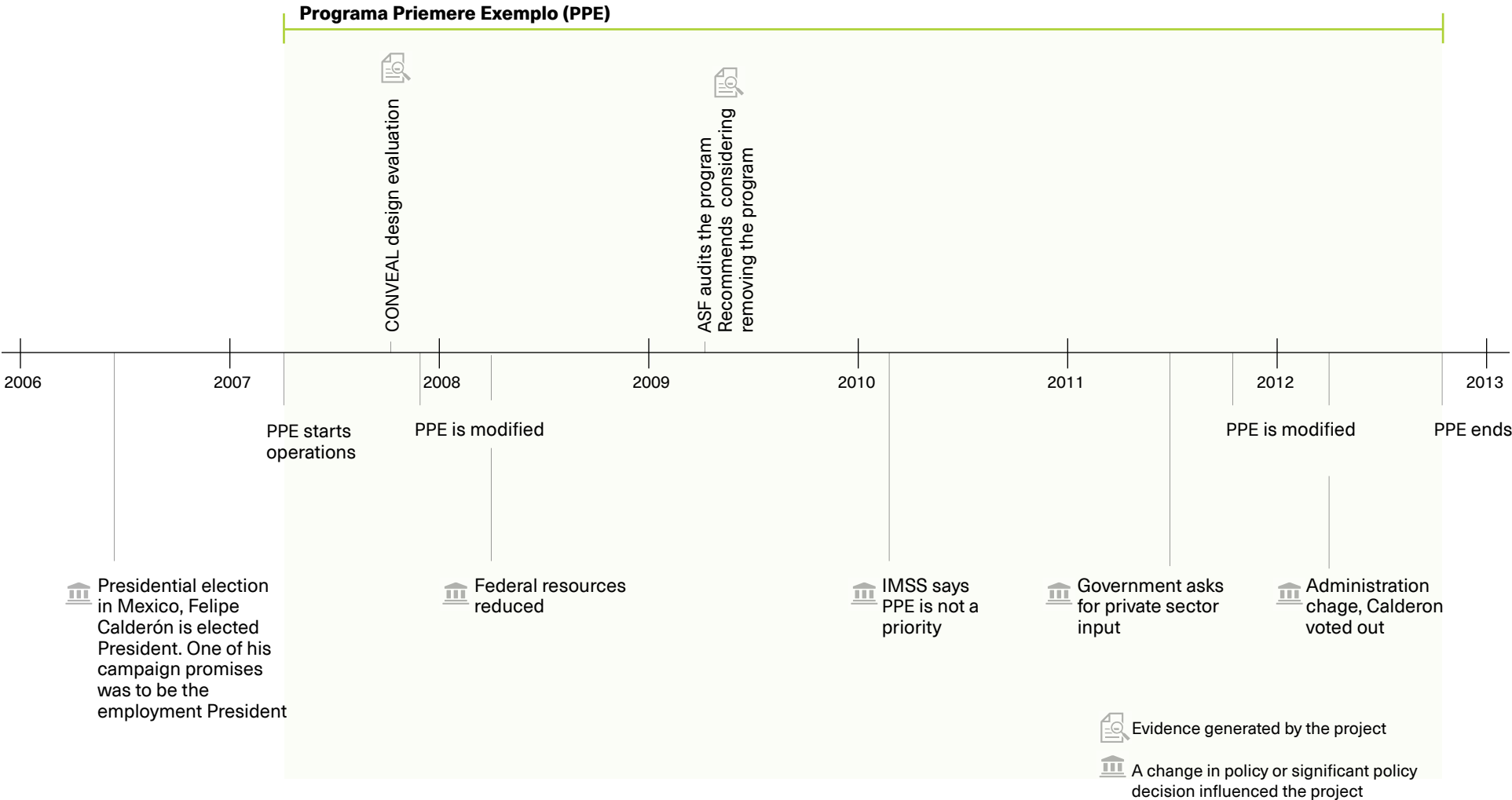
This stakeholder map is a visual representation of the major stakeholders involved with this project. The importance of each of the actors is defined by their relative size, and their proximity to the center of the project. Their role is defined by the color; multiple colors indicate multiple roles. 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.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Timeline

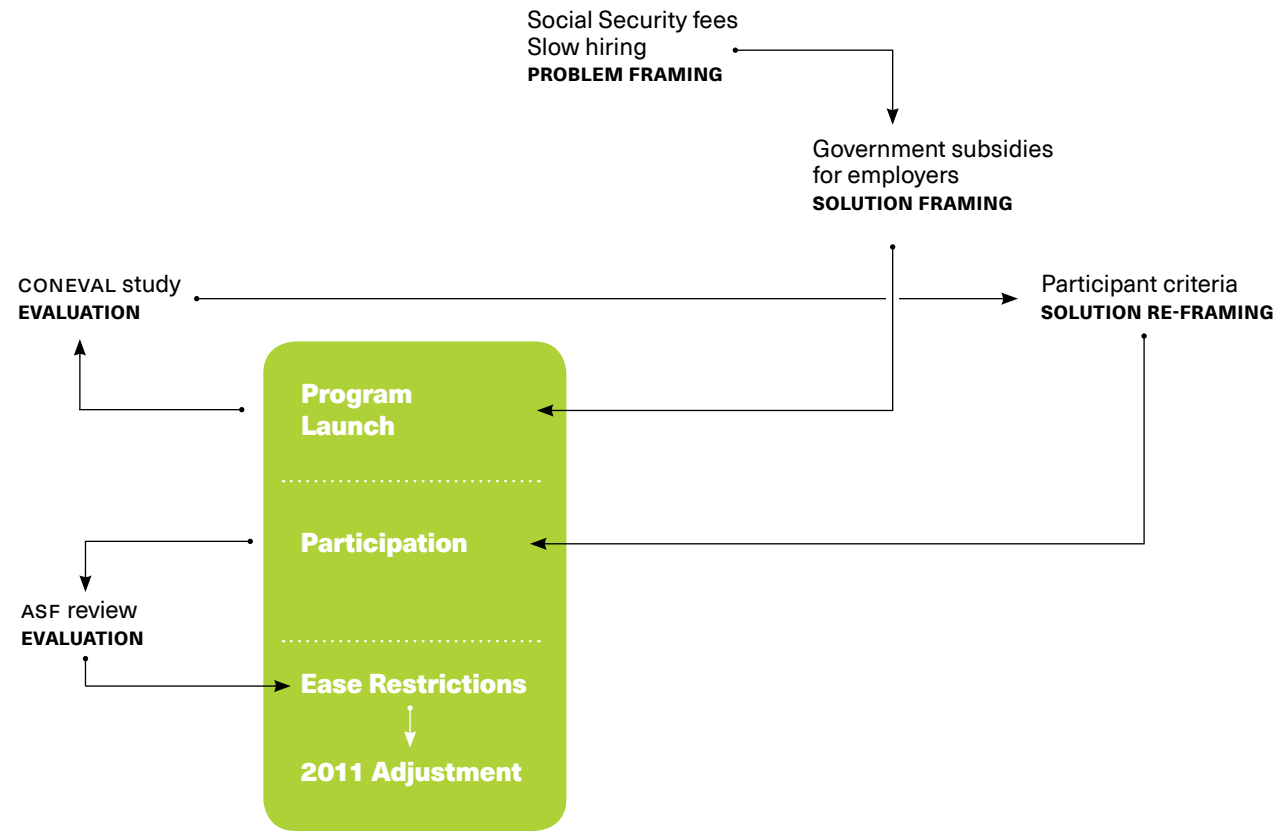




Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Process Diagram

Programa Primer Empleo was launched at a national **scale**.





Appendix 1

2007 Communication Actions for Programa Primer Empleo

Printed	1,001,000 brochures were distributed through IMSS sub-delegations, the President's residence when the program was launched and several unions.
Phone Campaign	From May 7 to June 22, 2007, through IMSS contact center, a telephone campaign was carried out to inform HR staff members at companies affiliated to IMSS about the benefits and requirements of Programa Primer Empleo. By the end of the campaign 78,344 companies had been contacted.
Media Campaign	From June 11 to July 10, 2007, a marketing campaign was carried out in the following media outlets: 27 radio groups, 2 television networks, 11 newspapers in Mexico City and 77 local newspapers, 24 magazines and Postcards and advertising banners.
Post Service	To promote the registration of new companies to the program, letters were sent during 2007. An invitation to all 822,978 employers was sent via the Mexican post service (SEPOMEX) (July and August) and 234,840 by email (May to July).
IMSS Webpage	A direct access to information about Programa Primer Empleo was enabled in the main IMSS webpage.
IMSS Contact Center	Through the IMSS contact center more than 100 thousand queries about the requirements and benefits of the program were answered.



**Evidence in
Practice**

Full Report

Print

Case Studies

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Page **183** of **220**

SOUTH AFRICA | EMPLOYMENT

Collaborative Analysis of Labor Intervention Effectiveness

An innovative collaboration between policymakers and researchers grapples with the pressing issue of unemployment.

By Katherine Wong





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part I: **The CALIE Story**

Unemployment is a critical political issue in South Africa. The unemployment rate in the country is extremely high at 26%. Youth unemployment (ages 18 to 35) is particularly dire at 37%.¹³¹ In particular, black South African women suffer from higher unemployment rates than black men, despite being more educated on average. Public officials, politicians, and organizations working to improve the economic and social environment in South Africa are eager to identify cost-effective ways to address unemployment, particularly among black women and youths.

The Department of Labor (DoL) is a key government unit responsible for ensuring that the labor market functions appropriately and is specifically tasked to reduce unemployment. One of the DoL's unemployment reduction programs, offered through labor centers in each province, provides counseling services to job seekers, seeks to match labor market supply with demand by connecting the unemployed with companies and organizations offering positions. While there had previously been efforts within the DoL to gather national-level data about the labor force, there had been no systematic follow-up data collected on job seekers after their contact with career counselors at the provincial level.

Background on J-PAL

The Abdul Latif Jameel Poverty Action Lab (J-PAL) is an internationally-known organization that aims to “reduce poverty by ensuring that policy is informed by scientific evidence.”¹³² The organization opened an office focused on research and policy in Africa in August 2010 at the Southern Africa Labor and Development Research Institute (SALDRU) at the University of Cape Town (UCT), with funding from The William and Flora Hewlett Foundation and other donors.

Because J-PAL has received stable, unrestricted funding from core donors (on top of its more variable, project-related funding), it has been able to spend some of its resources in a flexible manner that can be responsive to the policymaking environment. J-PAL has

used this funding to form strong relationships with the government of South Africa through workshops and other networking/knowledge sharing events, which facilitated the creation and development of the labor policy project described in this case study. J-PAL also benefits from a longstanding relationship between SALDRU, UCT, and government entities in South Africa. The relationships between J-PAL and other stakeholders are described in more detail in this case.

J-PAL Africa is structured into two main verticals—research and policy. The research team typically focuses on partnering with academic researchers to conduct studies, while the policy team focuses both on disseminating research emerging from the network to inform government and NGO partners who can implement the research findings and on assisting these partners with implementation and scale-up.

