ELIMU-SOKO

THE TEACHING INNOVATION LAB

Research and Innovation Agenda

January 2026

This document is a working draft. It represents our current understanding of what works in teacher professional development (TPD) for foundational literacy and numeracy (FLN), what we are hearing from governments, and what innovators can offer. It provides background context and direction for the Teaching Innovation Lab (TIL), and additional detail for the associated Expressions of Interest (EOI) and Request for Proposals (RFP) process. It is a living document and will continue to evolve as engagement deepens across countries and partners. We welcome feedback and input as we augment this document. Please reach out to: muna.ngenda@elimu-soko.org with any comments or questions.

Table of contents

Introduction

Teacher quality is the most important school-based factor affecting student learning. The difference between a highly effective teacher and a less effective one can mean the equivalent of an additional year of schooling for students. This finding, replicated across contexts and now well established in education research, makes teacher professional development one of the highest-leverage investments a government can make.

Two decades of rigorous research have identified what effective TPD looks like: Structured pedagogy programs that provide teachers with detailed lesson plans and quality materials, combined with ongoing instructional coaching, produce effect sizes[1,2], among the largest impacts documented in education interventions. Programs in Kenya, South Africa, India, and elsewhere have demonstrated that these approaches work.

Yet foundational learning outcomes across Africa remain critically low. More than 85 percent of children in Sub-Saharan Africa complete primary school unable to read a simple text or perform basic mathematics[3]. The gap between what research shows works and what education systems deliver remains vast. A World Bank review of 33 countries found that most government TPD programs incorporate fewer than half of the features that evidence identifies as effective.

The question is no longer what works. The question is what can work on a national scale, within government budgets, using government personnel. As external funding becomes less certain, governments are asking different questions. Not "what programs can donors fund?" but "what approaches can we own, operate, and sustain?"

This shift creates both urgency and opportunity. The conditions for a different kind of partnership are present. Innovators across Africa have demonstrated that cost reduction is technically possible. Governments are asking new questions, seeking approaches they can own rather than pilots they will inherit. The Teaching Innovation Lab exists to test whether proven TPD approaches can be adapted for government delivery without losing what makes them effective.

What Works in Teacher Professional Development

The research on effective teacher development has matured considerably over the past two decades. Kenya's Tusome program, reaching 7 million learners across nearly 24,000 primary schools, demonstrated effect sizes of 0.6 standard deviations for reading outcomes[4]. A meta-analysis of 60 causal studies found that instructional coaching improves teaching quality by 0.49 standard deviations on average, with effects persisting over time[5]. The Global Education Evidence Advisory Panel identifies structured pedagogy combined with ongoing support as one of the most cost-effective interventions available[6] . Nevertheless, a review of government TPD programs across 33 countries found that most incorporate fewer than half of the features that evidence identifies as effective[7] . The gap between research and practice persists.

Effective TPD is a cycle. Teachers change practice through a predictable pattern: awareness of new approaches, initial attempts at implementation, confrontation with obstacles, refinement through feedback, and eventual integration into routine[8]. Critically, teachers are more likely to sustain new practices when they see their students respond positively; without that feedback, change rarely sticks. Training that stops at awareness, without structured opportunities to try, reflect, and adjust, rarely produces lasting change[9]. Effective training allocates approximately 50 percent of time to active practice, 20 percent to modeling, 20 percent to discussion, and only 10 percent to lecture[10].

Most government training inverts these ratios, dedicating most of the time to presentation and discussion while leaving teachers to figure out implementation on their own. This training inversion helps explain why teachers attend workshops and return to their classrooms unchanged. try it on Monday, alone in their classrooms[11].

Four mechanisms support the full learning cycle. **Structured pedagogy** provides teachers with detailed lesson plans and quality materials that make effective instruction concrete and replicable, specifying what to teach, when, and how. By reducing cognitive load and providing everything teachers need to implement new practices, structured pedagogy increases the likelihood that teachers will try what they learned in training[12].

Ongoing coaching and feedback provide regular observation, specific guidance, and sustained support to address the implementation challenges that emerge when teachers attempt new techniques in their own classrooms. A multi-country study found that it is this ongoing support, not initial training, that determines whether teachers continue to implement new practices with fidelity and impact[13].