¹³¹ Martin Abel, Rulof Burger, and Patrizio Piraino, “The Value of Reference Letters- Experimental Evidence from South Africa,” Harvard University, 2016. [Access here.](#)

¹³² MIT, Abdul Latif Jameel Poverty Action Lab. [Access here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

In contrast to this typical bifurcation of research vs. policy in J-PAL Africa's organizational structure, policy outreach in the context of the project described in this case was largely undertaken by the project's principal investigators and J-PAL Africa's research team. The large amount of time and effort spent on policy work is also unusual for most projects in academia broadly. This is worth noting because such exceptional dedication to policy as well as research is one of the key factors that enabled evidence from the labor study to inform policies and practices at the DoL.

Launching the Labor Study

In January 2011—shortly after its inception—J-PAL Africa organized a workshop to share its mission with South African stakeholders and facilitate exchanges of ideas. A number of prominent policymakers, including Trevor Manuel, the former finance minister and then head of the National Planning Commission of South Africa (NPC),¹³³ were in attendance. Manuel knew Abhijit Banerjee, co-founder of J-PAL, from a previous Harvard-led research initiative on economic development policy in South Africa around 2006 to 2008.

At the workshop, Manuel was able to further develop his relationship with Banerjee, leading to a new partnership between

CALIE's ultimate goal was to identify initiatives that lead to positive outcomes, help inform policy decisions, and reduce unemployment.

the NPC and J-PAL, formalized as the Collaborative Analysis of Labor Intervention Effectiveness (CALIE). In order to address one of the NPC's key mandates—the issue of youth unemployment—J-PAL Africa raised funds and began to seek proposals for grants for projects related to labor policy. It was formally promoted as a joint research program between the NPC and J-PAL Africa designed to develop and implement rigorous impact evaluations of youth employment policy interventions. CALIE's ultimate goal was to identify initiatives that lead to positive outcomes, help inform policy decisions, and reduce unemployment.

In preparation of the workshop, the principal investigators and the J-PAL Africa research team surveyed studies on unemployment

within and outside of South Africa. This included data that found that firms typically look to their social networks to fill job vacancies. As a result, otherwise qualified applicants with weak social networks were disadvantaged, particularly if they were female (Schöer & Leibbrandt, 2004; Magruder, 2010). The principal investigators also reviewed academic papers focused on interventions designed to reduce the gap between intention and behavior in the health sector. These interventions included writing action plans, sending reminders to participants, and providing peer support. The research team sought to test the effectiveness of these mechanisms in the context of the labor market.

Eight months later, in August 2011, CALIE hosted a workshop (funded and run by J-PAL) focused on generating new research ideas on labor policy in South Africa. The workshop attendees included several international researchers (including Abhijit Banerjee), local researchers, and NPC members. The goal of the workshop was to link South African

¹³³ The NPC was established in 2010 with the objective of creating a long-term strategic plan for South Africa's development. Republic of South Africa, Department of Planning, Monitoring, and Evaluation. National Planning Commission. [Access here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

researchers with international academics who sought to work on similar research topics. It culminated in a document to solicit research proposals focused on reducing the South African unemployment rate. This document was later emailed to various high-caliber economics programs at several of the country's leading universities.

At the same time, Martin Abel—then Research Manager at J-PAL Africa and now a recent PhD graduate in Public Policy from Harvard University—led outreach presentations and workshops at universities in the Western Cape outlining J-PAL's mission, partnership model, and working methods—with an emphasis on the benefits of randomized controlled trials.

J-PAL's promotion efforts led sixteen academic teams to submit research proposals, including one from Professors Patrizio Piraino (from the University of Cape Town, or UCT) and Rulof Burger (from Stellenbosch University). The two professors, who went on to become the principal investigators of the project, had previously met through academic networking events and were interested in working with J-PAL because of its reputation and funding opportunities. They also shared a personal desire and sense of responsibility to work on the topic of unemployment in South Africa.

This motivation was enhanced by UCT's faculty evaluation criteria, which explicitly assessed the "social responsiveness" of an academic's research, which encouraged faculty to engage in socially relevant work. One professor stated that he "can only imagine doing work that may end up being useful for the government."

CALIE approved Piraino and Burger's preliminary research proposal, along with seven others, in December of 2011 and provided a USD\$10,000 grant from J-PAL Africa. In January CALIE hosted a workshop with all the accepted research proposal teams in order to provide feedback on the research questions and identify potential government partners. J-PAL Africa acted as an intermediary, using its strong relationships with government officials to connect Patrizio Piraino and Rulof Burger with the relevant people at the DoL to get the project off the ground. Given how unemployment was (and still is) such a crucial political issue in South Africa, the DoL was eager to explore alternative solutions and welcomed the research collaboration.

The two professors, J-PAL Africa, and the DoL worked together to implement the first pilot study at the Sandton regional Labor Center. J-PAL Africa provided funding for





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

This collaborative design process was crucial to the success of the project because it enabled the researchers to explicitly design the study to fit within the Department of Labor’s resource limitations, incentive structure, and existing work flow.

implementation, research assistants, and support managing the project's budget and logistics. J-PAL Africa's Martin Abel eventually left his position at J-PAL to begin his PhD at Harvard and was brought on to the team as a co-principal investigator.

Research Questions + Pilot Project

The research proposal sought to determine whether a “skills signaling” intervention could help disadvantaged and low-income job seekers overcome a lack of credibility that other job candidates with “credible signals”—such as a university degree or the right social network—might have.

Labor center career counselors regularly ran workshops aimed at job seekers, in line with their mission and the DoL guidelines. These workshops covered how to write CVs, networking skills, job applications, and other topics. For the pilot study, the principal investigators worked with J-PAL Africa and the DoL to design a workshop session on reference letters (in which they were able to collect rigorous data about their usefulness), and which the career counselors then added to the workshops at the Sandton Labor Center.

The principal investigators led the process to design and evaluate the interventions, but would regularly meet with J-PAL Africa

and the DoL to incorporate feedback and comments on their proposed design. While there was no formal name for this coalition, stakeholders credit Martin Abel for spearheading the meeting schedule, and acting as an intermediary. They also credited Laura Poswell, the executive director of J-PAL Africa, for driving the collaborative effort with the DoL by sharing information with the principal investigators when they were unable to be present for the joint meetings. Additionally, the research assistants who worked with the principal investigators conducted focus groups with the DoL in order to determine what intervention should be the focus of the research question. These focus groups led to the reference letters being added to the project.

This collaborative design process was crucial to the success of the project because it enabled the researchers to explicitly design the study to fit within the DoL's resource limitations, incentive structure, and the existing work flow of labor center employees. By designing the project in this way, the researchers tried to ensure a smooth implementation for the DoL that, if successful, could easily be scaled. For example, the researchers learned that, for the project to scale, its costs would have to fall within the existing DoL budget. They also



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

identified career counselors as a lynchpin of the intervention, so they structured the project to fit within the scope of the counselors' current performance metric of serving a certain number of job seekers each quarter—because labor center employees could use the intervention to better meet their performance goals, they had an incentive to support the project.

The DoL also hoped to leverage the project to access data on job seekers, given its own data collection limitations. Some employees stated that “there’s no tool or measurement in place that can actually measure the value of the career counselors or the career counseling that takes place” and that the DoL would “love to have more information ... to say it does have an impact, especially in the area of [job] placement.” Career counselors in particular expressed a strong desire for more data on the effectiveness of their services, such as information on job seekers' employment status after the career counseling sessions. There were a few minor disagreements in these discussions. Some government staff were concerned that the reference letters would not be helpful to those job seekers with a shorter job history. Additionally, there was some discussion over which interventions to include (e.g. reference letters versus the impact of counseling sessions



generally). Although J-PAL Africa and the principal investigators had to adjust their research focus slightly to get the government to support the project, they saw these compromises as critical in order to ensure that the research question was relevant and timely for the DoL, and to establish a

relationship of trust that would allow for more ambitious and extensive research in the future. In the pilot implementation, the career counselors would teach a group of workshop attendees about the potential benefits of using a reference letter, and provide reference letter templates to be completed by former employers. Career counselors would also teach the same workshop without the reference letter component in order to create a control group. Counselors and J-PAL Africa jointly accepted the filled-out reference letters. Then the J-PAL Africa research team applied each month to open positions on behalf of the work-seekers, using an experimental approach that varied reference letter inclusion. This variation helped to further test the effectiveness of the letters.

All of the participants felt positively about the design process and repeatedly described it as “very cooperative.” Everyone felt their suggestions were incorporated fairly. One of the principal investigators attributed this cooperative experience to the “shared interest in trying to meet the urgent demands this country has. [...] When you operate in an environment where the real questions are not only interesting academically, but they’re also urgent politically and socially, then it’s easier not to have conflicts... This project was the best-case scenario interaction between



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

government and academia because they were interested in our results, [and] they also informed what we were doing... There was a lot of input from the government." The DoL staff reiterated this shared personal interest; when asked about her motivation to work with the researchers, a career counselor said, "[I see] people on a day-to-day basis who are really struggling and needing intervention and needing tools to get somewhere... I think [the intervention] was quite a good initiative and very practical."

The initial results from the pilot study were positive, indicating that job-seekers who sent out reference letters with their CVs had a higher call-back rate than job-seekers who did not send out reference letters. The pilot also showed that an intervention of this type could be woven into the existing operations of the labor centers at basically zero added cost. However, the pilot did not show whether the intervention could operate at the scale needed to have a real impact in the labor market. Thus, the principal investigators and J-PAL Africa decided to apply for funding to scale-up the research project.

Larger Scale Impact Evaluation

The researchers and the DoL collaborated to determine how best to evaluate the intervention on a major scale, given funding constraints. The team scaled up the number

“This project was the best-case scenario interaction between government and academia because they were interested in our results, [and] they also informed what we were doing.”

RESEARCHER

of labor centers involved from one to four. Based on requests from DoL staff, the team also increased the scope of the study from just the impact of reference letters, to the impact of two interventions: (1) reference letters and (2) action plans—a template in which individuals would work with counselors to map out next steps for the job search, which could help job seekers increase the number and quality of job applications.

As in the pilot study, all participants had a positive and cooperative experience. The design process was similar to that for the pilot project: all stakeholders held regular meetings with each other to ensure their voices were heard, and the principal investigators incorporated suggestions into the design. To complement these ongoing meetings, there were two workshops with the DoL and career counselors to get their feedback on the design and to talk about the results.

The team applied for and received funding from two sources to enable this larger scale evaluation. The first, Programme to Support Pro-poor Policy Development Phase II (PSPPD II)—a joint program between the European Union and the Presidency of South Africa—had a call for research proposals focused specifically on reducing poverty and inequality. The second funding source was the World Bank Gender Innovation Lab which had a call for research proposals for projects focused on gender. In 2014, the project received funding from PSPPD II to run a full-scale, four-site randomized controlled trial. In 2015, the team received additional funding for the project from the World Bank Gender Innovation Lab. Overall, funding from these two sources totaled about USD\$250,000.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

In the 2015 implementation of the project, the four labor centers were: Soweto, Krugersdorp, and Polokwane, in addition to Sandton. They chose these labor centers based on conversations with the DoL. The criteria were which centers would be easiest to work with, given office logistics, location, and number of registered job-seekers. Offices in prime locations with a large number of job-seekers would make it easier for the researchers to obtain a larger sample size. Selection also took account of the DoL's interest in working with job seekers in different kinds of geographic areas (the Polokwane labor center was specifically included on this basis). The DoL also provided access to a database of people who had signed up for job counseling services, which the researchers used to invite job seekers to the labor centers.

Reference Letters

The reference letter component of the full-scale evaluation was an extension of the pilot, with the scale was expanded to test more participants. The study demonstrated that reference letters helped job seekers progress in their search to find employment, showing a 60% increase in callbacks. Employers found the reference letters a credible signal of the workers' skills. The study also showed that women were, on average, more diligent about taking the reference letter templates to their

The Department of Labor was particularly pleased with the results because the intervention did not require additional resources beyond the time of the career counselors, and it fit within the career counselors' existing incentives.

former employers and following the process through to completion, which could partly explain why the reference letters especially improved callbacks for women by 89%.

Action Plan

For the action plan component, the project evaluated four interventions. Each intervention was a combination of one or more of the following: the action plan itself (i.e. weekly goals and a plan to achieve those goals), a job counseling workshop, peer support (i.e. the research team would send a message to a friend of the job seeker describing the latter's goal), and reminders.

In total, the study worked with a sample of about 1,100 registered job seekers. These job seekers were primarily unemployed, age 18 to 35, who were actively searching for jobs but with low intensity (on average, 4 applications per month, with 8 hours of job searching spent per week). They were all black South Africans, and evenly split between males and females. The research team collected data from several phone surveys and tracked the number of applications submitted, hours searched, employment status over time, and other metrics.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The study found that, although the action plan had no effect on the number of hours spent searching for a job, it significantly increased the number of job applications submitted (by 15%) and the number of job offers (by 30%). It increased the effects of normal counseling sessions by helping job seekers follow through with their goals. The reminders and peer support interventions were found not to be effective.

The research project concluded in June 2016, and the principal investigators and J-PAL Africa shared their findings with the DoL and regional career service coordinators through a workshop in October 2016. The research team and J-PAL Africa also held debriefing meetings for the career counselors to share their findings and answer questions about the process.

The DoL was particularly pleased with the results because the intervention did not require additional resources beyond the time of the career counselors, and it fit within the career counselors' existing incentives (i.e. the intervention enabled career counselors to achieve their goal of seeing a certain number of job seekers each week). Additionally, the DoL was "quite happy... that the results are something that can be implemented" and are easily actionable (as opposed to showing how a certain intervention may not lead to an

expected outcome). This was beneficial from a morale perspective because the DoL had research findings that it could act on. While it is useful that the research showed that two of the interventions had no effect, the DoL interviewees stated this finding alone would not have generated much enthusiasm because it would have left the DoL with the same low placement rate and without any significant action items. One interviewee stated that when this happens, it is "quite difficult then because [the stakeholders] have gone through this whole process... but then the end result isn't something that you can clearly put forward" and implement.

The research paper was finalized in November 2016. Many of the interviewees noted that the project's turnaround time was incredibly short, just 16 months from the start of the scaled-up project to the finalization of the paper. One J-PAL Africa employee stated that he had "never seen that on any other project." This quick turnaround can be attributed to several factors: unemployment was, and is, an extremely pressing issue in South Africa, causing the DoL to be extremely supportive and responsive; Martin Abel wanted to use the project as his job-market paper in time for his graduation from Harvard; and there was strong buy-in for the project from all stakeholders at every level.

When pressed to describe any challenges faced over the course of the project, all stakeholders had only positive comments, reiterating that the project went smoothly and was cooperative. While there may have been minor disagreements about how best to design the project, all interviewees described such disagreements as discussions rather than conflicts. One important thing to note is that the study was always referred to as a DoL project, as opposed to a J-PAL one, in order to ensure that the Government of South Africa was located at the center of a project in which it was not only a key stakeholder, but whose engagement would be essential for following through on expanding implementation based on positive outcomes.

As of the time of publication of this case study, J-PAL Africa, the PIs, and the DoL are discussing an expanded roll-out of related policies and practices based the results of the project, with new plans expected in early 2018.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part II: **Key Themes + Insights**

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

Definition 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.

Broadly speaking, J-PAL defines evidence as impact measures coming from randomized controlled trial evaluations, but recognizes that this definition may be narrower than that used by other stakeholders, especially policymakers and development professionals.

In the context of this case study, all participants agreed about the validity of the research study. Generally, there was agreement both on the project design and implementation. Although J-PAL Africa and the principal investigators had to alter the research design slightly to get the DoL onboard, the changes were relatively minor and, from the perspective of the researchers, worthwhile in order to lay down a positive foundation for a long-term working relationship with the government. Furthermore, researchers also offered to collect and analyze data that was not strictly necessary for their research, but that was extremely valuable for their government partners, both because it would inform their work and because it would be collected with the project's—and not the government's—resources. This also helped to foster a positive, long-term working relationship.

Incentives

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.

Shedding light on the incentives of each stakeholder provides a clearer understanding of what mechanisms can be changed in order to facilitate the translation of evidence into policy and practice. In the CALIE case, the incentives of all the main stakeholders were aligned such that everyone could work towards the same goal, facilitating the success of the project.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Researcher Incentives

In the context of this case study, interviewees stated that the researchers were very involved in the projects on the ground and with disseminating their work after the study was completed. It is interesting to note that the policy team at J-PAL Africa did not work on this. All policy-related outreach was conducted by the principal investigators and the J-PAL research team. Many interviewees stated that this amount of dedication to policy and implementation considerations is unusual to see among researchers. This dedication can be attributed to several incentives.

Firstly, while this degree of attention to policy and implementation considerations is atypical among most academic researchers, J-PAL Africa's research team and the principal investigators saw themselves as the group best placed to work with the government policymakers and implementers. This kind of engagement is part of J-PAL's stated mission. Once the dissemination and scale-up design phase is complete, J-PAL Africa plans to have a policy team member work closely with the Department of Labour on the next phase of the project, assisting in the rolling out the intervention.

Secondly, the incentives of the academic sector in South Africa are structured to promote projects that have a social



benefit. South African academic institutions explicitly evaluate professors not only based on research, teaching, and administrative commitments, but also on social responsiveness. It is an important, structurally embedded way that the university encourages academics to conduct rigorous research that is socially relevant. The principal investigators involved in this study stated that their home universities evaluate them partially based on this metric, though there is some degree of self-selection, and they preferred to work in a university setting where social responsiveness is a priority.

Thirdly, personal motivations played a large role. The researchers emphasized that they felt a responsibility to “make a difference” on pressing issues because of the high poverty and unemployment in South Africa. The principal investigators involved committed above-ordinary resources and time to work on the policy aspect of the pilot study, given their personal commitment to the underlying issues. Stakeholders from all organizations (other researchers, the DoL, J-PAL Africa staff) mentioned several times that this level of effort on policy is very unusual for most researchers.

Another incentive for the researchers' interest in policy engagement was their desire to build a strong relationship with the DoL in order to conduct more research in the future. Policy-related data in South Africa is not always publicly available due to a lack of systematically-collected data and a lack of electronic infrastructure to publish data online. By forming a good relationship with the DoL, the researchers could build a relationship that would allow them to work on future projects that could both increase the utility of existing internal DoL data and generate new data from scratch. When asked about why the researchers conducted more policy advocacy work than usual, a J-PAL Africa employee explained: “I think what's interesting in that case is also the fact that a lot of them are local



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Another incentive for the researchers' interest in policy engagement was their desire to build a strong relationship with the Department of Labor in order to conduct more research in the future.

South Africans or have a strong link here... So there [are] probably two reasons why they are so involved... The local link, and the fact that they were involved with the DoL from the very beginning... There's been a continuous relationship and I think they want to continue having that relationship probably for other research projects that they do later."

Additionally, the pilot study was the basis for Martin Abel's PhD job-market paper and therefore had a great influence on his future professional prospects. As a result, he had a particularly strong incentive for the pilot study to be conducted rigorously, quickly, and ideally make a significant policy impact.

Implementer Incentives

The project was explicitly designed to be implemented by the DoL and to work within the existing incentive structure and work flow of labor center employees. Labor center employees were not resistant to working on it because the project enabled the employees to better achieve the targets on which they were assessed, specifically, the number of job seekers they work with each quarter. Labor center staff have high targets in terms of the number of job seekers they need to counsel per year (upwards of 1,800), with the definition of "counseling" including one-on-one consultations, workshops, and other group activities. A DoL employee said: "we've

got certain indicators that we [measure] on job seeker registration, career counseling, placements, and so on. There is a target and then there's a target for each official. Each official has to do X number a quarter or annually. That is basically what you get measured on... how performance is evaluated at this stage within the department."

Counselors expressed a strong desire for feedback on their services—updates and information on job seekers' employment status after the career counseling they received. Another important part of the pilot study's appeal was the inclusion of a feedback loop through which the researchers shared the results of the research with DoL and the labor center staff who participated in the study. Interviewed government staff said they were willing to work with the researchers because it would help them have more data on impact and quality (rather than only quantity) of services that they would not otherwise be able to collect. One J-PAL Africa employee stated: "The counselors, all of them in fact, ... were very committed to helping people so that was what I heard from the PIs that we were very lucky in that regard that it's not always the case with government or with any other organization that you have people who want the research project to go smoothly so that we can get the results and then from there figure out how to make things better."



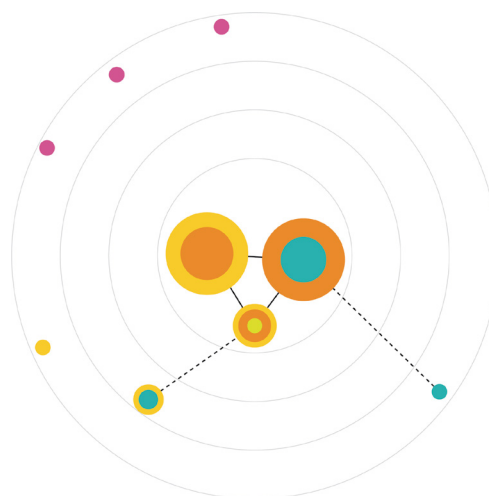
Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Intermediaries

The role of intermediaries is crucial to consider, as intermediaries are the actors that connect and facilitate communication between evidence generators and practitioners. Although intermediaries may not be the most visible participants when considering evidence and practice, many of our sources have explicitly or implicitly referred to intermediaries when citing examples of successful evidence-to-practice flows.

J-PAL Africa was cited by many interviewees as the key intermediary (“the glue”) that helped connect the researchers, the government, and other stakeholders. J-PAL Africa created the set of conferences and convenings where people across sectors could meet and learn of each other’s problems, methods, and ideas. They also provided three key resources: funding for the pilot project, administrative/ logistical/ budgetary support, and assistance connecting the principal investigators researchers with the government. In particular, J-PAL Africa had strong connections with the government, which attracted researchers to partner with it to run studies and communicate findings. Findings were disseminated through meetings, workshops, or other avenues where the government was a direct partner. As one

principal investigator said: “J-PAL is also very good at engaging with policymakers.... It identified key people in the department who we could approach for this idea ...” Another level of intermediation is embedded within J-PAL Africa’s structure itself. Because J-PAL Africa is physically housed at the University of Cape Town with the Southern Africa Labor and Development Research Unit (SALDRU), it benefited from informal linkages, notably existing connections between SALDRU and government officials.