Peer collaboration and communities of practice enable teachers to learn from each other through structured observation, joint lesson planning, and shared reflection, sustaining improvement without permanent external support.

Data-driven instructions equip teachers to assess student learning, interpret results, and adjust teaching, accordingly, targeting support where it is needed most.

Technology is an enabler across each of these mechanisms. It can extend coaching reach through video-based observation, enable peer learning at scale through messaging platforms, support structured pedagogy through digital lesson guides and audio support, and make data-driven instruction feasible through efficient assessment and visualization tools. A systematic review of 170 technology-mediated TPD studies found that hybrid approaches blending digital and face-to-face components consistently outperform purely online or purely traditional methods[14]. Programs designed for offline functionality and low-bandwidth environments achieve broader reach than those requiring reliable connectivity.

The evidence base is mature. The challenge is translating it into systems that governments can afford, scale, and sustain.

The Challenge: From Evidence to Scale

Education systems across Africa face competing demands: recruiting and deploying teachers, managing payroll, maintaining infrastructure, expanding access. Ghana allocates 67 percent of its education budget to teacher compensation, 21 percent to goods and services, and 12 percent to capital investments, leaving minimal fiscal space for dedicated teacher professional development15. Similar patterns are held across the continent. Quality improvement competes with basic operational requirements.

Effective TPD approaches often cost US$100 to $300 per teacher annually when implemented with full external support16. A country with 200,000 primary teachers would need $20 to $60 million annually for comprehensive TPD, amounts that dwarf current allocations. These cost structures place proven approaches beyond what most governments can sustain on a national scale.

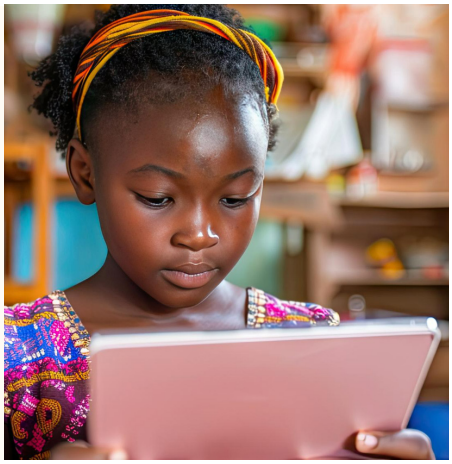
For decades, Official Development Assistance filled this gap. Bilateral and multilateral funding supported multi-year TPD programs, often with dedicated coaching cadres, intensive training models, and external technical assistance. As ODA to education contracts, that funding is no longer assured. The programs it supported, many of which operated parallel to government systems, were rarely designed for transition to domestic financing.

Cost is only part of the challenge. The cascade model, in which national trainers train regional trainers who train district officers who train teachers, loses fidelity at each level. By the time content reaches classroom teachers, it often bears little resemblance to the original design. Programs that produce strong effects under controlled conditions frequently weaken when delivered through government structures with less training, higher ratios, and competing priorities.

Historically, sophisticated TPD programs operated parallel to government systems and were rarely designed for sustainable scaling within domestic budget constraints. Programs that co-design with ministries from the outset, use government personnel and resources, and keep costs aligned with domestic financing levels are more likely to sustain practice change once external support ends. The opportunity now is to strengthen existing structures. Governments already operate teacher cluster meetings, inspection systems, pre-service training colleges, and data collection routines. The question is whether proven approaches can be embedded within these structures, and what it costs to transition from external delivery to government ownership.

This transition cost is often underestimated. Moving from a pilot to a national scale requires more than replication. It requires building government capacity, adapting to system constraints, and sustaining quality as delivery shifts from external implementers to ministry personnel. Understanding the cost of change, and not only the cost of delivery, is essential for realistic scale-up planning.

Innovators across Africa are already testing approaches that could bridge this gap. Organizations like Tari in Kenya, STIR Education in Uganda, and Eneza Education across multiple countries have demonstrated that cost reduction is technically possible, delivering teacher support through basic mobile phones, existing government meeting structures, and SMS-based platforms. A companion Innovator Landscape Report profiles the organizations working in this space and maps the range of approaches currently being tested across the continent.