[View all the stakeholder roles.](#)

Timing

The different and often discordant timeframes within which researchers, policymakers, and implementers operate often hobble efforts to coordinate, let alone collaborate, on evidence-informed approaches. Electoral cycles and political windows differ from NGO funding cycles and from academic publishing rhythms. Yet each actor is bound by the timeframes of her formal stakeholder group.

The CALIE project was carefully and strategically timed, enhancing its chances of success.

J-PAL Africa strategically chose to work on this labor-related project because the government would be more responsive, since unemployment was a politically critical issue in South Africa. As a result, the government was keen to work with the researchers in a timely way.

Additionally, there was strong buy-in from each of the stakeholders at every level. The researchers had a strong relationship with the DoL, and the DoL had buy-in at both a high level (heads of the provincial labor centers) and on the ground (career counselors). This buy-in was created through a lot of hard work by the researchers to be good working partners and through the



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

strategic design of the research project. The strong buy-in, coupled with a push from the principal investigators, created a positive feedback cycle, expediting the timing of the research project even more. By pushing hard to carry out the project in a timely way, the researchers increased the likelihood that key people from the DoL would still be in their positions, so the main decision-makers from the DoL on implementation were those who had worked directly on the project and it was “fresh on their mind.” The researchers followed through with feedback for all stakeholders by holding debriefing presentations for the career counselors to share their findings and answer questions about the process, further increasing buy-in and expediting further scaling.

Capacity and Resources

Few organizations provide incentives or carve out explicit time for managers to explore emerging evidence in their field. Even fewer assign staff to find relevant evidence and translate it into accessible formats for the organization. As a result, the role of preparing and sharing evidence that is timely, useful, and relevant for practitioners is sometimes explicitly played by formal intermediaries (e.g., certain think-tanks). More frequently, an actor who holds a formal role within another stakeholder group spontaneously takes on

If practitioners trust that the evidence generated is accurate and relevant to their needs, translation of that evidence is far more likely to be prioritized.

the (additional) responsibility for trying to integrate evidence, with no actor formally responsible for the process. Discovering and integrating evidence requires time, energy and funding.

In this case study, the researchers and the DoL had the necessary capacity and resources to work effectively together to design and implement the project with a focus on generating the evidence needed to evaluate and refine the intervention. J-PAL had received

steady funding, which had the flexibility to enable the organization to be responsive in how it invested its time and money. As a result, J-PAL Africa could afford to spend time forming strong relationships with the government, as well as underwrite research proposals explicitly aimed at addressing pressing policy concerns. Additionally, the DoL was excited about the project because the intervention responded to the self-identified priorities of the government, making use of existing public resources for implementation.

Trust and Convening Power

The cross-stakeholder collaborations required for evidence-informed policies and practices are often difficult to initiate, develop, and sustain. Particularly when institutional incentives are lacking, personal trust, respect, and buy-in between individuals across stakeholder groups become critical to fostering the effective flow of evidence into practice.

If practitioners trust that the evidence generated is accurate and relevant to their needs, translation of that evidence is far more likely to be prioritized. The trust between all of the stakeholders in the CALIE case facilitated J-PAL's use of its convening power to bring everyone together and ensure success.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

J-PAL's global reputation made it very easy for J-PAL Africa to gain access to the government in order to launch and coordinate the project. An employee stated: "As J-PAL, ... you can literally practically email anybody and they'll see you and you won't have to wait. It's very interesting... I think J-PAL just has built this unbelievable global brand name." On the ground, they also have strong personal relationships with officials within the government which they'd started building at their inception. J-PAL explicitly invested time and resources to cultivate and maintain these relationships in South Africa (thanks, in part, to its flexible funding, as previously described). Labor center staff said they enjoyed their partnership with the researchers and perceived them as "easy to work with," in large part thanks to J-PAL Africa's operational support on the ground and the respect they showed to government staff in the design and implementation of, as well as the follow-up communication about, the results of the program.

Additionally, J-PAL Africa was careful to refer to the reference letter study as a DoL project, as opposed to a J-PAL one, ensuring the Government of South Africa was correctly seen as the central stakeholder in the project and could receive the main credit for the project. As a result, there was

little competition for visibility between the government and the researchers. A J-PAL Africa employee said: "When you talk to the department and say... That's a J-PAL project—it's completely not the framing we use because it isn't. It's the labor centers' and... their reference letters and their action plans... I know that maybe it's more just a semantic thing, but we're very careful when we frame these things."

Learning From Failure

Potential consequences for risk-taking and experimentation with innovative approaches are generally seen as negative and dissuade the exploration of novel, evidence-informed interventions. Fear of failure can further hinder the incorporation of novel evidence into practice, even when stakeholders recognize the value and applicability of the evidence.

In the context of the CALIE project, all stakeholders were aligned in the goal of generating a positive impact on job-seekers. Nonetheless, all stakeholders understood that if the intervention had no impact on job-seeker results, it did not necessarily mean the project was a failure. Thanks to the co-design process, all stakeholders arrived at a shared goal of running the project smoothly in order to collect accurate data in order to





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

assess job-seeker outcomes. Fortunately, the project not only collected accurate data but also indicated the intervention had a positive impact on job-seeker outcomes.

Our sources did not explicitly describe how they would have responded had the project's hoped-for outcomes not been achieved. However, an indication that the stakeholders would have "learned" had the project been deemed a "failure" is that the funding to disseminate results for this project was not dependent on its success. J-PAL Africa's steady, unrestricted funding meant that results would be communicated – and that continued engagement with the DoL was not premised on the achievement of positive outcomes.

Conclusion

CALIE illustrates how the stakeholders approached the project as a collaboration in service of exploring a shared problem: what kinds of interventions might help address the high employment in the country? Key players within CALIE worked extensively to find an alignment in both institutional and personal incentives across stakeholder groups so they all shared a common goal. This shared objective guided the process of project design, implementation, evaluation, and follow-up. All stakeholders also shared a

commitment to learning from the project and to disseminating that learning.

Stakeholders also recognized what each had to offer one another. For example, researchers were extremely interested in the government data that the DoL had access to, a trove of long-term value to the researchers (and at no cost to the government). And J-PAL was more than willing to give credit for any success to where it would prove most useful: the government. These kinds of "negotiated exchange rates" can help bring disparate stakeholders together in exploration of a shared problem.

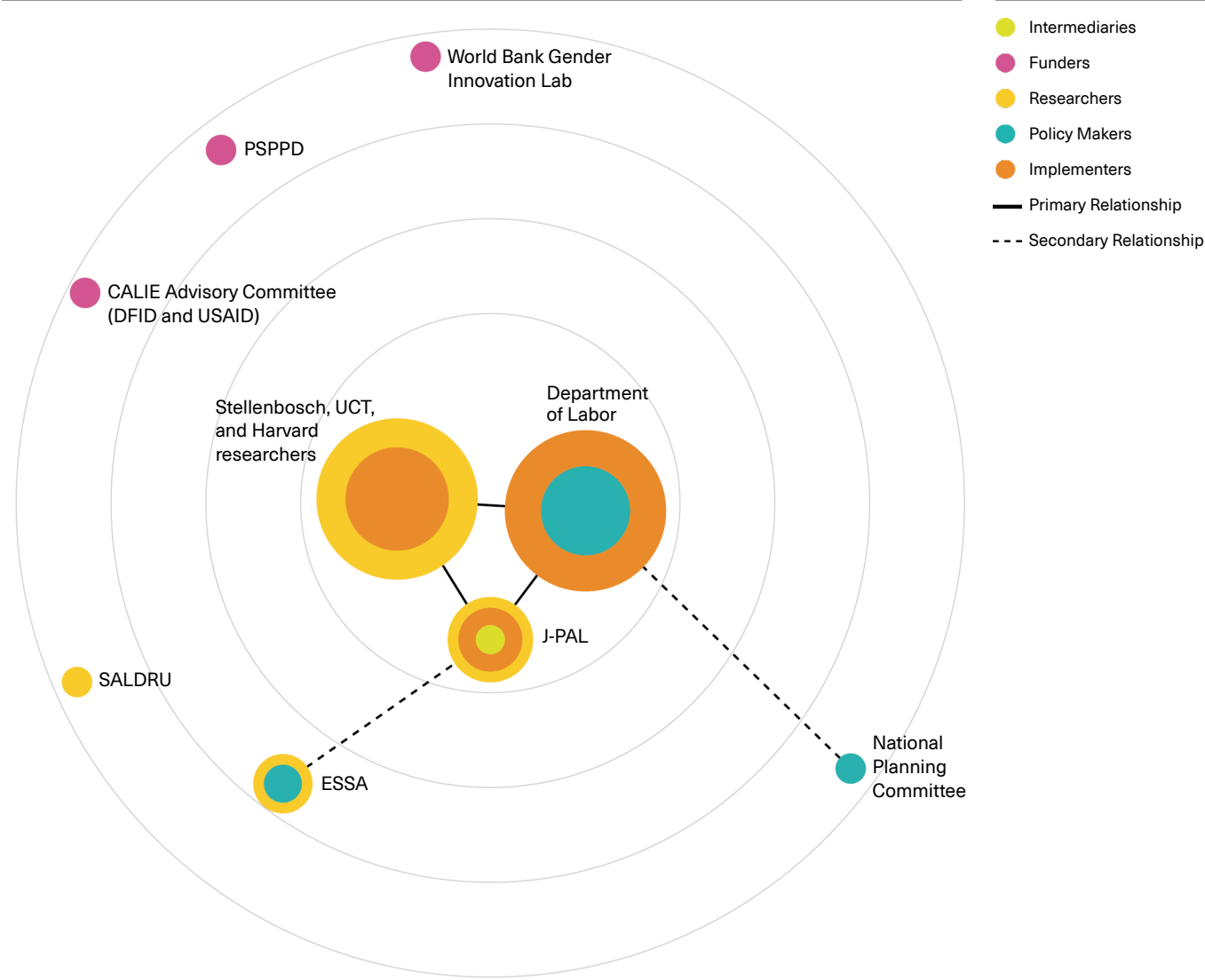
By identifying these key themes, this case study hopes to show that the lessons of CALIE's success can be adapted to other evidence-informed collaborations by other organizations.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Stakeholder Map