What Governments Are Asking For

Engagement with education ministries across multiple countries, alongside perspectives from more than thirty innovators, implementers, and technical organizations, has surfaced consistent themes. These conversations are ongoing, and the themes outlined here will continue to evolve as engagement deepens across countries and partners.

The message is clear: countries do not need another large, generic TPD program. They want sustainable mechanisms, tools, routines, and system roles that strengthen everyday interactions between teachers, school leaders, inspectors, and learners. These priorities reflect hard-won experience.

Many governments have hosted externally funded TPD programs that showed promising results during implementation but could not be sustained when funding ended. They have seen innovations that worked in pilot conditions fail when extended to rural areas or lower-capacity districts.

Innovation Themes

Four themes have emerged from government engagements, a review of the evidence base, and input from the Lab's technical advisors. Each represents an initial hypothesis about where innovation can address the gap between what systems have and what they need to deliver effective teacher support at scale. These themes will

form the basis of the first call for proposals the Lab will launch. The list will grow as the Lab engages with additional governments and learns from early pilots.

Theme 1: Individual Teacher Support

Core problem: Teachers lack practical, immediately usable tools for daily instruction in foundational literacy and numeracy. Systems emphasize workshops and compliance but fail to provide actionable guidance teachers can use during school hours. When training ends, teachers return to classrooms without ongoing support to translate what they learned into daily practice. Early grade reading and mathematics require consistent, structured routines that are built sequentially, yet teachers often lack the supports to deliver these with fidelity[17]. Lesson planning remains compliance-oriented rather than diagnostic. Available tools, where they exist, often ignore low-connectivity environments and the time constraints teachers face.

Rationale: Closing the feedback loop between student learning gaps and teacher practice gaps is central to improving early grade outcomes. Embedding improvement routines directly into the school day turn awareness into behavior change. Offline-first supports increased feasibility in low-connectivity contexts. Tools that fit within teachers' preparation time, such as micro-guides, short videos modeling practice in local languages, and simple digital aids, offer a path to sustainable support without requiring additional travel or workshops.

Hypothesis: Practical, low-cost, offline-first tools that fit within teachers' daily routines can sustain and reinforce structured pedagogy between formal training events and coaching visits, extending the impact of existing TPD investments at marginal additional cost.

Theme 2: Group Coaching and Peer Learning

Core problem: Traditional cascade training is too slow, too costly, and too fragile to sustain improvements at scale. Most African education systems include some form of teacher cluster meeting, professional learning community, or subject-based network. The infrastructure exists. Tanzania has MEWAKA, Kenya has School-Based Teacher Support (SBTS), Nigeria has PLC frameworks. But in practice, many cluster meetings lack clear protocols, facilitation support, or connection to instructional improvement. Attendance is inconsistent. Discussions drift into administrative matters. Impact on teaching practice is limited

Rationale: Teachers learn effectively from peers when collaboration is structured and focused on practice. Peer-led models build collective efficacy and reduce isolation. Group coaching cycles can institutionalize practice change using existing cadres such as head teachers and inspectors. Evidence across African contexts shows that ongoing, proximate support outperforms one-off workshops in improving teaching quality and student outcomes. Group formats are also cost-effective when protocolized and supported by low-cost digital tools[18].

Hypothesis: Digitally enabled, peer-led coaching models with clear protocols, light facilitation, and structured reinforcement can deliver sustained practice change at a fraction of the cost of traditional cascade training, using existing government cluster structures.

Theme 3: Closing the Data-to-Action Loop

Core problem: Where Theme 1 addresses tools teachers use directly, this theme focuses on how supervisors and school leaders use data to target and prioritize teacher support. Ministries collect vast amounts of data from classroom observations, student assessments, attendance records, inspection reports, and school improvement plans. Very little of it is used to guide teacher support. Middle-tier actors, including inspectors, curriculum support officers, district planners, and school heads, describe four persistent gaps: data quality is poor, with records incomplete, inconsistent and often inflated due to compliance pressures; data interpretation is limited, with insufficient time and training to analyze patterns; action routines are unclear, with no structures for follow-up after data is collected; and role clarity is lacking, with overlapping responsibilities across decentralized levels. The middle tier is where this bottleneck is most acute. But with limited coaching density, solutions must also equip school heads and other proximate actors to translate data into targeted teacher support. In Senegal, data from MOHEBS bilingual programs and inspector observations exists but is seldom used. Nigeria, Kenya, Tanzania, and Ghana report similar challenges.