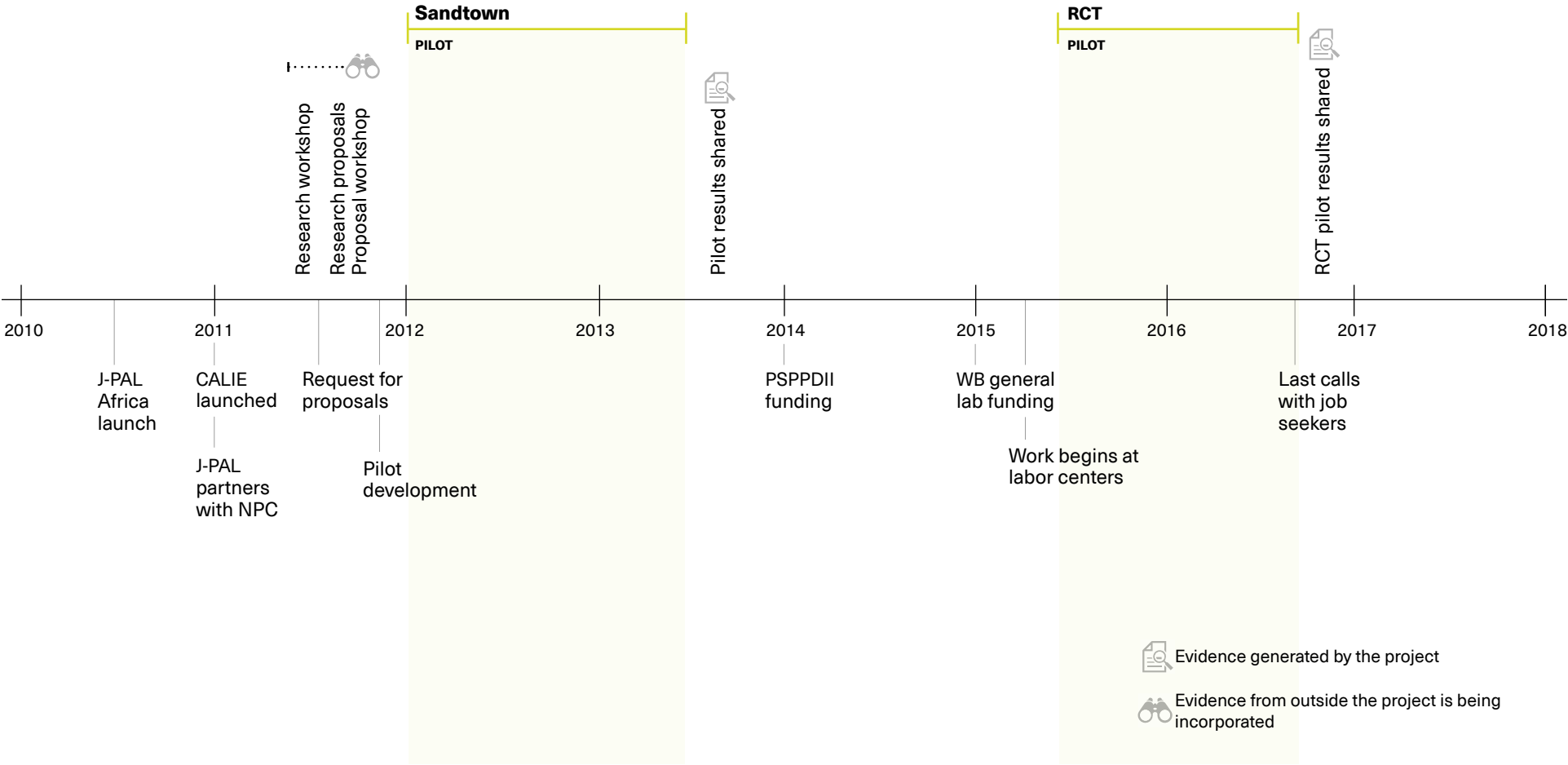
This stakeholder map is a visual representation of the major stakeholders involved with this project. The importance of each of the actors is defined by their relative size, and their proximity to the center of the project. Their role is defined by the color; multiple colors indicate multiple roles. 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.





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

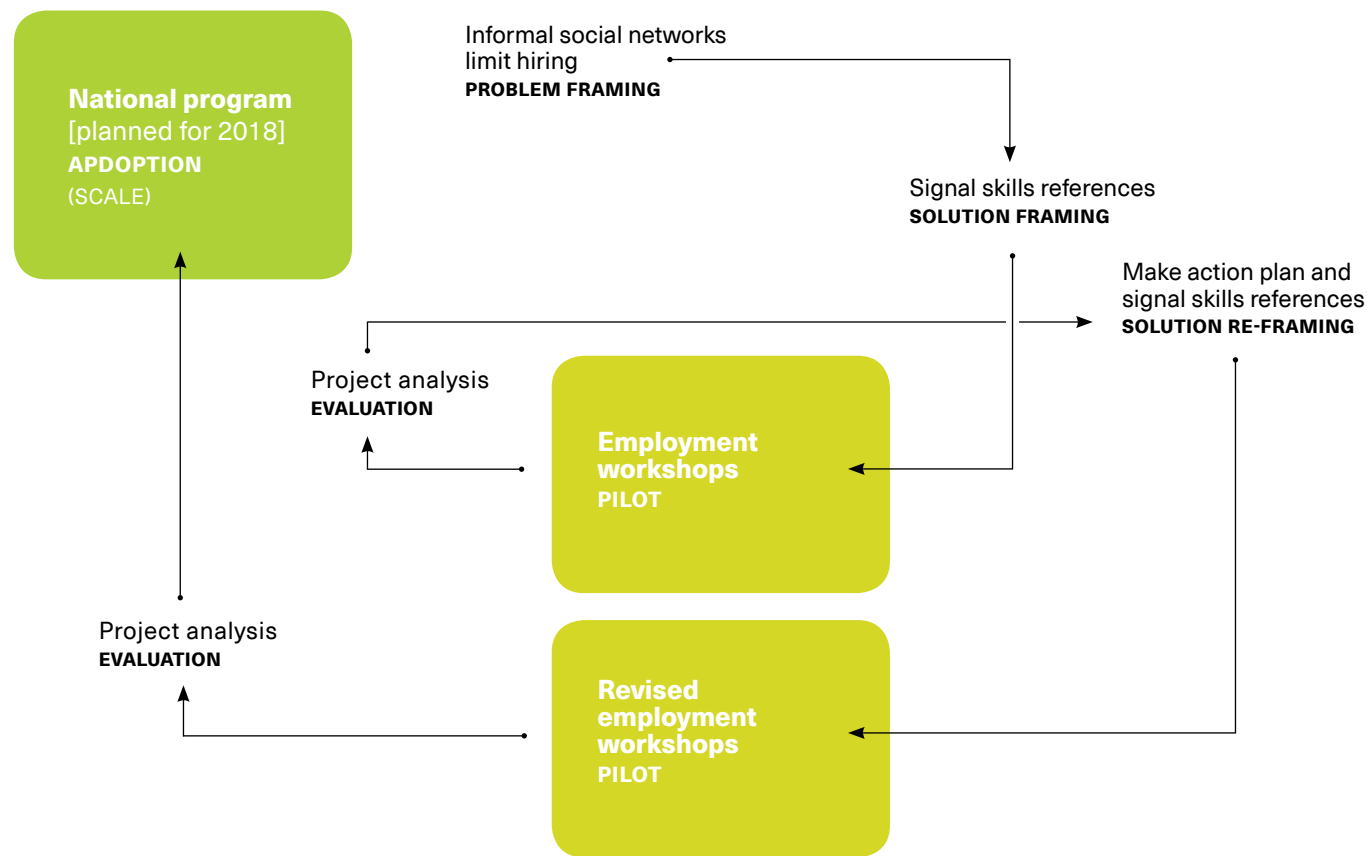
Timeline





Process Diagram

National adoption to reach **scale** is planned for 2018.





**Evidence in
Practice**

Full Report

Print

Case Studies

Page **202** of **220**

Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

SOUTH AFRICA | EDUCATION

FUEL: Feed, Uplift, Educate, Love

By creating a culture of accountability and evidence-based decision-making, FUEL is tackling how meals are distributed to students in schools across the country.

By Emilie Leforestier





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part I: **The FUEL Story**

School Nutrition in South Africa

Upon coming to power in 1994, Nelson Mandela launched a set of “Presidential Lead Projects” under the Reconstruction and Development Program. One of these projects was the Primary School Nutrition Program, which targeted the following objectives:

- improve education by enhancing active learning capacity, school attendance and punctuality by providing a nutritious meal early in the morning;
- improve health through micro-nutrient supplementation;
- improve health through parasite control/eradication;
- improve health through providing education on health and nutrition; and
- enhance broader development initiatives, especially in the area of combating poverty.¹³⁴

The Department of Health (DoH) designed the Program, which explains its initial focus on students’ health, nutrition and ability to focus. As a result, school officials had to report to both the Department of Education (DoE) on teaching-related objectives and to the DoH on nutrition-related objectives. For purposes of reporting and governance, the management of the program eventually shifted solely to the DoE in 2004. It was also renamed the National School Nutrition Program (NSNP) and phased into secondary schools from 2009 to 2011.

As of 2017, the program prioritizes 60% of all South African schools based on the socio-economic makeup of the areas they are situated in. It aims to provide a nutritious daily meal to every child in these schools.

Background on FUEL and the Project

This case study centers on a South African nongovernmental organization called Feed, Uplift, Educate, Love (FUEL). Its co-founders, Charles Luyckx and Gary Campbell, are fellows of the Africa Leadership Initiative (also known as ALI).¹³⁵ Luyckx was attending an ALI-related conference in South Africa in 2006 when he met with a woman working in the South African President’s Office. She explained that she had just come back from the South African province of KwaZulu-Natal and had observed that many of the students in poorer areas did not have access to a nutritional meal at school. Far from Nelson Mandela’s ambition of providing a full and balanced meal, some of the students she encountered were only receiving “old stale bread.”

Luyckx did not hear this story by accident. He was one of the founding directors of the restaurant chain Nando’s in the United Kingdom and had a solid background in food distribution. His fellowship with ALI included

¹³⁴ “Report on the Evaluation of the National School Nutrition Program,” Health Systems Trust. [Access here.](#)

¹³⁵ “Class III: Seriti.” Africa Leadership Initiative. [Access here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

a pledge to undertake a significant project to contribute to his community. Luyckx was a self-described “starter”—he enjoyed taking on new initiatives, being confronted with new challenges, and formulating and testing hypotheses to resolve these challenges. He was therefore on the lookout for an opportunity to give back to society, preferably in South Africa, focused on a compelling topic where he could add value. The problem of school nutrition sparked his interest and he resolved to learn more about the challenge of food distribution in South African schools.

Formulating FUEL’s Theory of Change

Luyckx’s initial hypothesis, heavily informed by his skillset and experience running Nando’s, framed the challenge as one of procurement and logistics. He thought that the procurement element could be solved by a large, centralized food purchasing hub that would aggregate the buying power of the schools’ entire food system and therefore realize large cost savings. Logistics had to do with delivery efficiency, which Luyckx suspected could be drastically improved.

Framing the problem as one of procurement and logistics was well-suited to the expertise of one of Luyckx’s trusted colleagues at Nando’s, Gary Campbell, Commercial Director of Nando’s in the UK, who had recently



decided to return to his native South Africa.

Together, Luyckx and Campbell set out to investigate how they could improve the NSNP program by leveraging their expertise in food management systems. The first few months were dedicated to building a diagnostic of the NSNP system, which they knew was complex: its governance, the links between national and provincial officials, and the connections between the latter and individual schools.