Hypothesis: Simple tools and routines, such as dashboards, PLC agendas, or feedback routines, that strengthen data reliability and help middle-tier actors and school heads translate it into where support is needed can improve instructional quality more cost-effectively than generating new data or building new systems. The key test is: Could a ministry plausibly issue this as guidance, train staff once, and expect it to be used repeatedly without the innovator's presence?

Rationale: Data only improves outcomes when it connects to action. Strengthening this connection requires attention to data quality, ensuring that what is collected is reliable and accurate, and to data use, ensuring that those who collect it have the tools, time, and routines to translate it into teacher support[19]. When school directors, inspectors, and district planners can see patterns in observation data and respond with targeted feedback, data collection becomes meaningful. Technical capacity exists. What is missing is a system routine that closes the loop between data and support, with clear and comprehensive workflows, decision rules and responsibilities that account for and align the roles, incentives, and operating environments of both middle-tier and learner-facing actors. The Teaching Innovation Lab builds on existing middle-tier implementation research[20] by piloting closed-loop feedback mechanisms that enable multiple layers of the education system to act on existing data and use it to guide teacher support. The Lab is interested in low-cost analytic tools, automated decision and action prompts, real-time dashboards, and feedback and accountability routines that help school heads, inspectors, and curriculum support officers translate information into specific coaching priorities and peer-learning agendas within existing government structures. We are explicitly interested in low-friction workflows that

replace or simplify existing routines, rather than adding parallel processes. In practice, many innovations under this window will enable and strengthen middle-tier work, but they must do so by re-engineering how data moves and is acted upon, not only by changing behavior through coaching or incentives.

Theme 4: Finding and Spreading What Already Works

Core problem: : Within every country, there are teachers and district officers achieving strong foundational learning outcomes despite operating under the same constraints as their peers. They have developed practices that work in their context, with their students, using available resources. But systems rarely identify and document these exemplary practices or create pathways for their diffusion. The challenge is not a lack of good practice — it is the absence of system mechanisms to surface, validate, spread, and sustain those practices. Teacher motivation suffers when there is no recognition for excellence and no pathway for professional contribution beyond the classroom. In the context of the Teaching Innovation Lab, positive deviance refers to observable, repeatable practices or routines used by system actors that lead to better instructional quality or learning outcomes, despite operating within the same policies, resources, and constraints as others

Rationale: Positive deviance methodology offers a concrete approach: identify outlier results through data or peer nomination, document supporting practices in local languages, and spread these through existing peer structures. This approach aligns evidence of use, peer learning, and scalable support. It builds on what already exists rather than introducing external models. When teachers see colleagues in similar circumstances succeeding, credibility increases[21]. Recognition strengthens professional identity. The payoff is improved practice and sustained engagement, as teachers see pathways to contribution and growth within their own systems. In this theme, positive deviance refers to realistic, replicable instructional practices demonstrated in typical classrooms operating under normal system constraints. This window explicitly focuses on practice, not personality, charisma, or exceptional talent. The Lab is not seeking exceptional "top 10%" teachers, heroic individual effort, or practices that depend on unusual resources. Instead, we are interested in practices that can plausibly be adopted by many teachers working in similar conditions.

Innovations under this window may focus on identifying positive deviance among:

1. Teachers

Examples of qualifying teacher-level positive deviance:

- Teachers whose students consistently outperform peers on foundational literacy or numeracy without additional resources
- Teachers who:
  - ◊ Use time differently (e.g. more active practice, better transitions)
  - ◊ Adapt lesson plans effectively while maintaining curriculum alignment
  - ◊ Use formative assessment informally but systematically
  - ◊ Maintain higher engagement in large or multigrade classes
- Teachers whose practices persist even without external coaching

What does not qualify:

- Teachers succeeding due to selective intake, extra tutoring, or external support
- Practices dependent on exceptional materials, technology, or unsustainable effort

1. Middle-Tier or Support Actors

Examples of qualifying teacher-level positive deviance:

Innovations may also target positive deviance among:

- ·inspectors
- Coaches
- Subject advisors
- District education officers

Examples include:

- Supervisors who provide unusually actionable feedback
- Advisors who tailor support based on teacher need rather than schedule

2. School Heads / Instructional Leader

Examples of positive deviance at the school leadership level:

- School heads who consistently improve instructional quality despite staffing shortages
- Leaders who
  - ◊ Structure PLCs or staff meetings to focus tightly on instruction
  - ◊ Create informal coaching or peer observation routines
  - ◊ Protect instructional time more effectively than peers
  - ◊ Use simple data to prioritize support without formal dashboards
- The focus here is what they do differently, not their formal authority or compliance with policy.

What Innovators Should Focus On:

- **Identification:** How will positive practices be identified using existing data, peer nomination, or supervisor nomination other system-embedded mechanism.
- **Documentation:** How practices will be captured as concrete instructional routines, lesson structures, or feedback approaches directly relevant to foundational literacy and numeracy.
- **Translation into tools:** How practices will be converted into usable formats such as micro-guides, short videos, step-by-step scripts, PLC facilitation materials, or coaching prompts, aligned to national curricula and local languages.
- **Diffusion:** How practices will spread through existing peer-learning structures, coaching routines, cluster meetings, or digital channels without creating parallel systems or relying of external facilitation

Hypothesis: Low-cost systems for identifying, documenting, and diffusing local exemplary practice can improve teaching quality while strengthening teacher's professional identity and engagement. Innovators applying under this window must design and implement systems, not one-off case studies. To maintain focus, proposals will be deprioritized if they:

- Import external "best practices" without local discovery
- Focus on individual excellence without a diffusion mechanism
- Require high incentives, awards, or promotions to motivate adoption
- Depend on high-end technology or professional media production
- Treat positive deviance as inspiration rather than instruction

These four themes represent starting points, not boundaries. As the Lab engages with additional governments, tests innovations, and learns from early pilots, new themes will emerge, and existing ones will sharpen. The hypotheses will be refined through evidence. What remains constant is the focus: identifying innovations that can deliver effective teacher support at costs governments can sustain.

Across all themes, two principles are paramount. Innovations must be designed with equity at the center, reaching marginalized groups through localized materials, attention to rural access, and scheduling that works for all teachers. Reaching underserved populations often increases costs, and these trade-offs must be made transparent. Innovations must also respect teacher welfare, fitting within school schedules rather than adding burdens, and building recognition into the design.