Based on these early interactions, the team started mapping out the processes, and

identifying the topics he was interested in exploring further. They came to understand that the NSNP’s implementation varied significantly across South Africa’s nine provinces. Some of the provinces had adopted centralized procurement models—which the team initially favored—while others had decentralized models in which the schools themselves bought the food they needed for their students, with funds provided through the program. This phase of the project, while not explicitly labelled as “evidence-gathering,” was aimed at building a fact base to inform. Its purpose was to build a fact base to inform the



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

design of a potential intervention. However, the hypothesis being tested was not whether procurement or logistics were areas to invest in and try to improve, but rather where the bottlenecks within these areas were, based on the assumption that the system could be more efficient under centralized procurement and logistics. The more the team researched the issue, however, the more often they were confronted with evidence that, in fact, some schools frequently had sufficient food delivered efficiently and effectively but they nonetheless failed to serve students nutritious meals on time. The problem, then, seemed to be less about logistics and more about accountability, reporting, and communication between stakeholders.

A Well-Connected Team

FUEL was funded by Luyckx and by the Yellowwoods Trust, an organization with strong social and economic ties to the business fabric of the country. These ties enabled FUEL to gain access to influencers within the education system that an organization with a different governance structure might not have been able to access, particularly so early in its history.

Some of these introductions turned out to be more closely aligned with FUEL's trajectory. In August 2007, after Campbell had gathered

The problem, then, seemed to be less about logistics and more about accountability, reporting, and communication between stakeholders.

a first evidence base, Robert Brozin, Nando's CEO in South Africa, sent a letter to the South African Minister of Education, indicating Nando's desire to help with challenges related to the NSNP.¹³⁶ Nando's brand name was a powerful way to gain access to the Department of Education. The letter enabled Brozin and Campbell to meet with the Director General of Education (head of the National Department of Education) and the Deputy Director General of Education. Their

presentation of what Campbell had uncovered to date triggered enough interest for them to be introduced to Neo Rakwena, Director of the NSNP at the DoE.

This top-down approach soon showed its limits. On the one hand, it had been very helpful for the team to be introduced to the NSNP team by high-ranking officials at DoE. On the other hand, there was no buy-in from the NSNP top team itself for FUEL's engagement, and doubts surfaced about its agenda and goals. As 2007 came to a close, they realized that the national DoE "wasn't the place to really start the detailed work" and instead they resolved to understand the system better at the provincial level. This dynamic would repeat itself a few more times over the following years, and lead FUEL each time to focus on actors ranking lower in the official NSNP hierarchy but ones who had more discretion and influence on the ground. Each occurrence was a lesson on the importance of engagement and buy-in at all levels, albeit a costly one in terms of time, resources, and organizational stamina.

¹³⁶ "Robert Brozin, Co-Founder, Nando's, on Building a Brand with Passion," INTA Bulletin 71, no. 12, 2016. [Access here.](#)



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Iterations on the Theory of Change

Following conversations with the NSNP at the national level, FUEL got the blessing of NSNP officials to visit all nine provinces and create a more advanced version of the system map. The team therefore set out to meet “as many people in the system” as they could.

The bulk of 2008 was dedicated to creating system descriptions, and, equally important, to building relationships with NSNP officials in key provinces. The team learned that the school system in the North West Province had recently decided to change its food delivery model from a centralized co-operative system to a decentralized one, and the FUEL team offered to lend its expertise to the effort, allowing it to test the hypothesis that influencing procurement processes could help leverage improvements in meals served to students.

In early 2009, FUEL started working with the Chief Financial Officer of the North West Province’s Department of Education, who managed financial affairs for all matters related to education (including school nutrition issues) and to whom they had been introduced by the head of the Department of Education. At the time, the NSNP was being phased into secondary schools in areas with the least privileged socio-economic



makeup, which represented an opportunity for the CFO to test a different delivery model. With FUEL’s support, the CFO decided to pilot a centralized procurement model for the supply of food in secondary schools, while leaving food procurement in primary schools decentralized. There was an effort to update the provincial department’s policies and practices, with the hope that changes to policies at the top would “trickle down” into the schools’ hierarchy and compliance structures.

This top-down approach, even though focused at the provincial level, showed two weaknesses. First, working with high-ranking provincial officials did not guarantee the adoption of the updated policies at the school level because principals have much discretion over how policies are implemented. Second, when high ranking officials, like the CFO or the department head, transitioned from their positions—a common occurrence in public service—the carefully-crafted strategy and procurement processes that FUEL had spent months helping develop fell apart, leaving few allies or contacts within the North West’s NSNP team. In other words, while moving from the national NSNP level to the provincial level originally seemed like the right approach to build on-the-ground support, it became apparent that the provincial level was still too far removed from what happened in the schools and how the NSNP was getting implemented on a day-to-day basis.

This led FUEL to rethink its approach and to focus more on the program implementers on the ground. Not only did procurement and logistics require top-down approval, which could be politically fragile, but more importantly two years into FUEL’s work the team had no evidence that altering procurement systems would translate into large improvements in the share of students receiving a nutritious meal on time.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

In FUEL's words, the field workers could become "levers [they] could use in the system to influence performance at all schools."

Everyone at FUEL refers to this period as a particularly difficult one for the team. At that point, the team had been active in South Africa for close to three years, but the only signs of progress he could point to were a system description and some growing relationships—reinforcing FUEL's sense of the urgency of the problem, but not yet providing sufficient insight into how best to address it. However, aided in part by strong team relationships and unwavering financial support, the team decided to carry on, building on what they had learned from the failure of its province-level approach.

Creating a Sense of Ownership

In Luyckx's own words, it was clear from the start that government, at all levels, was the "ten-ton gorilla in the room." Campbell believed that making the government more efficient could unlock massive resources.

The initial setback in the North West not only strengthened this belief, but convinced the team that nothing could be achieved without providing stakeholders at every level with a sense of ownership.

FUEL's initial experience in the North West specifically highlighted the risk of working with high-ranking officials without creating buy-in from the schools, where the NSNP actually was implemented. Given that the program covered about twenty thousand schools, FUEL needed to identify intermediaries it could engage with at the local level. They turned to the field workers, NSNP employees in charge of monitoring and implementing the program at a school level. There were 400 to 500 field workers across the country, which represented an accessible scale for FUEL. Field workers also presented the advantage of staying in their

positions longer than politicians or political appointees, and working in teams within established governance structures that make them less disrupted by the departure of senior managers. In FUEL's words, the field workers could become "levers [they] could use in the system to influence performance at all schools."

FUEL decided to keep its focus on the North West province as they did not have a track record or strong relationships in other parts of the country, unlike the trust built in the North West. The team met with field workers in different districts of the province to better understand their experience and needs. Their plan was to accompany a small set of field workers on their visits to the schools, identify issues and bottlenecks standing in the way of meal delivery, and co-create solutions with them through workshops and brainstorming exercises. Once success had been achieved in the North West province, workshops could be conducted in other districts and provinces to train field workers and scale the approach, all while providing a sense of ownership to field workers.

Conceptualizing the MRR Approach

The team realized during its observations of the field workers' routines that there was a disconnect between the NSNP's mission to serve students a nutritious morning meal



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

and what the field workers were actually measuring. They spent most of their time looking through the school's records to investigate how many students had been fed, but they were not collecting reliable evidence on the quality, size, or timeliness of the meal. An alternative hypothesis started to form, one which revolved around accountability and measurement mechanisms. This fit well with FUEL's own focus on measurement. From the start, the team had been focused on the way FUEL's impact could be measured, which, in August 2009, had led them to commission a baseline study to observe the rates at which students were actually being fed. It also fit well with one of FUEL's recruits, Laura Poswell. Poswell had been working as a researcher for the Development Policy Research Unit at the University of Cape Town, and became a part of the team in 2007. She was instrumental in helping to design and codify what would later be known as the MRR approach.

The team came up with the concept of "non-threatening accountability," which consisted of "[pushing] people towards the right behavior, but without the punitive side." Embedded in this concept was a strong intuition that field workers needed to see value in the changes that FUEL would promote, and that there should be no negative incentives to reporting deficiencies. This



meant helping schools address their challenges and promoting improvements, rather than highlighting or much less punishing their deficiencies.

This approach led to a concept called "Monitoring, Reporting, and Responding" (MRR), which became the foundation for FUEL's engagement in the provinces. Some structural features of MRR were already in place: field workers were monitoring aspects of program implementation, but reports were mostly narrative, so they could not provide a standardized, objective measure of program implementation. Observations were difficult to compare across schools or districts, as they were not documented against consistent

standards. For example, NSNP field workers would evaluate the cleanliness of schools without having a standard baseline of kitchen cleanliness; they monitored food quantity guidelines, but recommended food quantities were often expressed in kilogram or liter units, which were hard to assess without the reliable presence of scales and measuring cups in schools.

Lastly, because current monitoring systems lacked mechanisms for field workers to provide recommendations or action steps, reports could only be used as static snapshots, rather than as dynamic working tools. If field workers were not able to identify precise issues that could make a school's



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

NSNP performance improve, or if the metrics for a school's performance were not well defined, schools did not have the ability or incentives to rectify issues. In other words, there was no clear way to define "success" and no road map to improve performance.

Testing the Approach, and Starting from Scratch Once Again

After building relationships with field workers and rebuilding relationships with the NSNP officials at the district and provincial levels, FUEL organized a series of workshops with 48 field workers. This resulted in a co-created MRR tool, with performance indicators crafted by the field workers themselves based on their experience of school visits. This newly designed MRR tool would allow field workers to assess whether schools were serving nutritious meals on time and meeting NSNP guidelines related to food groups and quantities. The indicators were also aligned with the expected outputs and outcomes as articulated in the NSNP framework.¹³⁷ The tool was designed as a physical document and manifested as printed copies of a checklist used by field workers to evaluate the schools' food preparation serving processes. Most important, it was designed to provide a framework for collaboration towards improvement, primarily at the school level, between field workers and principals. It was also designed to allow for simple aggregation

Embedded in this concept was a strong intuition that field workers needed to see value in the changes that FUEL would promote, and that there should be no negative incentives to reporting deficiencies.

of performance scores to facilitate analysis at higher levels: within school management teams, across schools and districts, and between districts.

137 The NSNP is not funded directly through the Department of Education using normal funding channels. It is funded through a Conditional Grant that is channeled from Treasury to the Department of Education on the basis of a direct cabinet decision.

Six months into this new approach FUEL experienced a new setback: in the district they had been focused on, due to an unanticipated and rare confluence of external factors, 43 field workers out of the 48 that FUEL had been working with were promoted due to department restructuring—exactly the situation FUEL had hoped to avoid by working directly with field workers. This left only half a dozen field workers trained on the MRR methodology to continue the effort.

The team was shaken by this setback. By now, they had tried working with the NSNP at a national level, then at a provincial level, and then with the field workers at a school level. FUEL had revised its theory of change from a procurement orientation to a change management approach. They had been convinced that working with field workers would be more sustainable, because historically there had been little turnover amongst that group. Though FUEL understood that this wave of field worker promotions was atypical, the setback nonetheless depleted its collective energy and enthusiasm.

FUEL was reconsidering whether they wanted to continue the program. Their initial push for impact measurement turned out to be a pivotal element for the organization. As mentioned previously, a baseline study had been conducted early on in the project to



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

assess what share of the schools in each province managed to serve a nutritious meal on time. Six months had passed since the field workers had started using the MRR approach. FUEL commissioned an assessment of about 150 schools to evaluate to what extent schools were consistently serving a nutritious meal to their students on time. The study showed that the average score of the schools had jumped by 16% in less than a year (from 55% to 65% of students). There was now hard evidence that the approach co-created with the field workers actually worked. As one team member put it: “A new tool, newly trained field workers with a slightly different approach going to the schools, is actually shifting the performance.”

Based on the results, the FUEL team went back to work. They shared the evidence with provincial officials and lobbied for the appointment of new field workers to replace the promoted group. They were very hands-on, even contributing to the writing of job specifications, which enabled them to incorporate the MRR process into each job description. The departed field workers were replaced within three months, and new appointees trained shortly thereafter. Six months after these new field workers started using the MRR approach, FUEL commissioned a third field study, and the average performance score of schools in the province had gone up by another 17%: from the baseline of 55% in 2008, it had increased

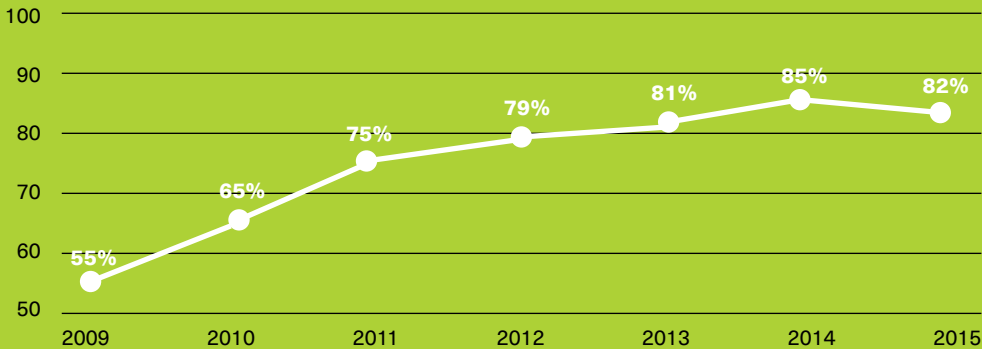
to 64% after the first intervention, and 75% after the second intervention. Figure 1 tracks the improvement from 2009 through 2015.

Embedding MRR in the System

The progress demonstrated by the consecutive field studies in the North West Province prompted FUEL to expand the approach to other provinces. Since co-creation of the MRR tool had been successful in both bringing about change and bringing people on board the project, FUEL wanted to continue experimenting with the methodology.

In parallel with its efforts in the North West, FUEL had been building relationships in all other provinces and creating system descriptions for each one. Based on these relationships, they approached the DoE leadership in each province to advocate for adoption of the MRR methodology. The first practical step towards adoption in each province was an MRR workshop, where the methodology was laid out, and a new monitoring tool co-designed. FUEL worked with attendees to co-create a local version of the MRR system, adapted to the constraints and needs of each province. Additionally, workshops helped address difficulties faced by field workers in the assessment of school performance—problems like translating quantities observed at school from volume to

Figure 1. Primary Schools serving a nutritious meal on time (%)





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Part II: **Key Themes + Insights**

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

Definition 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.

In FUEL’s case, the evidence used did not stem from academic research or externally-produced data. Rather, the evidence was collected by NSNP field workers through the MRR indicators and associated scores from the schools. The MRR system was the result of an iterative process of learning what was most (and least) salient when considering the most effective improvements to the quality of the NSNP program—reflecting the

team’s openness to modifying theories based on evidence. The MRR methodology was designed to capture the key performance metrics that FUEL identified as being crucial to gauge the success of the NSNP program. FUEL therefore supports the creation of accurate evidence and its use throughout the NSNP hierarchy to ensure that the evidence is reliable, and to make it as transparent as possible: each province’s indicators and scores are shared at quarterly progress meetings with other provinces.

Through its own efforts, FUEL learned that there are key decision makers in the school system that have enormous importance and discretion, most notably principals. Yet this has been a consistent and robust finding across many interventions in education across a diversity of settings. It is an example of how existing evidence potentially could have saved the organization both effort and resources, but did not reach decision makers in a timely and useful form.

Incentives

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.

The MRR tool was designed to fit within existing DoE incentive frameworks. Specifically, it enabled school officials to measure and achieve their job performance goals more effectively. For example, field workers were required to accurately assess a certain number of schools each month. The MRR approach allowed them to achieve their



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

The FUEL team carefully managed perceptions and intentionally gave ownership of the customized tool to those stakeholders, which facilitated broader adoption.

quantitative goals through reliable metrics, and enhanced the value they added by encouraging them to tie their observations to recommendations. Additionally, school principals now have a more accurate gauge of how well they are feeding their students, and they have clearly articulated ways to improve their score (e.g., rather than simply “increase the number of students fed,” school principals can ensure “students are fed before 10AM” and ensure “X grams of rice per student are cooked each day”). National-level officials can also monitor the progress of the NSNP program in more detail and can easily back up their progress with detailed quantitative data for each school.

One key element of FUEL’s strategy has been that the MRR tool be customized for each province through a co-design process involving the relevant stakeholders, including field workers, district and provincial officials, and principals. The FUEL team carefully managed perceptions and intentionally gave ownership of the customized tool to those stakeholders. This not only facilitated broader adoption due to the customized design, but also created personal and reputational incentives for workshop participants to ensure that the rollout of the MRR methodology happened as smoothly and as effectively as possible.

The FUEL team also made sure that the MRR logic resonated with NSNP stakeholders at every level. This all contributed to a culture of compliance and performance, driven by an alignment between the MRR methodology and NSNP actors’ eagerness to do well. In this context, the MRR tool tapped directly into a desire to improve performance.

At the provincial level, an informal accountability mechanism was built in over the years, ensuring that schools have an incentive for their nutrition efforts to be well-ranked. While there are no direct or official consequences to a poor score, progress is monitored both by district and provincial officials, with support from FUEL. If issues or lower-than-expected performance arise, a discussion takes place between the field worker and the school about ways to improve. Further discussions take place on a quarterly basis at the district and province levels, i.e., between field workers and their supervisors, or between NSNP officials from different districts in the same province. Thus, though there are no formal sanctions for poor scores, incentives to improve are strong and aligned to motivate performance at all levels.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Capacity, Resources, and Time

Few organizations provide incentives or carve out explicit time for managers to explore emerging evidence in their field. Even fewer assign staff to find relevant evidence and translate it into accessible formats for the organization. As a result, the role of preparing and sharing evidence that is timely, useful, and relevant for practitioners is sometimes explicitly played by formal intermediaries (e.g., certain think-tanks). More frequently, an actor who holds a formal role within another stakeholder group spontaneously takes on the (additional) responsibility for trying to integrate evidence, with no actor formally responsible for the process. Discovering and integrating evidence requires time, energy and funding.

FUEL benefits from steady, flexible, and reliable funding. The deeply-rooted personal and professional relationship between Charles Luyckx and Gary Campbell was also an important element, as it led Luyckx to trust FUEL's methods and ability to have an impact. While this constellation of favorable factors is unique to FUEL, a core lesson is that sustained and flexible funding is often essential for a new, evidence-informed approach to bear fruit.

Several benefits stem from FUEL's reliable, flexible, and durable sources of funding.

First, it impacts the competitiveness of FUEL's employment conditions. The organization is able to attract diverse talent and experience, and team members have a mix of private sector and community or nonprofit backgrounds. Second, it enables FUEL to have a longer time horizon than if the organization were bound to short- or medium-term timelines, ones typically associated with philanthropic or government grants. This longer timeline, combined with the deep trust existing between Luyckx and the FUEL team, has created an environment where the periodic failures associated with trial-and-error experimentation have been treated as temporary setbacks on a longer path to success. It played a particularly important role when FUEL encountered setbacks early in its history. An organization with less stable and trusting funding could not have taken three years to build relationships, refine its theory of change, and demonstrate its impact. The flexibility also allows FUEL to cultivate long-lasting relationships within the education system, and to build trust by demonstrating a continued commitment to the NSNP throughout the years. Third, the reliable source of funding have liberated FUEL from having to dedicate a significant share of its resources to fundraising or mandatory reporting, and so enabled the team to

focus on programmatic activities including stakeholder engagement, workshops and trainings, and progress monitoring.¹³⁸

Other, albeit less tangible, benefits have stemmed from Charles Luyckx's heavy involvement since FUEL's early days. Luyckx has said that he, and by extension FUEL, operates under an unofficial mantra borrowed from U.S. President Truman: "there's no limit to what you can achieve if you don't care who gets the credit." This enables FUEL to give the credit to its partners at all levels of the NSNP for the progress accomplished, which propels the collaborative approach vital to success. Of course, should Charles Luyckx or the Yellowwoods Trust choose to withdraw their support to FUEL, the organization's dependence on so few sources of funding could make it vulnerable.

The team has also prioritized flexibility and adaptability by choosing to remain small, spend its resources cautiously, and leverage

¹³⁸ Note that the team has given thought to seeking compensation from the NSNP or public stakeholders in exchange for the services provided, mostly as a commitment device. This has not been implemented.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love



external resources when possible. For instance, FUEL has developed a partnership with Eric Schollar and Associates, a research agency, to run surveys in order to gauge the fidelity with which field workers use the MRR tool. While Schollar originally only conducted surveys, FUEL came to contract the service provider to train field workers, give them regular feedback, and monitor their implementation capacity. This has allowed FUEL to maintain a minimal in-house staffing level while having the flexibility to respond to fluctuating external demands.



On the NSNP side, the MRR methodology has required virtually no additional investment of government resources, while enabling district, provincial and national-level officials to demonstrate higher performance, and get credit for the efforts made. In other words, the approach was designed, explicitly, to insert itself into existing processes and structures to better leverage and multiply the effectiveness of those resources and processes. Policy-makers were engaged in developing the intervention based on their needs and those of the stakeholders closest to the issues,



while not being asked to devote any additional funding. In addition, the risk associated with implementing the MRR tool was perceived to be low: FUEL took on the risk of achieving positive results through its initial pilots, and thereby offered a way for stakeholders to demonstrate improvements in the performance of their areas of responsibility with little reputational or financial risk.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Timing

The different and often discordant timeframes within which researchers, policymakers, and implementers operate often hobble efforts to coordinate, let alone collaborate, on evidence-informed approaches. Electoral cycles and political windows differ from NGO funding cycles and from academic publishing rhythms. Yet each actor is bound by the timeframes of her formal stakeholder group.

A particularity of the FUEL case is the long-term approach that the team was able to take from the onset. This has been tied to its sources of funding. It was also tied to the team's patience and commitment, which stemmed from the personalities at play and was embedded into the organization's culture.