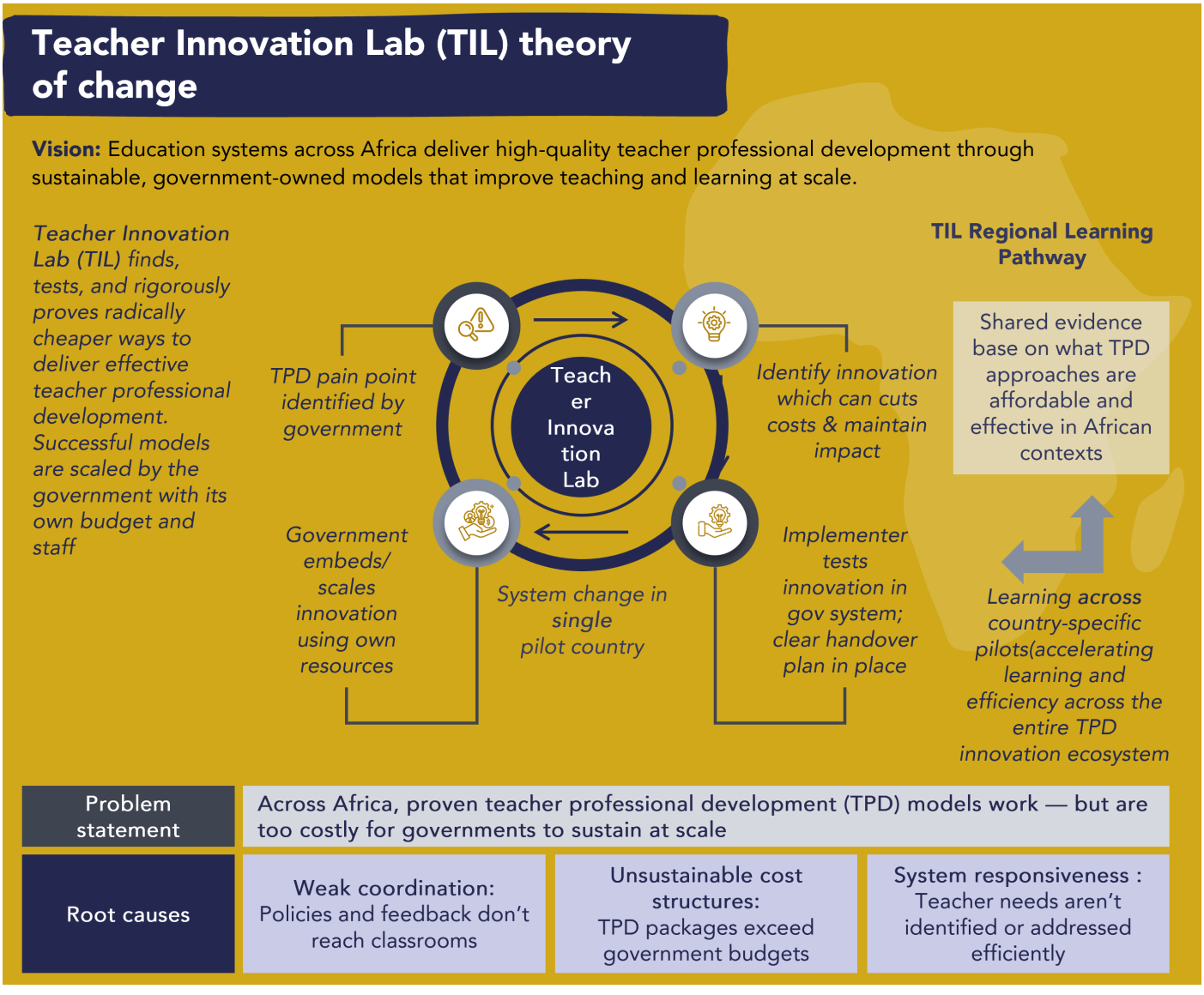
These themes are interconnected. Individual teacher support tools are most effective when reinforced through peer learning structures. Peer learning works best when driven by real data. Closing the data-to-action loop depends on middle-tier actors and school heads with clear roles and routines. And finding what already works feeds content back into all three: exemplary practices become the tools, the topics for PLCs, and the evidence that guides targeted support. The next section outlines how the Lab will test these hypotheses.

How the Teaching Innovation Lab Works

Theory of Change

The Lab envisions African education systems delivering high-quality teacher professional development through sustainable, government-owned models that improve teaching and learning at scale. The theory of change begins with governments identifying pain points in their teacher support systems, such as high delivery costs, weak instructional fidelity, or poor reach to rural teachers. The Lab then supports the identification and testing of innovations that address these pain points while maintaining impact. Innovations are tested within government systems, not in parallel structures, and every pilot includes a clear handover plan for government ownership. Innovators must articulate a theory of scale, not only a theory of change: how the model works within government systems, what the key cost drivers are, and what assumptions underpin affordability at national reach.

The diagram below illustrates this pathway from government-identified pain points through to system-level change.



Proven TPD approaches contain both essential components that must be preserved and adaptable components that can be modified for cost and context. Effective programs work through specific mechanisms: structured materials ensure consistent content delivery, coaching provides personalized feedback, peer learning sustains motivation, and data use enables targeting. Some of these mechanisms may be achievable through lower-cost means. Others may require specific investments that cannot be reduced without losing effectiveness. The challenge is identifying which is which. By testing variations in delivery while measuring both fidelity and outcomes, the Lab generates evidence on what preserves effectiveness at lower cost.

Once proven effective and affordable, innovations are scaled by governments using their own personnel and budgets. This leads to system-level changes where TPD approaches are no longer donor-dependent but fully institutionalized. The process is iterative and adaptive, with failed innovations terminated quickly and lessons documented to prevent resource waste.

As pilots are implemented across different countries, the Lab facilitates cross-country synthesis of evidence, accelerating learning and reducing duplication. Governments can make investment decisions based on peer evidence rather than repeating trial and error. Over time, this builds a shared evidence base on what works, under what conditions, and at what cost. Key actors include governments, who define priorities, host pilots, and scale successful models; innovators, who adapt and test models for affordability and system fit; and regional partners, who support knowledge transfer and policy alignment.

How Pilots Work

- The Lab funds six-to-twelve-month pilot implementations, in partnership with Ministries of Education.
- Each pilot will reach a minimum of 200 teachers across diverse school contexts, including urban, peri-urban, and rural environments.
- All pilots are embedded within government systems rather than parallel structures and apply quasi-experimental designs with baseline and endline data collection.
- Measurement goes beyond whether an intervention works. The Lab tracks teacher practice, student learning outcomes, and implementation fidelity.

It also examines variation within treatment groups. When some schools show strong effects and others show weak effects under the same intervention, understanding the difference reveals what implementation factors are essential and what can be adapted for context or cost.

Cost analysis is central to every pilot. The Lab tracks three categories: setup costs, meaning one-time investments in content, technology, and training of trainers; delivery costs, meaning ongoing expenses for personnel, materials, and logistics; and transition costs, meaning the investments required to shift from external to government delivery.

Understanding the cost of change, not just the cost of delivery, is essential for realistic scale-up planning. Cost data will need to be framed in government-relevant units, such as cost per learner or incremental cost to existing TPD budget lines, to support realistic adoption conversations.

Scalability is assessed across multiple dimensions: technical feasibility, institutional compatibility, political viability, and financial sustainability. These assessments draw on frameworks from the Management Systems International Scaling Community of Practice and the Brookings Institution.

Ultimately, cost reduction alone is not the goal. The goal is to identify approaches that maintain effectiveness at price points governments can sustain. The Lab establishes minimum effectiveness thresholds below which scale-up is not recommended, regardless of cost savings.

An innovation that cuts costs by eliminating what produces impact is not a success. Some components are essential: remove them and the intervention fails. Others are adaptable: they can be modified without losing impact. Distinguishing between the two is the Lab's central contribution.

What the Lab Does Not Do

The Lab is a testing and learning mechanism, not an implementing organization. It does not develop curriculum, implement programs directly, or manage schools. It does not replace government systems but works within them. It does not guarantee scale-up funding but provides evidence for scale-up decisions.

Risk and Limitations

Several factors could limit the Lab's ability to generate useful evidence or contribute to sustainable scale. Simplifications necessary for affordability may reduce effectiveness below meaningful thresholds, as South Africa's experience with virtual coaching illustrates cost savings were achieved, but with smaller effects and some unintended consequences[22]. The Lab is designed to identify these trade-offs explicitly rather than assume cost reduction and quality maintenance are simultaneously achievable. At the same time, approaches that work in one country may not be applicable to others with different infrastructure, languages, or system structures, and pilots across diverse contexts can only partially address this challenge by identifying which innovations are robust to variation and which require significant adaptation.

Measurement presents its own constraints. Six-to-twelve-month pilots provide useful signals about feasibility and early effects but cannot demonstrate long-term sustainability or system-wide impact. The Lab generates evidence for decision-making, not definitive proof, and longer-term follow-up will be essential to understand whether pilot effects persist. Similarly, innovations that rely on connectivity, devices, or digital literacy may work in pilot conditions but fail when extended to lower-resource settings. The Lab requires innovations to demonstrate offline functionality and alternative delivery modes, yet infrastructure constraints remain a fundamental challenge in many contexts.

Perhaps the most significant risk is pilot purgatory. The education sector is littered with promising pilots that never achieved scale, and generating evidence is necessary but not sufficient for adoption. The Lab prioritizes innovations with clear government champions and realistic pathways to scale, yet closing the gap between evidence and adoption requires more than evidence alone. Government capacity constraints, changes in leadership, and competing priorities can derail promising initiatives even when the evidence is strong. The Lab works within existing government structures and builds handover planning into every pilot, but sustained political commitment remains essential for success at scale.

Conclusion

The Teaching Innovation Lab exists to address a specific challenge: translating what works in teacher professional development into approaches that African governments can afford, deliver, and sustain. The evidence base is mature. The gap is in implementation at scale.