The organization's timing was not, however, always aligned with that of its stakeholders, and in particular with political actors. With so many key government players rotating so frequently, FUEL learned the hard way that it was crucial to engage with actors operating on longer timelines and less tied to political objectives and calendars (though even here FUEL faced a major challenge when so many field officers departed). This lesson is relevant for other organizations engaging with political actors, who are more likely than career personnel and field-level staff to go through

With so many key government players rotating so frequently, FUEL learned the hard way that it was crucial to engage with actors operating on longer timelines and less tied to political objectives and calendars.

pressures to gain momentum, and activity slowdowns, depending on election cycles and political trends.

Learning from Failure

Potential consequences for risk-taking and experimentation with innovative approaches are generally seen as negative and dissuade the exploration of novel, evidence-informed interventions. Fear of failure can further hinder the incorporation of novel evidence into practice, even when stakeholders recognize the value and applicability of the evidence.

FUEL encountered several setbacks in developing and rolling out its approach. As discussed above, when the organization was created, it started building relationships with high-ranking officials within the DoE and NSNP. While this approach initially enabled FUEL to get traction and buy-in from top levels, the team found itself having to "start from scratch" when a reorganization changed the structure of top echelons in the education system. FUEL then lost the buy-in and momentum it had developed, and no longer had support for or interest in its activities. Both challenges sparked reflection among the team and were used as learning experiences.



Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

FUEL is not as “afraid of failure” as other organizations since it has stable funding that allows for some trial-and-error, and a commitment to learn from “failures” in order to develop its methodology.

Trust, Respect, and Buy-in Among Stakeholders

The cross-stakeholder collaborations required for evidence-informed policies and practices are often difficult to initiate, develop, and sustain. Particularly when institutional incentives are lacking, personal trust, respect, and buy-in between individuals across stakeholder groups become critical to fostering the effective flow of evidence into practice.

From its start, the FUEL team focused on establishing and cultivating relationships even with people who may have initially been skeptical of their intentions. For instance, one of our interviewees at the NSNP national level told us that upon assuming her current position in 2009, she initially didn’t trust Gary Campbell or believe that his agenda was aligned with the NSNP’s best interests. However, she slowly warmed up to FUEL when hearing about the team’s track record and long-standing engagement with the education system in South Africa.

Another factor that helped build trust has been FUEL’s insistence that the MRR tool be tailored and co-designed with province and district officials as well as field workers. As a result, the MRR tool looks slightly different in each province.

Conclusion

This case study illustrates how evidence informed the design and development of South Africa’s National School Nutrition Program. FUEL established deeply collaborative relationships with national, provincial and local government leaders in order to improve the country’s flagship in-school nutrition program. By making use of the evidence that emerged from each successive stage of its pilots, FUEL re-shaped its approach, seeking to align the incentives and timeframes of the various stakeholders. Even though FUEL took an evidence-based approach, it took them several program iterations before they were able to move beyond their initial hypothesis (about how logistical challenges lay at the heart of the problem). As we have learned from other case studies in this series, integrating other stakeholders and evidence from the beginning in a truly open and flexible manner—not just to evaluate the

effectiveness of a hypothesized solution but also to reframe how the question is actually asked—might have saved FUEL time and other valuable resources.

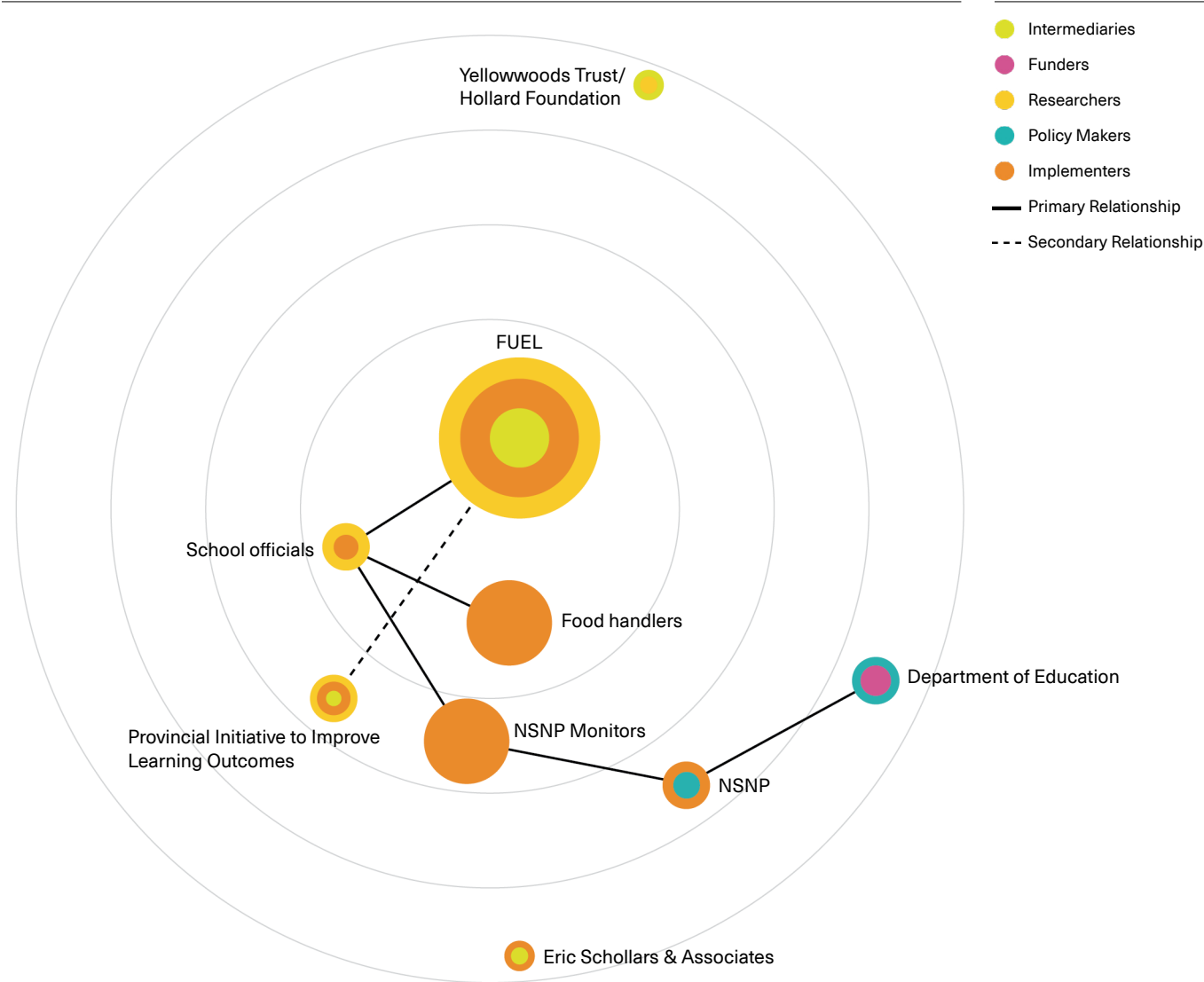
Through flexible, committed funding, FUEL was able to weather the periodic setbacks it encountered. The school nutrition program ultimately reached significant national scale, and FUEL has expanded its collaboration with the South African government into grappling with other key challenges in the education system.



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.

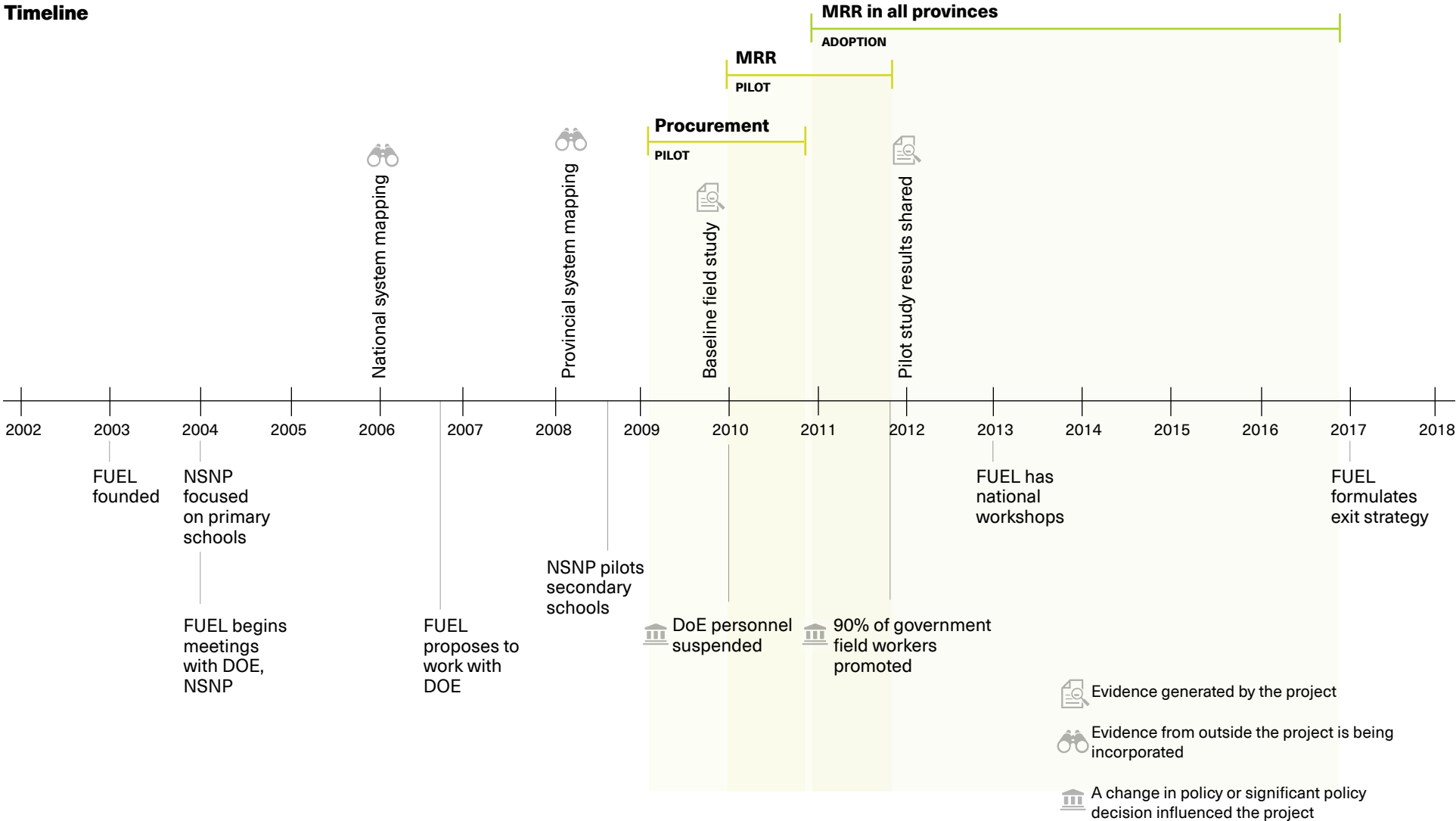
NSNP stands for National School Nutrition Programme





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Timeline





Teacher Community Assistant Initiative
Graduating the Ultra Poor
Teaching at the Right Level
Aqua+
Progresa | Oportunidades
Programa Primer Empleo
CALIE
FUEL: Feed, Uplift, Educate, Love

Process Diagram

South Africa is in the process of **scaling** MMR nationally.