The stakes are significant. More than 85 percent of children in Sub-Saharan Africa complete primary school, unable to read a simple text or perform basic mathematics. Every year this problem goes unsolved; another cohort passes through schools without acquiring foundational skills. The cost is borne by children, by families, and by economies that cannot afford to leave human potential unrealized.

The themes identified here reflect current government priorities and evidence gaps: individual teacher support; group coaching and peer learning; data, assessment and use of evidence; and finding and spreading what already works. These themes will evolve as pilots generate findings, and new challenges emerge. Some innovations will prove effective and affordable. Others will reveal that certain investments cannot be reduced without losing their impact. Both findings are valuable for governments making resource allocation decisions.

Success will be measured not by pilot results alone but by government adoption. The ultimate test is whether ministries of education scale proven approaches using their own systems and budgets. The Lab invites partnership: from governments seeking sustainable solutions, from innovators willing to adapt their approaches for system integration, from funders interested in evidence-based investment, and from researchers committed to understanding what enables scale. The challenge of teacher professional development in Africa will not be solved by any single actor. Progress requires collaboration, evidence, and the willingness to learn from both success and failure.

End notes

[1] Piper, B., Destefano, J., Kinyanjui, E. M., & Ong'ele, S. (2018). Scaling up successfully: Lessons from Kenya's Tusome national literacy program. Journal of Educational Change, 19(3), 293-321.

[2] Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence.

[3] World Bank (2019/2022). "Learning Poverty" briefs

[4] Piper, B., Destefano, J., Kinyanjui, E. M., & Ong'ele, S. (2018). Scaling up successfully: Lessons from Kenya's Tusome national literacy program. Journal of Educational Change,

[5] Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence.

[6] Global Education Evidence Advisory Panel. (2023). Cost-effective approaches to improve global learning (Smart Buys). World Bank & FCDO.

[7] Popova, A., Evans, D. K., Breeding, M. E., & Arancibia, V. (2022). Teacher professional development around the world: The gap between evidence and practice

[8] Guskey, T. R. (1986). Staff development and the process of teacher change.

[9] Mejia, J. (2021). Teacher professional development: Teacher training. RTI International.

[10] Mejia, J. (2021). Teacher professional development: Teacher training. RTI International.

[11] Mejia, J. (2021). Teacher professional development: Teacher training. RTI International.

[12] Mejia, J. (2021). Teacher professional development: Teacher training. RTI International.

[13] Piper, B., Mejia, J., & Betts, K. (2020). Do's and don'ts of improving teaching through instructional support: Findings from a multi-country study of coaching and communities of practice. CIES.

[14] Hennessy, S., D'Angelo, S., McIntyre, N., et al. (2022). Technology use for teacher professional development in low- and middle-income countries: A systematic review.

[15] UNICEF. (2023). Ghana Education Budget Brief.

[16] Estimates based on program cost data from Tusome (Kenya), PRIMR (Kenya), and Funda Wande (South Africa). See Evans, D. K. & Popova, A. (2016). What really works to improve learning in developing countries?

[17] Mejia, J. (2021). Teacher professional development: Teacher training. RTI International.

[18] Tachie, S. A. (2022). The impact of peer collaboration on teacher confidence and classroom practice in South African schools African Journal of Teacher Education, 11(2), 45–62. https://doi.org/10.21083/ajote.v11i2.6789

[19] Bell, S. (2025). Senegal middle tier study report. Science of Teaching. https://scienceofteaching.site/wp-content/uploads/2025/08/Senegal-Middle-Tier-Study-Report.pdf

[20] Such as the Brink Foundation and What Works Hub for Global Educations partnership on Implementation Research at the Middle Tier, supported by the Gates Foundation

[21] Marsh, D. R., Schroeder, D. G., Dearden, K. A., Sternin, J., & Sternin, M. (2004). The power of positive deviance. BMJ, 329(7475), 1177–1179.

[22] Cilliers, J., Fleisch, B., Kotzé, J., Mohohlwane, N., Taylor, S., & Thulare, T. (2020). Can virtual replace in-person coaching? Experimental evidence on teacher professional development and student learning in South Africa RISE Working Paper Series No. 20/050). Research on Improving Systems of Education (RISE). https://doi.org/10.35489/BSG-RISE-WP_2020/050